Properties of Inclusive $B \to \psi$ Production

A thesis presented
by

Daniel Young-Joon Kim

to
The Department of Physics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Physics

Harvard University
Cambridge, Massachusetts
January 2002

©2002 - Daniel Young-Joon Kim

All rights reserved.

Thesis advisors
George Brandenburg
Richard Wilson

Author
Daniel Young-Joon Kim

Properties of Inclusive $B \to \psi$ Production

Abstract

Using the combined CLEO-II and CLEO-II.V data sets of 9.1 fb^{-1} at the $\Upsilon(4S)$, we measure properties of ψ mesons produced from decays of the B meson. ("B" denotes any of B^+, B^-, B^0, or $\bar{B}^0$; and "ψ" denotes either $J/\psi(1S)$ or $\psi(2S)$.)

First, we measure the inclusive branching fractions $\mathcal{B}(B \to \psi X)$, where X is any final state. This requires subtracting the contribution of $\psi(1S)$ produced in continuum (non-$B\overline{B}$) processes. We also derive $\mathcal{B}(B \to \psi(\text{direct})X)$ by subtracting the contributions of $\psi(1S)$ with intermediate parents $\psi(2S)$ or χ_{c1}. Our branching fraction results are consistent with and significantly improve upon previous CLEO measurements, and are competitive with recent measurements from BaBar.

Second, we obtain momentum distributions for $\psi(1S)$ and $\psi(2S)$ produced directly from B decays. The momentum distributions are corrected for measurement smearing, detection efficiency, and continuum $\psi(1S)$ production. The results do not support recent speculations of a narrow bump at $p_{\psi(1S)} \approx 500$ MeV.

Third, the ψ dilepton decay modes are used to measure α, the parameter describing the average polarization of directly produced ψ. We obtain $\alpha_{\psi(1S)} = -0.30 \pm 0.07 \pm 0.04$ and $\alpha_{\psi(2S)} = -0.45^{+0.22}_{-0.19} \pm 0.04$, where the first error is statistical and the second error is systematic; these results rule out the Color Evaporation Model of ψ production. Our measurements of α are the first to include only the ψ which are produced directly from B decays.

The polarization and momentum spectrum measurements have the potential to shed light on the nonperturbative parameters of Non-Relativistic QCD (NRQCD), the effective field theory which provides the current theoretical understanding of the production and decay of heavy quarkonium. However, the extraction of these parameters is currently limited by theoretical uncertainties.

Contents

Title Page i
Abstract iii
Table of Contents iv
Citations to Previously Published Work vii
Acknowledgments viii
Dedication xiv

1 Theoretical Background **1**
1.1 Motivation 1
1.2 Hadronic B Decays 2
1.2.1 Weak Decay 2
1.2.2 Inclusion of Gluons 4
1.2.3 Penguin Diagrams 6
1.2.4 Operator Product Expansion 6
1.2.5 Numerical Values of the Wilson Coefficients 7
1.3 Pre-NRQCD Models 10
1.3.1 Color Singlet Model 10
1.3.2 Color Evaporation Model 12
1.4 NRQCD 12
1.4.1 Motivation 12
1.4.2 The NRQCD Lagrangian 13
1.4.3 Factorization Formula for Inclusive ψ Production 16
1.4.4 Power Counting for State Transitions 18
1.4.5 Inclusive $b \to \psi X$ Production 19
1.5 Effects of B Hadronization 20
1.6 $B \to \psi X$ 21
1.6.1 Outline: State of Current Calculations 21
1.6.2 Inclusive Branching Fraction 22
1.6.3 Momentum Distribution 25
1.6.4 Review: Polarization Definitions 26
1.6.5 Polarization 28

2 Experimental Apparatus **31**
2.1 CESR . . . 31
2.1.1 Historical Background: 1977-1980 . . . 31
2.1.2 CESR Operation . . . 33
2.1.3 CESR Luminosity . . . 35
2.2 CLEO . . . 38
2.2.1 Overview . . . 38
2.2.2 Tracking . . . 41
2.2.3 Time of Flight Counters . . . 45
2.2.4 Time of Flight Bunchfinder . . . 48
2.2.5 Crystal Calorimeter . . . 50
2.2.6 Magnet . . . 51
2.2.7 Muon Detectors . . . 52

3 Experimental Techniques for Charmonium Measurement **55**
3.1 Obstacles to Theoretically Relevant Measurements . . . 55
3.1.1 Reference Frames: B vs. Laboratory . . . 55
3.1.2 Mismeasurement due to Bremsstrahlung and Detector Resolution . . . 57
3.1.3 Intermediate Parents . . . 58
3.2 Charmonium Reconstruction: Overview . . . 60
3.3 Bremsstrahlung Recovery . . . 61
3.4 Selection Criteria . . . 62
3.4.1 Dataset . . . 62
3.4.2 Events . . . 63
3.4.3 Tracks . . . 63
3.4.4 Calorimeter Showers . . . 64
3.4.5 $\psi \to \ell^+\ell^-$ . . . 65
3.4.6 $\psi(2S) \to \psi(1S)\pi^+\pi^-$ . . . 65
3.4.7 $\chi_{c1} \to \psi(1S)\gamma$ . . . 66
3.5 Data Yields . . . 66
3.6 Signal Monte Carlo . . . 67

4 Measurements of ψ Momentum Distributions and Inclusive Branching Fractions **78**
4.1 Analysis . . . 78
4.1.1 Analysis Overview . . . 78
4.1.2 Signal Lineshapes . . . 79
4.1.3 Invariant Mass Fits . . . 81
4.1.4 Yields, MC (Efficiency) . . . 83
4.1.5 Yields, Data . . . 84
4.1.6 Normalization; Inclusive Branching Fraction . . . 85

4.2 Systematic Error Study . . . 89
4.2.1 Overview . . . 89
4.2.2 Uncertainty in Unmeasured Modes . . . 89
4.2.3 Invariant Mass Fit Procedure . . . 90
4.2.4 Monte Carlo Low-Level Accuracy . . . 95
4.2.5 Monte Carlo High-Level Accuracy . . . 98
4.2.6 Luminosity-Related Uncertainty . . . 99
4.3 Results with Systematic Errors . . . 100
4.3.1 Inclusive Branching Fraction . . . 100
4.3.2 Momentum Distributions . . . 101
4.4 Bin Migration Correction . . . 105
4.4.1 General Results . . . 105
4.4.2 Application to Data . . . 108
4.5 Final Results, Momentum Distribution . . . 112
4.6 Discussion of Results . . . 112
4.6.1 Comparing to CLEO Monte Carlo . . . 112
4.6.2 Comparing to Other Measurements and Theory . . . 114

5 Polarization Measurement **132**
5.1 Analysis . . . 132
5.1.1 Analysis Outline . . . 132
5.1.2 Signal Lineshapes . . . 134
5.1.3 Invariant Mass Fits . . . 134
5.1.4 Continuum Correction [$\psi(1S)$ only] . . . 137
5.1.5 Illustration: $\alpha_{\psi(1S)}$ for Inclusive $B \to \psi(1S)X$ . . . 137
5.1.6 Correcting for Feed-Down, $\psi(1S)$ . . . 138
5.1.7 Fitting for α . . . 140
5.2 Systematic Error Study . . . 140
5.2.1 Overview . . . 140
5.2.2 Uncertainty in Unmeasured Modes . . . 141
5.2.3 Invariant Mass Fit Procedure . . . 142
5.2.4 Monte Carlo Accuracy, Luminosity-Related Uncertainty . . . 142
5.2.5 Procedure for Feed-Down Correction and α Extraction . . . 143
5.2.6 Results . . . 145
5.2.7 Low-Statistics $\psi(2S)$ Partitions . . . 146
5.3 Results and Discussion . . . 152
5.3.1 Cross-check: Lepton Universality . . . 152
5.3.2 Comparing to Other Measurements . . . 153
5.3.3 Conclusions . . . 153

List of References **170**

Citations to Previously Published Work

The experimental results presented in this thesis are also described in an internal note for CLEO collaborators, CBX 01-49 [72], and will be submitted to a journal for publication in February 2002.

Acknowledgments

A hundred steps from my Ithaca desk lies one of the few places on the planet that have hosted the fleeting presence of beauty quarks. This, along with the fact that this particular spot has been host to ten million B meson pairs, means that this thesis must have been written at the turn of the millennium, in an era of experimental particle physics in which the data that produce graduate theses are made possible by the ingenuity and labor of hundreds of others. If truly properly done, this section would extend for many more pages than it already does.

George Brandenburg has been a wonderful thesis advisor. His intuitive sense of how to proceed in questions of physics, statistics, and life in general saved me on numerous occasions from the ruts that inevitably occur in data analysis. With seemingly unlimited patience, he helped me to motivate myself when it seemed that there was little to get motivated about. As Director of HEPL, he has also ensured that a generation of students had excellent computing facilities and work environments, were able to go to physics conferences and schools, and could get advice and help for all manner of logistical and bureaucratic problems.

Likewise it has been a privilege to study under Richard Wilson. A member of CLEO since its first publication [1], he has inspired me with his broad scientific wisdom, the result of having been personally involved in many of the major events in the last half-century of particle physics, all backed up by an encyclopedic memory for historical detail. Despite his perpetually busy schedule as a statesman of science, he has consistently kept Ithaca (and the Harvard students working there) on his international itinerary. The list of his former PhD students is as illustrious as it is lengthy and I am honored to join it.

Many thanks to Cumrun Vafa and Masahiro Morii for rounding out the thesis committee. From his memorable teaching of the first undergraduate physics course I took at Harvard, Prof. Vafa convinced me, freshman year, that I was studying the right subject at the right place. Prof. Morii, continuing the tradition of e^+e^- physics at Harvard with a new BaBar group, agreed to be on my committee without having met me beforehand – a true sign of generosity.

I am also indebted to several CLEO collaborators who strengthened the analysis immeasurably. As chair of the paper committee for this analysis, Prof. Giancarlo

Moneti patiently read through the many drafts of the writeup – even as the page count pushed into triple digits – and provided numerous comments and suggestions. Thanks also to Prof. Roy Briere, who also volunteered to join the paper committee; the theory chapter has benefitted greatly from his scrutiny, and his incisive style of clear thinking has been an inspiration. Alexey Ershov persuaded me to pursue inclusive $B \to \psi$ as a thesis topic; his data skim and analysis code got me started on a fast track.

This thesis is based on efforts made beginning in Fall 2000. The initial 80% of my graduate career includes many other people to thank.

Hitoshi Yamamoto first offered me a summer research job the summer before senior year at Harvard. The project eventually resulted in my first publication. In between, two of his superbly taught courses opened my eyes to the beauty of particle physics. Regardless of his personal schedule, he would think nothing of spending hours patiently explaining to me insights that only come naturally to a very few: the ones who are brilliant, work incredibly hard, and do so over decades.

My contribution to CLEO software, the Time-of-Flight Bunchfinder, grew out of an early prototype program written by Ted Liu, who also was an inspiring and generous mentor during my first summer at Ithaca.

The Si3 Databoards which now do their work in the CLEO III DAQ crates were done with a great deal of collaboration and support from John Oliver, Charlie Strohman, George Brandenburg, Chris Bebek, Elliot Lipeles, and the Si3 Group.

The patrons who made this experiment (and the knowledge that comes from it) possible also deserve recognition. Particle physics research in America is funded in its entirety by the taxpayers of the United States via annual Congressional appropriations bills, which fund the Department of Energy and National Science Foundation, which in turn award research grants to HUHEPL and CLEO. I am also indebted to Harvard University, which helped support me through the Harvard National Scholarship during my twenty-one semesters at Harvard.

The first year of graduate school is special because one is 22, bright-eyed, and flattered to be receiving a paycheck to study physics. Staying up all night with friends made for some of the best memories of graduate school. Of course, the primary reason

for the gatherings was to finish (and, sometimes, begin) problem sets that were due the next morning; but in hindsight, there was a cherishable purity in our efforts to gain knowledge. Perhaps someday, a delirious 3:30 a.m. lunch will once again be had at a Boston IHOP with Robyn Madrak, Michael Temkin, and Chungsok Lee. Skiing day trips to Vermont with Marc Hill were another feature of that unique year, before we all scattered from Cambridge to join experiments.

Since arriving in Ithaca, much of my growth has taken place in the parallel universe of classical music, so I must thank members of it here here as well. Linda Case, a gifted teacher and violinist, refined my playing to new levels, transforming me from a student into a performer and teacher. Thanks also to the late Prof. Edward Murray, who truly believed in me; he made the Cornell Symphony Orchestra fertile ground for my musical growth as well as that of a generation of others. Prof. John Hsu showed me much about chamber music and historical performance practice. Prof. Steven Stucky encouraged me to discover the world of new music, and Prof. Xak Bjerken enlivened life with his artistry and zany personality. Dane Marion helped me play my best in Barnes Hall by shaping the performance atmosphere with his enthusiasm and studio acumen. Finally, thanks to the many wonderful, dedicated career musicians of the Syracuse Symphony Orchestra – among them, Greg Quick, Eric and Anita Gustafson, and Lindsay Groves – whose moving performances and genuine kindness inspired me to "take the plunge" and join them as a full-time professional.

A shared interest in music led to my meeting many wonderful friends in Ithaca as well, all of whom will surely make their mark upon society. It has been a great pleasure to get to know Lisa Leong and Chris Arrell, both gifted musicians and highly observant, thoughtful people. Thanks also to Victoria Paterson, who always understood exactly what I was talking about whenever we discussed life as a twenty-something violinist in Ithaca. Best wishes to my old chamber music colleagues Nat Miller, Tracy Decker, Victor Lin, Victor Kam, and Vic Spain. And, of course, a hearty "cheers" to the Wednesday Night Orchestra Clique – Lisa Tinanoff, Andrea Lee, Noah DeGarmo, Lauren French, and Devon Stout.

Two fellow musician-scientists need to be mentioned. Thanks to Werner Sun, CLEO office neighbor and housemate, for his innumerable droll observations, late-

night conversations, loans of hard-to-find recordings from his vast CD collection, and prompt help in ferreting out tough (well, to me) bugs in my CABS code.

Anna Herforth, a savvy botanist, fantastic oboist, and my closest friend since we met three years ago, has enriched my life immensely with her unwavering good humor, empathy, and support. I am glad to be navigating the future with her. Thanks also to Anna's parents for their kind gestures and lively conversation.

It is impossible to thank my family enough but I will make an attempt here. My father, Nam Kyun Kim, a professor of chemical engineering, helped inspire my interest in mathematics and science from a very early age. My mother, Hae Sook Kim, was my first piano teacher and has always done everything that could possibly be done to create an ideal environment for the growth of her children. Their decisions to go to graduate school in, and subsequently become citizens of, the United States, have opened many doors for me; their always wholehearted support of whatever endeavor I find interesting, has opened even more. My sister, Nancy Sunhee Kim, whose Ph.D. defense in cognitive psychology took place at Yale the day before mine, has also been an constant inspiration and support. Finally, the members of my extended family have always made their support and encouragement clear, despite geographical separation.

With my graduation Harvard will be leaving the continuously evolving CLEO collaboration, which, to quote Richard Wilson, "has been a wonderful host to Harvard for the past thirty-two years." As the last Harvard graduate student to have been educated at this productive and well-respected experiment, I submit that everyone listed in Tables 0.1 and 0.2 would agree with Prof. Wilson.

Experiment	Year	Author	Thesis
CLEO I	1981	John Haggerty	Inclusive Kaon Production at Upsilon Resonances
CLEO I	1982	Joseph Izen	High Momentum Pion and Kaon Production in the Upsilon Region
CLEO I	1985	Martin Hempstead	Aspects of Charm Production in the Decay of B mesons
CLEO I	1990	Dong Xiao	Charmless Semileptonic Decays of the B Mesons
CLEO I	1991	Jeffrey Wolinski	Study of Lambda-like Baryons in e^+e^- Annihilation
CLEO II	1995	Tiehui Liu	$D^0\overline{D}^0$ Mixing and Doubly Cabibbo Suppressed Decays
CLEO II	1995	Michael Saulnier	Measurement of Neutral B Meson Properties
CLEO II-II.V	2001	Alexey Ershov	Beauty Meson Decays to Charmonium
CLEO II-II.V	2001	Daniel Kim	Properties of Inclusive $B \to \psi$ Decays

Table 0.1: Harvard graduate students whose Ph.D. theses were based on data collected at CLEO, 1980-2001.

	Year																					
	1980+										1990+										2K	
Name (Affiliation)	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
R. Wilson (F)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
F. M. Pipkin (F)	●	●	●	●	●	●	●	●	●	●	●	●										
C. Bebek (P)	●	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
J. M. Izen (G)	●	●	●																			
J. Haggerty (G)	●	●	●																			
W. A. Loomis (F)	●	●																				
W. Tanenbaum (F)	●	●																				
R. Kline (P)	●																					
W. W. MacKay (P)		●	●	●	●																	
J. Rohlf (P-F)		●	●																			
C. Longuemare (P)		●																				
J. Hassard (P)			●	●	●	●	●															
M. Hempstead (G)			●	●	●	○	○	○	○	○												
K. Kinoshita (P-F)				●	●	●	●	●	●	●	●	●	●	●	○	○	○	○	○	○	○	
R. Giles (P)				●	●	●	●															
T. Bowcock (P)						●	●	●	●	●												
J. Wolinski (G)								●	●	●	●	●										
D. Xiao (G)								●	●	●	●	●										
M. Procario (P)								●	●	●	●	○	○	○	○							
H. Yamamoto (F)													●	●	●	●	●	●	●			
T. Liu (G)													●	●	●	●	●					
M. Saulnier (G)													●	●	●	●	●					
D. Cinabro (P)													●	●	●	●	○	○	○	○	○	○
S. Henderson (P)													●	●	●	○	○	○	○	○	○	○
G. Brandenburg (F)																●	●	●	●	●	●	●
R. Briere (P)																	●	●	●	●	○	○
Y. Gao (P)																		●	●	●	●	○
A. Ershov (G)																			●	●	●	●
D. Y.-J. Kim (G)																		●	●	●	●	●

Table 0.2: Harvard-affiliated members of CLEO, 1977-2001. Affiliations: "F" = Faculty/Professor, "P" = Postdoctoral Fellow, "G" = Graduate Student. A solid circle signifies that the person was affiliated with Harvard in the author list of the last Physical Review Letter to be published by CLEO that year. (Therefore, the indicated dates may lag the actual years of employment at Harvard.) A hollow circle signifies that the person was a member of the CLEO-CESR collaboration for that year but with another university, indicating a continuing contribution to the experiment.

For physicists who admire music,
musicians who admire physics,
and everyone in between.

Chapter 1

Theoretical Background

1.1 Motivation

The Standard Model is comprised of Quantum Chromodynamics (QCD), based on the gauge group SU(3), and the SU(2)×U(1) standard electroweak model. Together these theories describe three of the four fundamental forces and have enjoyed phenomenal success. At the time of this writing, experimental efforts to measure processes disallowed by the Standard Model invariably result in upper limits, while precision measurements of fundamental parameters invariably match Standard Model calculations to high precision – often to three, four, even eight (in the case of the anomalous magnetic moment of the electron) digits.

The field of B physics is now at a point where experimental sensitivity is fast approaching the range in which non-Standard Model effects have a reasonable chance of being observed. The observation of $b \to s\gamma$ at CLEO, the measurements of $\sin 2\beta$ at BaBar and Belle, and the prospect of much greater experimental sensitivities at future experiments such as BTeV and LHC-B, all are motivated by a desire to search for inconsistencies in the Standard Model which would provide insight on what lies beyond.

As measurements become increasingly precise, the unfortunate fact that b quarks are only observed as color-neutral bound states takes on greater importance. While calculable in principle from full QCD, a fully quantitative understanding of nonper-

turbative gluons has proven to be difficult to attain. Current theoretical technologies consist of effective field theories (HQET, ChPT, NRQCD), lattice QCD and QCD sum rules, and process-specific models (Isgur-Wise, BSW, ACCMM.) Calculations based on the first two categories are preferable because the degree to which they differ from full QCD is, in principle, calculable. Unfortunately, these approaches are only applicable to a subset of the measurable properties of B decays; they do not apply to a large class of phenomena, for example, exclusive hadronic weak decays.

This thesis is a study of the inclusive decay $B \to \psi X$, where ψ refers to either the $J/\psi(1S)$ or $\psi(2S)$, and X is any final state. Experimentally, we measure the inclusive branching fraction, momentum distribution, and polarization of the ψ. Phenomenologically, the decay $b \to \psi X$ is calculable in the non-relativistic QCD (NRQCD) effective field theory framework. By comparing measurement to theory we can provide constraints on the NRQCD matrix elements, which apply universally to all methods of ψ production.

1.2 Hadronic B Decays

This section is intended to be a brief review. For further reading, Ref. [2] provides comprehensive details.

1.2.1 Weak Decay

Decays of the B meson are governed by the weak charged current:

$$\mathcal{L}_{CC} = -\frac{g}{2\sqrt{2}} J_\mu W^\mu \quad (+h.c.) \tag{1.1}$$

where W^μ is the W-boson field, and the left-handed fermion fields, represented by four-component Dirac bispinors, couple via the current

$$J^\mu = (\overline{\nu}_e \;\; \overline{\nu}_\mu \;\; \overline{\nu}_\tau)\, \gamma_\mu\, (1-\gamma_5) \begin{pmatrix} e_L \\ \mu_L \\ \tau_L \end{pmatrix} + (\overline{u}_L \;\; \overline{c}_L \;\; \overline{t}_L)\, \gamma_\mu\, (1-\gamma_5)\, V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} \tag{1.2}$$

with the mass and weak eigenstates of flavor related by the unitary matrix V_{CKM}:

$$V_{CKM} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.3}$$

$$\approx \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix} \tag{1.4}$$

The approximation shown in Equation 1.4, where $\lambda = \sin\theta_C \approx 0.22$, is useful in understanding the intergenerational coupling strength (suppressed by λ^2 or λ^3 for b decays, thus explaining the relatively long B lifetime) but ignores the complex phases in the CKM which lead, for example, to CP violation.

The quark level process behind $B \to \psi X$ is $b \to c\overline{c}q$, (where $q \in \{s, d\}$), which proceeds primarily through the "tree" Feynman diagram in Figure 1.1(a). The matrix element is given by

$$\mathcal{M} = i\frac{g^2}{8} \underbrace{V_{cb}V_{cq}^*}_{CKM} \underbrace{[\overline{c}\gamma_\mu(1-\gamma_5)b]}_{b\to cW} \underbrace{\frac{g^{\mu\nu} - \frac{p^\mu p^\nu}{m_W^2}}{p^2 - m_W^2}}_{W} \underbrace{[\overline{q}\gamma_\nu(1-\gamma_5)c]}_{W\to\overline{c}q} \tag{1.5}$$

where the operators b, c, and q annihilate a quark (or create an antiquark) of the specified flavor, and p^μ is the 4-momentum of the W propagator.

The kinematics of $b \to cW$ constrain $p^2 \sim m_b^2 << m_W^2$. Since $p^2/m_W^2 << 1$, we may expand the denominator of the W propagator term to obtain

$$\mathcal{M} = i\frac{g^2}{8} \underbrace{\frac{1}{m_W^2}}_{W} \underbrace{V_{cb}V_{cq}^*}_{CKM} \underbrace{[\overline{c}\gamma_\mu(1-\gamma_5)b]}_{b\to cW} \underbrace{[\overline{q}\gamma^\mu(1-\gamma_5)c]}_{W\to\overline{c}q} \underbrace{\left(1 + \frac{p^2}{m_W^2} + \frac{p^4}{m_W^4} + \ldots\right)}_{W} \tag{1.6}$$

The leading factor of $\frac{g^2}{8m_W^2}$ is the Fermi decay constant $\frac{G_F}{\sqrt{2}}$. So, when ignoring $\mathcal{O}\left(\frac{p^2}{m_W^2}\right)$ terms, Equation 1.6 is the effective four-fermion interaction, depicted in Figure 1.1(b) and defined by the following effective Hamiltonian:

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{cb}V_{cq}^* [\overline{c}\gamma_\mu(1-\gamma_5)b][\overline{q}\gamma^\mu(1-\gamma_5)c] \tag{1.7}$$

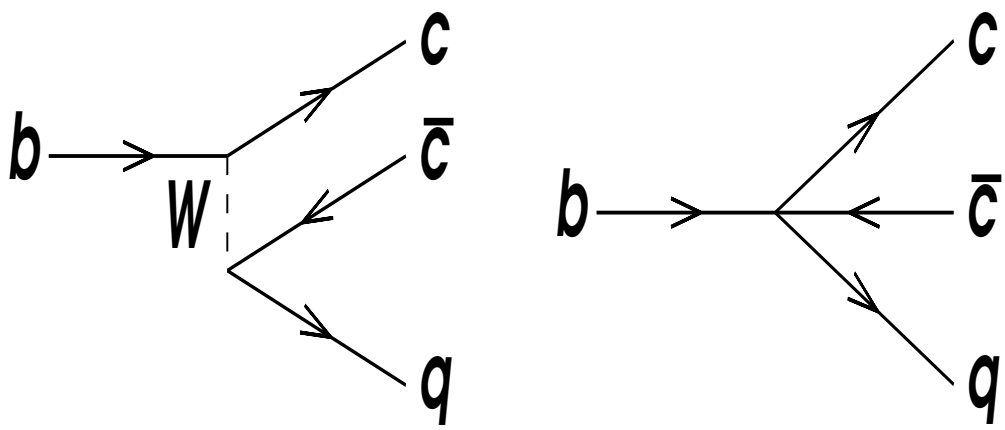

Figure 1.1: The dominant tree-level production mechanism for $b \to c\overline{c}q$, where q is either s or d. (a) Left: full theory, with W propagator. (b) Right: the effective four-quark interaction.

Explicitly adding the color indices to the quark creation and annihilation operators in Equation 1.7 yields

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{cb} V_{cq}^* Q_2 \tag{1.8}$$

where

$$Q_2 \equiv [\overline{c}_i \gamma_\mu (1-\gamma_5) b_i][\overline{q}_j \gamma^\mu (1-\gamma_5) c_j] \tag{1.9}$$

where $i, j \in \{1, 2, 3\}$. The W propagator carries no color charge, so the quarks that it produces ($\overline{c}q$) have the same color index. If we naively assume that production of a ψ bound state occurs only if the c and $\overline{c}$ are created in a color singlet state ($i = j$), then there is a suppression factor of 3, relative to decays in which the W quark pair hadronizes (no restrictions on i or j.) This is why $b \to \psi$ production is said to be "color-suppressed."

1.2.2 Inclusion of Gluons

Up to now the dynamical role of gluons has been ignored. We now include hard (short-distance) gluons with energies in the range $\mathcal{O}(M_B) < E_g < \mathcal{O}(M_W)$, where the lower limit must be greater than $\Lambda_{QCD} \sim 500$ MeV in order to enable a perturbative solution, and the upper limit exists because we use the 4-fermion vertex. When

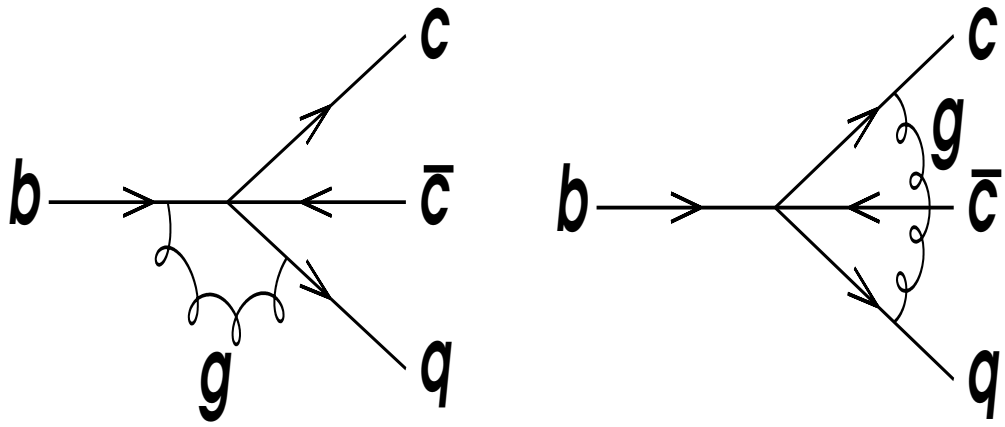

Figure 1.2: Examples of gluonic interactions which modify the tree-level effective four-quark interaction.

including these gluons, the effective Hamiltonian becomes

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{cb} V_{cq}^* (C_1 Q_1 + C_2 Q_2) \tag{1.10}$$

where a new operator Q_1, includes diagrams (such as the ones in Figure 1.2) in which hard gluons mix the color indices:

$$Q_1 \equiv [\overline{c}_i \gamma_\mu (1-\gamma_5) b_j][\overline{q}_j \gamma^\mu (1-\gamma_5) c_i], \tag{1.11}$$

and C_1 and C_2, the Wilson coefficients, encode all of the short-distance physics, including hard gluons.

Since we are interested in ψ production, we choose an operator basis which separates the $c\overline{c}$ color singlet and $c\overline{c}$ color octet contributions, as follows:

$$Q_{(1)} \equiv \frac{1}{3}[\overline{c}\gamma_\mu(1-\gamma_5)c][\overline{q}\gamma^\mu(1-\gamma_5)b] \tag{1.12}$$

$$Q_{(8)} \equiv [\overline{c}\gamma_\mu(1-\gamma_5)\lambda_a c][\overline{q}\gamma^\mu(1-\gamma_5)\lambda_a b] \tag{1.13}$$

where λ_a are the Gell-Mann matrices. Using Fierz identities to express $Q_{(1)}$ and $Q_{(8)}$ in terms of Q_1 and Q_2 (of Eqn. 1.10) yields [4]

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{cb} V_{cq}^* [C_{(1)} Q_{(1)} + C_{(8)} Q_{(8)}] \tag{1.14}$$

$$C_{(1)} = 3C_1 + C_2 \tag{1.15}$$

$$C_{(8)} = 2C_2 \tag{1.16}$$

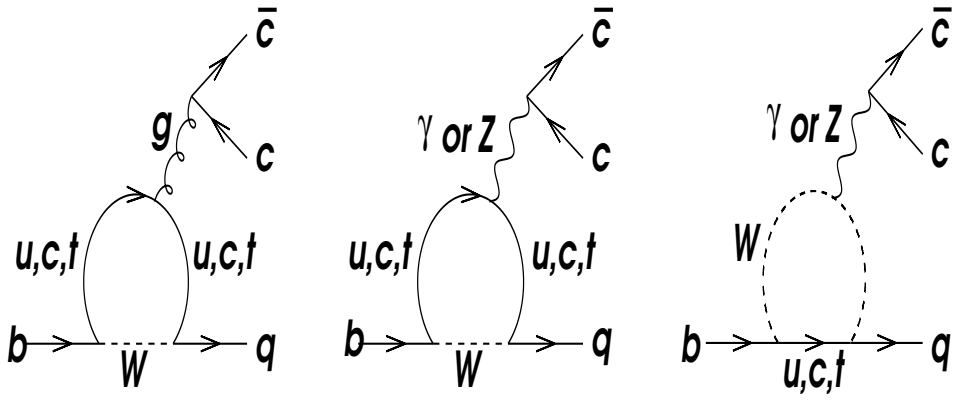

Figure 1.3: Penguin diagrams contributing $b \to c\bar{c}q$, where q is either s or d. (a) Left: QCD penguin diagram. (b) Center, Right: Electroweak penguin diagrams.

1.2.3 Penguin Diagrams

The $b \to c\bar{c}q$ transition may also occur by the "penguin" diagrams shown in Figure 1.3. While not listed here, Ref. [2] defines the 4-fermion operators $Q_3 \ldots Q_6$ for gluonic penguins and $Q_7 \ldots Q_{10}$ for the electroweak penguins, which enter into the effective Hamiltonian with corresponding Wilson coefficients $C_3 \ldots C_{10}$. Penguin decays are sensitive to non-Standard Model particles in the loop and are also important in studying CP violation in B decays; however, their contribution to the inclusive $b \to \psi X$ rate is very small relative to the tree diagram and is therefore ignorable.

1.2.4 Operator Product Expansion

The process used to obtain Equation 1.10 illustrates the Operator Product Expansion ("OPE") technique [5], in which the amplitude of a general process, such as a weak decay, may be written as

$$A = \langle \mathcal{H}_{eff} \rangle = \sum_i C_i \langle Q_i \rangle \tag{1.17}$$

where Q_i is a series of local operators, and C_i are the Wilson coefficients that describe their coupling strength.

Equation 1.17 effectively separates the problem into short-distance (C_i) and long-distance (Q_i) effects. The Wilson coefficients C_i contain all the short-distance strong-

interaction physics. Because of asymptotic freedom, these coefficients may be perturbatively calculated. The long-distance matrix elements $\langle Q_i \rangle$ include the effects of soft (long-distance) gluons; they must be experimentally determined, since they cannot be calculated in perturbation theory.

The boundary between "long" and "short" is given by the distance $1/\mu$. Clearly, the numerical values of C_i and $\langle Q_i \rangle$ both depend on the particular choice of the QCD renormalization scale μ. However, the final physical result, A (of Equation 1.17), must not depend on μ.

1.2.5 Numerical Values of the Wilson Coefficients

Although the Wilson coefficients C_i are calculable, the answer to the natural question "what are their values?" is complicated. The summary below emphasizes the factors that influence the numerical values of C_i; for a thorough theoretical review, see Ref. [2].

When calculating C_i in QCD perturbation theory, terms of the form $\ln(M_W/\mu)$ appear, which render the expansion invalid. Use of the Renormalization Group [6] corrects this problem by modifying the perturbation theory, so that the leading order (LO), or leading-logarithm approximation (LLA), solution to C_i is obtained by summing terms of the form $[\alpha_S(\mu)\,\ln(M_W/\mu)]^n$ over all $n \in \{0, \ldots, \infty\}$. Similarly, the next-to-leading order (NLO) or next-to-leading order logarithmic approximation (NLLA) solution is a sum of terms of the form $\alpha_S(\mu)\,[\alpha_S(\mu)\,\ln(M_W/\mu)]^n$. In order for these series to converge, $\alpha_S(\mu)$ must be sufficiently small, i.e., $\mu > \mathcal{O}(1\text{ GeV})$.

In calculating these coefficients it is more natural to work in the orthogonal operator basis $Q_\pm \equiv \dfrac{Q_1 + Q_2}{2}$ [7]. The renormalization group (RG)-improved solution for $C_\pm \equiv C_2 \pm C_1$, to LO, is [2]

$$C_\pm(\mu) = \left[\frac{\alpha_S(M_W)}{\alpha_S(\mu)}\right]^{\frac{9(\pm N-1)}{N(11N-2f)}} \tag{1.18}$$

where N is the number of quark colors and f is the number of quark flavors. For B decays, $N = 3$ and $f = 5$; choosing $\mu = M_B = 5.28$ GeV, and using the PDG 1999 average values for α_S [8] we obtain $\alpha_S(M_W) \approx 0.121$, $\alpha_S(M_B) \approx 0.211$, $C_- \approx 1.34$,

$C_+ \approx 0.86$, which by definition yield $C_1 \approx -0.24$ and $C_2 \approx 1.1$. Substituting this into Equations 1.15 and 1.16 yields

$$C_{(1)} \approx 0.4 \qquad \text{and} \tag{1.19}$$

$$C_{(8)} \approx 2.2 \qquad (\text{LO}, \mu = M_B) \tag{1.20}$$

which shows that the weak decay $b \to c\bar{c}q$ favors color-octet production of the $c\bar{c}$ pair. We note that the small value of $C_{(1)} = 3C_1 + C_2$ comes about because the opposite signs of C_1 and C_2 cause a partial cancellation.

There are four factors which affect the numerical values of the Wilson coefficients:

1. the renormalization scale μ;
2. choice of renormalization scheme (for NLO calculations);
3. LO or NLO;
4. the value of $\alpha_S(M_Z)$.

The sensitivity of the coefficients to these factors can be drastic for coefficients such as $C_{(1)}$, which are small and result from a subtraction; A comparison of the relative effects of these factors can be gained by examining values of the coefficients which are listed in Ref. [2].

Factor (1), μ, has a larger effect on the coefficients than any other factor. For example, $C_{(1)}^{LO}(\mu = 4.0 \text{ GeV}) = 0.23$ while $C_{(1)}^{LO}(\mu = 8.0 \text{ GeV}) = 0.47$. The variation is less severe for the NLO calculations, on the order of 30% for μ ranging from 4.0 to 8.0 GeV.

Factor (2), the renormalization scheme, affects the coefficient values at the NLO level. Two possible choices are known as "naive dimensional regularization" (NDR) and the 't Hooft-Veltman scheme (HV) [9]. The two differ by 10-20% in $C_{(1)}$, with closer agreement at higher μ.

Factor (3) is the degree of precision of the calculation; of course NNLO preferable to NLO, which is better than LO. As there are no full NNLO calculations currently available for $b \to \psi$, we will prefer the more recent calculations that go to NLO.

As far as the difference between LO and NLO in the Wilson coefficients, it was mentioned above that the NLO result has the benefit of being less dependent on the renormalization scale μ. The NLO calculations lead to a larger value of $C_{(1)}$ relative to the LO results by factors of 1.3-2.2, depending on the renormalization scheme and μ. Again, at higher values of μ, the agreement between LO and NLO results is closer.

Factor (4), the overall "normalization" of α_S, is given by current knowledge of $\alpha_S(M_Z)$, or the related quantity $\Lambda^{(5)}_{\overline{MS}}$ [10]. This factor is therefore driven by experimental error. The PDG 2001 values $\alpha_S(M_Z) = 0.1181 \pm 0.002$, or $\Lambda^{(5)}_{\overline{MS}} = 208^{+25}_{-23}$ MeV are sufficiently precise to cause only small variations in $C_{(1)}$: 2-8% for LO and 1-3% for NLO[1], with the higher variations coming from low μ.

To summarize, for $b \to c\bar{c}q$ decays, the largest factor affecting the values of Wilson coefficients is the renormalization scale μ. The lower the value of μ, the larger the difference caused by varying the other factors of LO/NLO, renormalization scale, and $\alpha_S(M_Z)$.

However, the sensitivity of the Wilson coefficients to the four factors is highly dependent on the particular coefficient under consideration. For example, the rate of suppression of color singlet $c\bar{c}$ production, relative to color octet production, is found to be

$$\left[\frac{C_{(8)}}{C_{(1)}}\right]^2 \approx \begin{cases} 21 - 117 & \text{for LO} \\ 9 - 30 & \text{for NLO} \end{cases} \tag{1.21}$$

where the variation is overwhelmingly due to varying μ between 4 GeV and 8 GeV. However, applying the same analysis to the coefficient C_2 results in values of $1.085 - 1.137$ for LO and $1.046 - 1.104$ for NLO – a maximum variation of only 9%.

As mentioned previously, the numerical value of μ is completely arbitrary (when within a sensible range); however, predictions of physical rates, which have the form $\sum_i C_i \langle Q_i \rangle$, must not depend on μ. Therefore, when studying the long-distance

[1] At the time of publication of Ref. [2], the world average was $\Lambda^{(5)}_{\overline{MS}} = 225 \pm 85$ MeV; this paper lists values of C_1 and C_2 for $\Lambda^{(5)}_{\overline{MS}} = 140$, 225, and 310 MeV. The observation $C_n(\Lambda^{(5)}_{\overline{MS}}$ =225 MeV) = $[C_n(\Lambda^{(5)}_{\overline{MS}}$ =140 MeV) + $C_n(\Lambda^{(5)}_{\overline{MS}}$ =310 MeV)]/2 at the few percent level, justified the use of linear interpolation to estimate $C_n(\Lambda^{(5)}_{\overline{MS}}$ =208 MeV). The uncertainty in C_n due to uncertainty in $\Lambda^{(5)}_{\overline{MS}}$ was estimated as $C_n(\Lambda^{(5)}_{\overline{MS}}$ =208 MeV) − $C_n(\Lambda^{(5)}_{\overline{MS}}$ =225 MeV).

hadronic matrix elements $\langle Q_i \rangle$, one must specify the renormalization scale μ to draw quantitatively meaningful conclusions.

To this point, we have considered the effective low-energy Hamiltonian for the $b \to c\overline{c}s$ transition, Equation 1.14. This equation includes the contribution of hard gluons via the Wilson coefficients $C_{(1)}$ and $C_{(8)}$, where the 1 and 8 refer to color singlet and color octet production of the $c\overline{c}$ pair. The effective Hamiltonian is an example of the Operator Product Expansion technique, Equation 1.17.

Although it separates the $c\overline{c}$ production into color singlet and color octet components, Equation 1.14 falls short of making predictions that can be compared to experiment, because it does not describe the hadronization of the $c\overline{c}$ into the experimentally observed ψ bound state.

As of this writing, the Non-Relativistic QCD (NRQCD) framework provides the most accurate and theoretically reliable understanding of the seemingly simple transition $c\overline{c} \to \psi$. Before reviewing its main ideas, we take a short detour to briefly discuss the two models which preceded it, the Color Singlet Model (CSM) and Color Evaporation Model (CEM).

1.3 Pre-NRQCD Models

1.3.1 Color Singlet Model

In the CSM, ψ is produced only when by $c\overline{c}$ pairs are produced with zero relative velocity in a color singlet state. Long-distance gluons are assumed to have no effect on either the color or angular momentum quantum numbers of the $c\overline{c}$ pair. These simplifying assumptions have two major implications. First, the production rate of $b \to \psi$ is directly related to that of $b \to c\overline{c}_{(1)}$, which is calculable in perturbation theory. Second, the type of charmonium produced ($J/\psi, \eta_c, \chi_{c1}$, etc.) is determined completely by the angular momentum quantum numbers of the original $c\overline{c}$ pair.

A longstanding theoretical problem for the CSM was first reported in 1980 [11]: the NLO solution for P-wave charmonium exhibited infrared divergence. Twelve years later, Bodwin *et al.* [4] solved the problem by including color octet terms.

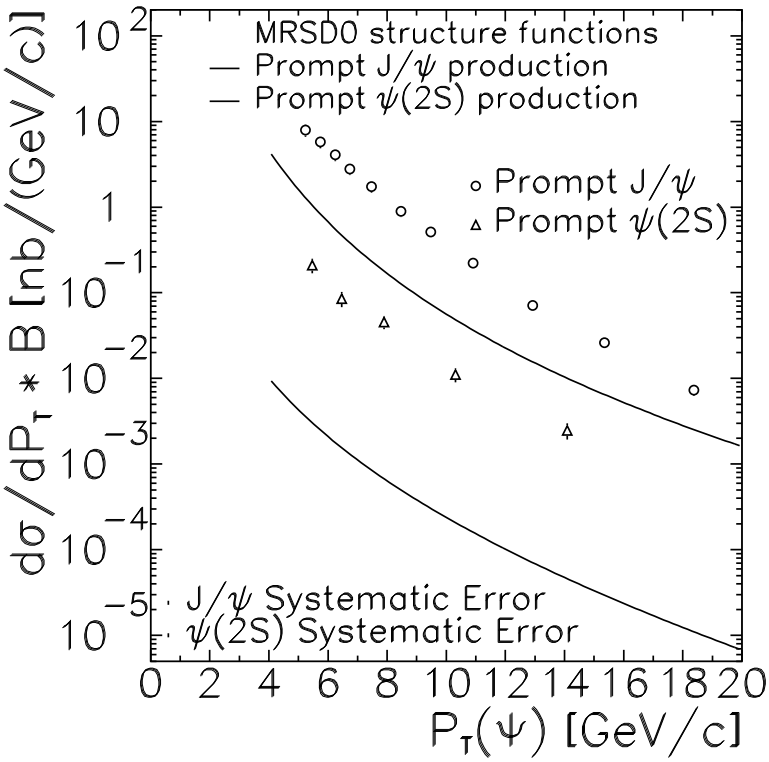

Figure 1.4: CDF experimental evidence ruling out the Color Singlet Model. The figure and its caption are taken from Ref. [13]: "The differential cross section times branching ratio $\mathcal{B}(\psi \to \mu^+\mu^-)$ for $|\eta^\psi| < 0.6$ for prompt ψ mesons. The vertical error bars are the statistical and the P_T-dependent systematic uncertainties, added in quadrature. Circles: J/ψ; triangles: $\psi(2S)$. The lines are the theoretical expectations based on the color singlet model."

Solid experimental evidence disfavoring the CSM was reported in 1992; CDF [12] observed production cross sections for prompt $\psi(1S)$ and $\psi(2S)$ at large p_T that were significantly higher than predicted by the CSM. A subsequent CDF analysis [13] made use of a silicon vertex detector to veto ψ that were produced from b decays; again they observed a a large excess ($\sim 50\times$) in the production rate of direct $\psi(2S)$ relative to the CSM.

NRQCD was able to accommodate the excess in prompt $\psi(1S)$ and $\psi(2S)$ production at CDF; some of the more accurate extractions of the matrix element $\langle \mathcal{O}_8(^3S_1)\rangle$ are derived from the CDF results. NRQCD also predicts that the polarization of

$\psi(2S)$ should become increasingly transverse as p_T increases [14]. Recent measurements from CDF [15], however, showed the opposite trend in central values, although with large errors.

1.3.2 Color Evaporation Model

The CEM [16] postulates that the exchange of soft gluons dominates the charmonium production process. The color and spin states of the $c\overline{c}$ pair at the time of creation therefore "evaporate" and have no effect on the state of the charmonium that eventually results. Another consequence of this postulate is that the production of charmonium states will occur in fixed ratios that are completely independent of the $c\overline{c}$ production mechanism. The role of kinematical variables in determining the production ratios of charmonium states is therefore ignored [17], placing the CEM on less firm theoretical ground than the CSM.

The CEM predicts that production of the χ_{cJ} system will universally occur in the ratios χ_{c2}:χ_{c1}:χ_{c0} = 5:3:1 (that is, proportional to $2J+1$.) A recent CLEO measurement [18] of $\frac{\Gamma(B \to \chi_{c2}(direct)X)}{\Gamma(B \to \chi_{c1}(direct)X)} < 0.44$ at 95% C.L., effectively rules out the CEM.

The CEM also makes a prediction that a measurement made in this thesis can cleanly test; namely, that charmonium mesons will be produced with no net polarization.

1.4 NRQCD

1.4.1 Motivation

In 1995, the pioneering work of Bodwin, Braaten, and Lepage [19] applied the effective field theory of Non-Relativistic QCD (NRQCD) to derive a rigorous OPE for the decay and production of heavy quarkonium. Just as we would expect for an OPE (Equation 1.17), NRQCD factorizes $b \to \psi$ production into a short-distance part, calculable in perturbation theory in powers of α_S, and a nonperturbative long-

distance part. Additionally, NRQCD expands the long-distance part into a series in v^2, where v is the typical velocity of the c quark inside the ψ bound state. Therefore a hierarchy of long-distance matrix elements is established in powers of v; consequently, one may consider only the leading-order v^6 terms (minimal NRQCD) or also add the contributions of the v^8 terms (the "v^2-improved" corrections). The situation with the long-distance part of the OPE therefore becomes analogous to that of the short-distance part, in which one may calculate the Wilson coefficients to LO or NLO.

NRQCD can in principle be made to yield identical predictions to full QCD, to any desired accuracy, by keeping the terms in the double-infinite expansion up to a sufficiently high power of v, then tuning the resulting input parameters (the number of which scales with the number of powers in the v^{2n} series) to match the full theory. The CSM and CEM, on the other hand, are models which make plausible approximations which are designed to simplify calculations, but at the price of dispensing with QCD. The NRQCD approach allows for the possibility that soft gluons will change the spin and color states of the originally produced $c\overline{c}$ pair, and allows for much subtlety in describing the effects of these soft gluons.

In this section we first describe the ideas behind NRQCD, then show how the CEM and CSM may be derived from it. Like the first section, this is a brief summary; for more details, there exist numerous excellent review articles and conference papers [14] [17] [20] [21] [22] [23].

1.4.2 The NRQCD Lagrangian

The average velocity v of the $c\overline{c}$ quarks in a ψ bound state, where this velocity is measured in the center-of-mass frame of the quarks, may be estimated from quark potential models [24] as $v^2 \sim 0.3$ for charmonium and $v^2 \sim 0.1$ for bottomonium. These values are confirmed by lattice QCD calculations [25]. A back-of-the-envelope calculation to explain this follows: $c\overline{c}$ quarks in a color singlet state that are sufficiently close to each other will experience an attractive, Coulombic color force, thus enabling an analogy to atomic physics [14]: the kinetic energy of the quarks is analogous to the Rydberg energy, which determines the energy level spacings of excited atomic states.

This leads to $M_c v^2 \approx M_{\psi(2S)} - M_{\psi(1S)} \approx M_{\chi_{c0}} - M_{\psi(1S)} \approx 500$ MeV, or $v^2 \sim 0.3$.

The hierarchy of energy scales in charmonium is then

$$\Lambda_{QCD} \sim M_c v^2 < M_c v < M_c \tag{1.22}$$

Λ_{QCD} is the the energy scale of nonperturbative QCD, $M_c v^2$ is the scale of the charmonium mass structure, $M_c v$ is the inverse of the typical size of the charmonium bound state, and M_c sets the energy scale for $c\overline{c}$ creation. This hierarchy also holds, and is even more differentiated for the $b\overline{b}$ system, but is invalid for $s\overline{s}$. In NRQCD the division between "long" and "short" distances in the OPE occurs at a renormalization scale Λ, which is chosen to be between $M_c v$ and M_c.

For the "high" energy scale of $\mathcal{O}(M_c)$, NRQCD dispenses with the usual 4-component Dirac spinor for the c quark field, replacing it with two 2-component Pauli spinors: ψ, which annihilates a c quark; and χ, which creates a $\overline{c}$ (it is a somewhat confusing coincidence that ψ and χ are also the names of certain charmonium mesons). Since v is small, they are treated nonrelativistically, in a Schrödinger field theory. The energy scales of order $M_c v$ and smaller are governed by the full QCD Lagrangian. This energy scale division is accomplished by setting a ultraviolet momentum cutoff Λ (same as the renormalization scale), which excludes the existence of relativistic c quarks in the theory.

By requiring that SU(3) gauge symmetry, rotational symmetry, **C** symmetry, **P** symmetry, and heavy-quark phase symmetry apply to the above fields, the technique of effective field theory [26] yields the "minimal NRQCD" Lagrangian [19]

$$\mathcal{L}_{minimal} \equiv \mathcal{L}_{light} + \mathcal{L}_{heavy} \tag{1.23}$$

$$\mathcal{L}_{light} = -\frac{1}{2}\mathrm{tr}\, G_{\mu\nu}G^{\mu\nu} + \sum \overline{q}\, i\not{D}\, q \tag{1.24}$$

$$\mathcal{L}_{heavy} = \psi^\dagger \left(iD_t + \frac{\mathbf{D}^2}{2M} \right) \psi + \chi^\dagger \left(iD_t - \frac{\mathbf{D}^2}{2M} \right) \chi \tag{1.25}$$

where

$$D^\mu \equiv \partial^\mu + igA^\mu \tag{1.26}$$

$$A^\mu = (\phi, \mathbf{A}^\mu) \tag{1.27}$$

$$g = \sqrt{4\pi\alpha_S} \tag{1.28}$$

and $G^{\mu\nu}$ is the gluon field strength tensor.

Going beyond the minimal Lagrangian requires the addition of correction terms $\delta\mathcal{L}$, the most significant of which are bilinear in the heavy quark fields [19]:

$$\begin{aligned}
\mathcal{L}_{improved} &\equiv \mathcal{L}_{minimal} + \mathcal{L}_{bilinear} && (1.29)\\
\mathcal{L}_{bilinear} &= \frac{c_1}{8M^3}[\psi^\dagger(\mathbf{D}^2)^2\psi - \chi^\dagger(\mathbf{D}^2)^2\chi] \\
&\quad \frac{c_2}{8M^2}[\psi^\dagger(\mathbf{D}\cdot g\mathbf{E} - g\mathbf{E}\cdot\mathbf{D})\psi + \chi^\dagger(\mathbf{D}\cdot g\mathbf{E} - g\mathbf{E}\cdot\mathbf{D})\chi] \\
&\quad \frac{c_3}{8M^2}[\psi^\dagger(i\mathbf{D}\times g\mathbf{E} - g\mathbf{E}\times i\mathbf{D})\cdot\boldsymbol{\sigma}\psi + \chi^\dagger(i\mathbf{D}\times g\mathbf{E} - g\mathbf{E}\times i\mathbf{D})\cdot\boldsymbol{\sigma}\chi] \\
&\quad \frac{c_4}{2M}[\psi^\dagger(g\mathbf{B}\cdot\boldsymbol{\sigma})\psi - \chi^\dagger(g\mathbf{B}\cdot\boldsymbol{\sigma})\chi] && (1.30)
\end{aligned}$$

where the chromoelectric and chromomagnetic field operators are denoted by

$$\begin{aligned}
E^i &= G^{0i} && (1.31)\\
B^i &= \frac{1}{2}\epsilon^{ijk}G^{jk}. && (1.32)
\end{aligned}$$

One powerful feature of NRQCD is that an estimate of the magnitude of every operator which appears in the heavy-quark Lagrangian (α_S, ψ, χ, D_t, $\mathbf{D}$, $g\mathbf{E}$, $g\mathbf{B}$, $g\phi$, and $g\mathbf{A}$) may be calculated in powers of M and v. Applying this "power counting" technique shows that each term in $\mathcal{L}_{heavy}$ (Equation 1.25) is of order M^4v^5, whereas the terms of $\mathcal{L}_{bilinear}$ (Equation 1.30) are of order M^4v^7. Since the volume of the ψ meson scales as $1/(Mv)^3$, $\int\mathcal{L}_{heavy}d^3x$ scales as Mv^2 and $\int\mathcal{L}_{bilinear}d^3x$ scales as Mv^4. Comparing these factors, we see that the $\mathcal{L}_{bilinear}$ provides an $\mathcal{O}(v^2)$ correction to the minimal NRQCD Lagrangian. In fact, $\mathcal{L}_{bilinear}$ contains all of the order Mv^4 terms [19]. Further correction terms (of order Mv^6, Mv^8, etc.) may be added to the NRQCD Lagrangian, thus enabling calculations made in NRQCD to match those of full QCD to any desired order of accuracy in v [27].

The minimal NRQCD Lagrangian (Equation 1.23) is symmetric in heavy-quark spin, and therefore predicts, for example, a mass degeneracy in the $^3S_1$ and $^1S_0$ charmonium states (which is incorrect: $m_{\psi(1S)} - m_{\eta_c} = 117 \pm 2$ MeV [10].) The bilinear correction term (Equation 1.30) includes terms with the Pauli spin matrices, $\boldsymbol{\sigma}$, which split this degeneracy.

A number of phenomenological input parameters remain which must be tuned to match full QCD. For minimal NRQCD, the parameters M and g require tuning; in the v^2-improved theory, the dimensionless parameters c_1, c_2, c_3, and c_4 also must be tuned. Although calculable in principle from perturbation theory, current efforts to determine these parameters rely on a combination of the perturbative and nonperturbative matching techniques [20].

Once the tuned parameters are obtained, lattice simulations of NRQCD provide useful tests of the nonrelativistic approximation and the validity of the power counting rules. Lattice calculations of the v^2-improved Lagrangian, Eqn. 1.29, reproduce the mass spectrum of experimentally observed charmonium states (1^1S_0, 1^3S_1, 2^3S_1, 2^3P_0, 2^3P_1, and 2^3P_2) with relative errors of $\sim 10\%$ (approximately v^4) for spin-independent splittings, and $\sim 30\%$ (approximately v^2) for spin-dependent splittings [25]. That the errors correspond with powers of v^2 is predicted by the power-counting rules of NRQCD. Further supporting evidence is found by analyzing the bottomonium system in a similar way [28]; the spectroscopic accuracy of the lattice calculation (using the same Lagrangian) is significantly better than the charmonium case, corresponding to the fact that v^2 is lower for Υ by about a factor of 3.

1.4.3 Factorization Formula for Inclusive ψ Production

Using the NRQCD Lagrangian, Bodwin, Braaten, and Lepage [19] obtain rigorous factorization formulas for the inclusive annihilation decay rates and inclusive production rates of heavy quarkonium. The formulas follow the general form of the OPE (Equation 1.17); the long-distance matrix elements are expressed in terms of the NRQCD heavy-quark creation and annihilation operators ψ and χ, sandwiched in between $\langle H|$ and $|H\rangle$ (annihilation), or $\langle 0|$ and $|0\rangle$ (production), where H is the bound-state charmonium hadron. Since this thesis is concerned only with studying properties of ψ production, the reader interested in the details of NRQCD's treatment of ψ decay may refer to the previously mentioned review articles for details. Likewise, many details involved in the derivation of the inclusive production factorization formula are omitted here, and we simply quote the result [19]:

$$\sigma(H) = \sum_n \frac{F_n(\Lambda)}{M^{d_n-4}} \langle 0|\mathcal{O}_n^H(\Lambda)|0\rangle \tag{1.33}$$

$$\mathcal{O}_n^H = \chi^\dagger \mathcal{K}_n \psi (a_H^\dagger a_H) \psi^\dagger \mathcal{K}_n' \chi \tag{1.34}$$

$$a_H^\dagger a_H = \sum_X \sum_{m_J} |HX\rangle\langle HX| \tag{1.35}$$

$$n \leftrightarrow {}_c(^{2S+1}P_J) \tag{1.36}$$

In Eqn. 1.33, σ is the cross-section for $i \to f$, where the initial state i is anything, and the final state f consists of the charmonium bound state H, plus any other particles X. The sum is over every possible state n (Eqn. 1.36) of color, spin, and angular momentum in which the $c\overline{c}$ emerges after being created and interacting with hard gluons. The F_n are dimensionless short-distance coefficients, analogous to the Wilson coefficients. Λ is the factorization scale. $\langle 0|\mathcal{O}_n^H(\Lambda)|0\rangle$ is the probability that a $c\overline{c}$ quark pair in the n color-spin state will eventually transform into a charmonium bound state H, with the possible assistance of soft gluons.

This equation exhibits all of the OPE characteristics: the scale Λ separates the problem into short- and long-distance factors, each of which is dependent on Λ in such a way that the product, being a physical observable, is independent of Λ; the short-distance coefficients F_n are calculable in QCD perturbation theory; the nonperturbative long-distance coefficients describe the contributions of soft gluons.

This formula applies to *all* mechanisms of ψ production, and that the particular details of the production are encoded in the short-distance coefficients F_n; therefore the long-distance matrix elements $\langle 0|\mathcal{O}_n^H(\Lambda)|0\rangle$ are universal and may be constrained using data from experiments which create ψ in very different environments.

In Eqn. 1.34, the factors $\mathcal{K}_n$ and $\mathcal{K}_n'$ are combinations of Gell-Mann matrices, Pauli spin matrices, and polynomials in the covariant derivative $\mathbf{D}^\mu$, such that $|c\overline{c}(n)\rangle = \psi^\dagger \mathcal{K}_n' \chi|0\rangle$. Here the operators χ and $\psi^\dagger$ create the quarks, and the $\mathcal{K}_n'$ manipulates their quantum state n. Table 1.1 lists the definitions of the production operators of leading and relative-v^2 orders.

In Eqn. 1.35, the operator a_H destroys a quarkonium H that exists at any time following the creation of the $c\overline{c}$ pair. Therefore the string of operators in Equation

Order	$c\bar{c}$ Color Singlet			$c\bar{c}$ Color Octet		
in v	Oper.	$\mathcal{K}_n$	$\mathcal{K}'_n$	Oper.	$\mathcal{K}_n$	$\mathcal{K}'_n$
v^6	$\mathcal{O}_1(^1S_0)$	1	1	$\mathcal{O}_8(^1S_0)$	λ^a	λ^a
v^6	$\mathcal{O}_1(^3S_1)$	σ^i	σ^i	$\mathcal{O}_8(^3S_1)$	$\sigma^i\lambda^a$	$\sigma^i\lambda^a$
v^8	$\mathcal{O}_1(^1P_1)$	$-\frac{i}{2}D^i$	$-\frac{i}{2}D^i$	$\mathcal{O}_8(^1P_1)$	$-\frac{i}{2}D^i\lambda^a$	$-\frac{i}{2}D^i\lambda^a$
v^8	$\mathcal{O}_1(^3P_0)$	$-\frac{i}{2\sqrt{3}}\mathbf{D}\cdot\boldsymbol{\sigma}$	$-\frac{i}{2\sqrt{3}}\mathbf{D}\cdot\boldsymbol{\sigma}$	$\mathcal{O}_8(^3P_0)$	$-\frac{i}{2\sqrt{3}}\mathbf{D}\cdot\boldsymbol{\sigma}\lambda^a$	$-\frac{i}{2\sqrt{3}}\mathbf{D}\cdot\boldsymbol{\sigma}\lambda^a$
v^8	$\mathcal{O}_1(^3P_1)$	$-\frac{i}{2\sqrt{2}}\mathbf{D}\times\boldsymbol{\sigma}$	$-\frac{i}{2\sqrt{2}}\mathbf{D}\times\boldsymbol{\sigma}$	$\mathcal{O}_8(^3P_1)$	$-\frac{i}{2\sqrt{2}}\mathbf{D}\times\boldsymbol{\sigma}\lambda^a$	$-\frac{i}{2\sqrt{2}}\mathbf{D}\times\boldsymbol{\sigma}\lambda^a$
v^8	$\mathcal{O}_1(^3P_2)$	$-\frac{i}{2}D^{(i}\sigma^{j)}$	$-\frac{i}{2}D^{(i}\sigma^{j)}$	$\mathcal{O}_8(^3P_2)$	$-\frac{i}{2}D^{(i}\sigma^{j)}\lambda^a$	$-\frac{i}{2}D^{(i}\sigma^{j)}\lambda^a$
v^8	$\mathcal{P}_1(^1S_0)$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}(-\frac{i}{2}\mathbf{D})^2$	$\mathcal{P}_8(^1S_0)$	$\frac{1}{\sqrt{2}}\lambda^a$	$\frac{1}{\sqrt{2}}(-\frac{i}{2}\mathbf{D})^2\lambda^a$
v^8	$\mathcal{P}_1(^3S_1)$	$\frac{1}{\sqrt{2}}\sigma^i$	$\frac{1}{\sqrt{2}}\sigma^i(-\frac{i}{2}\mathbf{D})^2$	$\mathcal{P}_8(^3S_1)$	$\frac{1}{\sqrt{2}}\sigma^i\lambda^a$	$\frac{1}{\sqrt{2}}\sigma^i(-\frac{i}{2}\mathbf{D})^2\lambda^a$

Table 1.1: NRQCD charmonium production operators, from Ref. [19]. The most important operators (those of order v^6 and v^8) are listed. The second and fifth columns list the operator being defined; the definition is given by Equation 1.34, with factors $\mathcal{K}_n$ and $\mathcal{K}'_n$ defined in columns 3, 4, 6, and 7. The $\mathcal{P}$ operators are secondary operators of higher order in v^2 relative to the $\mathcal{O}$ operators.

1.34 is proportional to the probability that the $c\bar{c}$ state will eventually become the charmonium state H.

1.4.4 Power Counting for State Transitions

Through power counting, NRQCD provides a qualitative picture of the relative probability for a given $c\bar{c}$ state to materialize into a bound-state charmonium H. As an example, consider $\psi(1S)$, the $^3S_1$ state. A $c\bar{c}$ pair which emerges from the processes of creation and hard gluon exchange in the $_1(^3S_1)$ state will naively have a relatively high probability of hadronizing as a $\psi(1S)$, as its color-spin quantum numbers match those of the $\psi(1S)$.

A $c\bar{c}$ pair which enters the long-distance realm in the $_8(^3S_1)$ state, however, is in a color octet state, and therefore will need to emit or absorb at least one soft gluon in order to become a singlet $\psi(1S)$ state. Since gluons carry one unit of angular momentum, gluon exchange must change the spin state of the $c\bar{c}$ pair. Furthermore, the fact that soft gluon exchange is required to obtain a color-singlet $c\bar{c}$ must somehow

affect the probability of hadronization.

Both of these points are addressed in NRQCD, which identifies two types of gluonic transitions which change the quantum state n [20]. "Electric" transitions are dominated by very soft dynamical gluons, satisfy the selection rules $\Delta L = \pm 1$ and $\Delta S = 0$, and occur with a probability of order v^2. "Magnetic" transitions are dominated by dynamical gluons with momenta of order Mv, satisfy selection rules $\Delta L = 0$ and $\Delta S = \pm 1$, and occur with a probability of order v^3. Both transitions change color octet states in to color singlet states and vice versa.

Therefore, by a series of these two types of gluonic transitions, any $c\bar{c}$ state can become any charmonium state H, regardless of the similarity or dissimilarity between the color-spin quantum numbers of the $c\bar{c}$ and H. In NRQCD, the power counting rules enable one to rank the $c\bar{c}$ states in order of their likelihood of eventually hadronizing as a given charmonium H.

1.4.5 Inclusive $b \to \psi X$ Production

The weak decay formalism of $b \to c\bar{c}s$ of the previous section may be combined with the NRQCD formalism of this section, to obtain an expression for the inclusive production rate for $b \to \psi(mS)X$, where $m \in \{1, 2\}$:

$$\Gamma(b \to \psi(mS)X) = \sum_n K_n \langle 0 | O_n^{\psi(mS)} | 0 \rangle \tag{1.37}$$

This infinite sum is a double expansion in α_S and v^2. To date, the most complete calculations [29] calculate the K_n to NLO. The long-distance terms in this calculation include the dominant color-singlet $(c\bar{c})_1(^3S_1)$ term, as well as terms of relative order v^4 from color-octet $c\bar{c}$ states:

$$\begin{aligned}\Gamma(b \to \psi(mS)X) &= K_1(^3S_1)\,\langle \mathcal{O}_1(^3S_1)\rangle \\ &+ K_8(^3S_1)\,\langle \mathcal{O}_8(^3S_1)\rangle + K_8(^3P_J)\,\langle \mathcal{O}_8(^3P_J)\rangle + K_8(^1P_1)\,\langle \mathcal{O}_8(^1P_1)\rangle\end{aligned} \tag{1.38}$$

K_n is the probability that the $b \to c\bar{c}s$ transition will produce the $c\bar{c}$ pair with a small relative velocity and in the n color-angular momentum state. The definition of the previously discussed Wilson coefficients $C_{(1)}$ and $C_{(8)}$ [see Eqns. 1.14, 1.21]

is identical to this, except that there are no requirements on the relative velocity or angular momentum of the $c\overline{c}$ pair. It should not be surprising, then, that the coefficients $K_8(^3S_1)$, $K_8(^3P_J)$, and $K_8(^1P_1)$ are all comparable to $C^2_{(8)}$, and similarly, $K_1(^3S_1) \sim C^2_{(1)}$. Qualitatively, Eqn. 1.38 shows that $b \to \psi$ production comes from a color singlet part which is enhanced in long-distance hadronization probability but suppressed in production rate; and a color octet part which is enhanced in production rate but suppressed in the long-distance hadronization probability.

As an aside, the CSM holds that soft gluons do not play a part in changing the quantum state n of $c\overline{c}$ pairs. Therefore the CSM version of Eqn. 1.38 is obtained by dropping all terms in the infinite sum except for first term, $K_1(^3S_1)\,\langle\mathcal{O}_1(^3S_1)\rangle$.

Before discussing the current state of knowledge of each term in Eqn. 1.38, we will briefly discuss the final missing link needed to relate theory to experiment. The discussion up to now has considered $b \to \psi X$, whereas experiments can observe only $B \to \psi X$. The fact that the b quark is itself in a bound state B meson must be considered.

1.5 Effects of B Hadronization

The B meson, where "B" in this thesis refers to the B^+, B^-, B^0, or $\overline{B}^0$ mesons, consists of a bound state of a heavy b quark (antiquark) and a light u or d antiquark (quark). The B meson may be visualized as a heavy b quark surrounded by an amorphous, complicated, and continuously changing cloud of light quarks and gluons.

Just as NRQCD expands the small parameter v in a power series, another effective field theory called Heavy Quark Effective Theory (HQET) employs an expansion in the parameter $1/m_b$, exploiting the fact that the b quark mass is about an order of magnitude larger than the scale of perturbative QCD, Λ_{QCD}. The literature on HQET is possibly even more voluminous than that of NRQCD; an oft-cited review article [30] provides more details than are summarized here.

In the limit $1/m_b \to 0$, i.e. to leading order in HQET, the decay of a b quark inside a B meson will be identical to the decay of a free b quark. In this limit, the decay width for all b hadrons will be identical. This approximation is justified at about

the 30% level, which is the range of experimentally measured lifetimes of the B^+, B^0, B_s, and Λ_b hadrons. Including the next-to-leading-order terms of HQET would yield corrections that quantify the difference between b and B decay, thus treating rigorously the Fermi motion of the b quark in the B meson.

An often-used alternative to the HQET formalism is to model the Fermi motion of the b quark in the B meson with the ACCMM model [31]. In this model, the meson's light quarks and gluons are replaced by a single spectator quark of mass m_s with a Gaussian momentum distribution, the width of which is characterized by the Fermi momentum p_F. Given these two parameters, conservation of energy and momentum then uniquely determine the distribution of the b quark's "floating mass."

1.6 $B \to \psi X$

1.6.1 Outline: State of Current Calculations

It should be clear at this point that strong interactions render the decay $B \to \psi X$ highly nontrivial to quantitatively describe. Nevertheless, the introduction of the NRQCD formalism in 1995 unleashed a large theoretical effort to calculate predictions of ψ production and decay in a diverse array of processes. This section summarizes the recent theoretical results for $B \to \psi X$, which make predictions of the $B \to \psi X$ inclusive branching fraction (Sec. 1.6.2), the momentum distribution of ψ created from B decays (Sec. 1.6.3), and the polarization of ψ created from B decays (Sec. 1.6.5). Regardless of what is being predicted, the calculations may be categorized in how they handle the following three issues:

- Precision of short-distance coefficients (order in α_S) (Sec. 1.2)
  - 1: LO, or LLA
  - 2: NLO, or NLLA
- Number of long-distance matrix elements (LDME's) (order in v^2) (Sec. 1.4)
  - 1 (CSM): leading color-singlet only (Sec. 1.3.1)

- 4: leading color-singlet and color-octet
- 8: leading and derivative color-singlet and color-octet

- Handling of the b quark within the B meson (Sec. 1.5)
  - 0: Treat b quark as being free
  - 1: Model b quark Fermi motion with ACCMM model
  - 2: Include first correction term from HQET expansion

When discussing predictions in the following sections, we will use the notation (x, y, z) to describe the level of precision of a given publication, where each of the numbers is taken from the above list. For example, $(1, 4, 1)$ will mean that a prediction is leading-order in α_S, sums over 4 long-distance matrix elements, and models the b quark Fermi motion in the ACCMM model. Historically, successive papers improve the level of precision, but the older papers are still useful, since examining the change introduced by going to the next level of accuracy gives some indication of the level of confidence to attribute to the current state-of-the-art calculations.

1.6.2 Inclusive Branching Fraction

Ko *et al.* [32] provided a (1,3,0) analysis soon after the NRQCD paper was published. In 1997, Beneke, Maltoni, and Rothstein [29] extended the analysis to (2,4,0), pointing out that the large sensitivity of the Wilson coefficient $C_{(1)}$ to the renormalization scale necessitated a NLO(α_S) calculation. This calculation also showed that the color-singlet $^3S_1$ term $K_1(^3S_1)\langle\mathcal{O}_1(^3S_1)\rangle$ contributes negatively to the decay rate at NLO. This was patched up by including some, but not all, NNLO correction terms. The fact that the LO, NLO, and quasi-NNLO calculations of $K_1(^3S_1)$ are positive, negative, and positive, respectively, are a sure sign that the true value of this coefficient is near zero – in other words, the short-distance production of color-singlet $c\bar{c}$ pairs in $b \to c\bar{c}s$ is very highly suppressed. The latest results, published in 2000 by Ma [33] (2,4,2), derive correction terms resulting from the b quark Fermi motion.

The expression for the inclusive branching fraction is [33]

$$\begin{aligned}
\mathcal{B}(B \to \psi X) &= 0.00073\langle \mathcal{O}_1(^3S_1)\rangle \\
&\quad +0.19\langle \mathcal{O}_8(^3S_1)\rangle + 0.33\langle \mathcal{O}_8(^1S_0)\rangle + \frac{0.34}{m_c^2}\langle \mathcal{O}_8(^3P_0)\rangle \quad (1.39) \\
\Delta_B\mathcal{B}(B \to \psi X) &= -0.00068\langle \mathcal{O}_1(^3S_1)\rangle \\
&\quad -0.07\langle \mathcal{O}_8(^3S_1)\rangle - 0.03\langle \mathcal{O}_8(^1S_0)\rangle - \frac{0.05}{m_c^2}\langle \mathcal{O}_8(^3P_0)\rangle \quad (1.40)
\end{aligned}$$

where $\Delta_B\mathcal{B}(B \to \psi X)$ denotes the correction to the branching fraction introduced by the NLO HQET correction. Note that these equations follow the form of Eqn. 1.37. All of the LDME's $\langle \mathcal{O}\rangle$ are in units of GeV^3. We now discuss each of the above terms, in order.

First Term, $\langle \mathcal{O}_1(^3S_1)\rangle$: The long-distance matrix element $\langle \mathcal{O}_1(^3S_1)\rangle$ is the only one that is calculable directly from the NRQCD heavy quark radial wavefunction [34]:

$$\langle \mathcal{O}_1(^3S_1)\rangle = \frac{9|R(0)|^2}{2\pi} = \begin{cases} 1.16 \text{ GeV}^3 & \text{for } \psi(1S) \\ 0.76 \text{ GeV}^3 & \text{for } \psi(2S) \end{cases} \quad (1.41)$$

With this expression for $\langle \mathcal{O}_1(^3S_1)\rangle$, we may now obtain the CSM prediction for the inclusive branching ratio. Substituting the above result into Eqn. 1.39, and dropping all of the other (color octet) terms, we obtain $\mathcal{B}_{CSM}(B \to \psi(1S)X) = 0.08\%$ and $\mathcal{B}_{CSM}(B \to \psi(2S)X) = 0.06\%$. This contrasts with measurements from a previous CLEO analysis [35]:

$$\begin{aligned}
\mathcal{B}(B \to \psi(1S)X) &= (0.80 \pm 0.08)\% \qquad \text{and} \\
\mathcal{B}(B \to \psi(2S)X) &= (0.34 \pm 0.05)\%. \quad (1.42)
\end{aligned}$$

Therefore the CSM predicts inclusive branching fractions that are a factor of 6-10 smaller than experimentally observed.

Second Term, $\langle \mathcal{O}_8(^3S_1)\rangle$: The color-octet matrix element $\langle \mathcal{O}_8(^3S_1)\rangle$ is extracted from analyses of ψ transverse momentum distributions at the Tevatron. The extraction used in Refs. [29] and [33], for example, is taken from [36]; here, the first error

is the statistical (experimental) error, and the second error is due to varying the renormalization and factorization scales μ:

$$\langle \mathcal{O}_8(^3S_1)\rangle = \begin{cases} (1.06 \pm 0.14^{+1.05}_{-0.59}) \times 10^{-2}\ \text{GeV}^3 & \text{for } \psi(1S) \\ (0.44 \pm 0.08^{+0.43}_{-0.24}) \times 10^{-2}\ \text{GeV}^3 & \text{for } \psi(2S) \end{cases} \tag{1.43}$$

Unfortunately, in the dozen or so extractions of this matrix element that have been carried out to date, the variation in results is rather large. A recent review article [23] summarizes all of the extractions with a range of possible values:

$$\langle \mathcal{O}_8(^3S_1)\rangle = \begin{cases} (0.3 \sim 2.0) \times 10^{-2}\ \text{GeV}^3 & \text{for } \psi(1S) \\ (0.1 \sim 1.0) \times 10^{-2}\ \text{GeV}^3 & \text{for } \psi(2S) \end{cases} \tag{1.44}$$

We may now calculate the contribution to the inclusive branching fraction from the color-octet mechanism $B \to (c\bar{c})_8(^3S_1)X \to \psi X$. Substituting Eqn. 1.44 into Eqn. 1.39, we obtain branching fraction contributions of $(0.06 - 0.38)\%$ from $B \to (c\bar{c})_8(^3S_1)X \to \psi(1S)X$ and $(0.02 - 0.19)\%$ from $B \to (c\bar{c})_8(^3S_1)X \to \psi(2S)X$. This particular color octet mechanism contributes significantly to the total inclusive branching fraction, although the errors are very large.

Third and Fourth Terms, $\langle \mathcal{O}_8(^1S_0)\rangle$ and $\langle \mathcal{O}_8(^3P_0)\rangle$: The remaining terms in Eqn. 1.44 involve the less-well-known matrix elements $\langle \mathcal{O}_8(^1S_0)\rangle$ and $\langle \mathcal{O}_8(^3P_0)\rangle$. By substituting Eqns. 1.41 and 1.43 into Eqn. 1.44, and setting the sum equal to the 1995 CLEO measurements (Eqn. 1.42) as well as the matrix element values in Eqns. 1.41 and 1.43 to obtain an expression for a combination of $\langle \mathcal{O}_8(^1S_0)\rangle$ and $\langle \mathcal{O}_8(^3P_0)\rangle$, thus providing a constraint on these matrix elements. The factor of 10 uncertainty in $\langle \mathcal{O}_8(^3S_1)\rangle$, however, results in a prediction for these matrix elements of a similar accuracy.

Outlook Presently, the theoretical uncertainties in extracting values of the hadronic matrix elements $\langle \mathcal{O}\rangle$ are much larger than the experimental uncertainties. Because of this, the improved measurements of $\mathcal{B}(B \to \psi X)$ that are presented in this thesis, although useful in reducing the errors of other measurements that depend on them, are not likely to improve the current understanding of these matrix elements.

1.6.3 Momentum Distribution

A full prediction of the ψ momentum distribution needs to incorporate two effects which smear the ψ momentum. First is the Fermi motion of the b quark in the B meson, for which the smearing is of order Λ_{QCD} [37]. Second is the exchange of soft gluons, which is required for the hadronization of color-octet $c\bar{c}$; the soft gluons carry momenta of order $Mv^2 \approx \Lambda_{QCD}$. Although the NRQCD LDME's incorporate the role of soft gluons in hadronization, the inclusion of soft gluon *kinematics* is "usually a higher-order effect in the non-relativistic expansion" [38].

In 1997 Palmer, Paschos, and Soldan [37] (1,4,1) made a first attempt to fit the spectrum. This study only includes smearing due to Fermi motion. The calculated momentum distribution which best matches the 1995 CLEO data [35] corresponds to a Fermi momentum of 570 MeV.

Most recently, Beneke, Schuler, and Wolf [38] (1,4,1) model both smearing effects. The soft-gluon contributions are modeled with an Gaussian energy distribution ansatz inspired from the ACCMM model [31]. The soft-gluon coupling strength in the ansatz is tailored for each angular momentum and color state, based on the necessary electric and magnetic gluon transitions needed to reach the ${}_1({}^3S_1)$ state. The ACCMM model is applied to simulate the Fermi motion of the b quark; however, it is asserted that the ACCMM model fails for the two-body decays $B \to \psi K^{(*)}$.

The result is a series of p_ψ distributions, as a function of p_F, Λ, and n. Here p_F is the Fermi momentum, Λ is the mean energy of soft gluons, and $n \in \{{}_8({}^1S_0),\ {}_8({}^3P_0),\ {}_8({}^3S_1)\}$ is the quantum state of the generated $c\bar{c}$. The calculated distributions are compared to the measured $p_\psi(1S)$ distribution from CLEO [35], excluding the contribution of the $B \to \psi K^{(*)}$ decays. The best fits are observed with $\Lambda = 300^{+50}_{-100}$ MeV, $p_F \sim 300$ MeV, and a roughly equal mixture of $c\bar{c}$ states n. The result for Λ is in excellent agreement with the NRQCD scaling rules, $\Lambda \sim m_c v^2 \sim \Lambda_{QCD}$ (c.f. Eqn. 1.22). Although the fit results for the mixture of n could conceivably be used to determine the LDME's $\langle \mathcal{O}_8({}^1S_0)\rangle$, $\langle \mathcal{O}_8({}^3P_J)\rangle$, and $\langle \mathcal{O}_8({}^3S_1)\rangle$, these extractions are not reported.

Outlook The measurements of the ψ momentum distributions presented in this thesis are based on higher statistics and correct for detector measurement smearing, compared to the previous CLEO measurement [35]. An improved measurement may make it possible to better determine the NRQCD long-distance color octet matrix elements.

In a development completely unrelated to NRQCD, a recent paper [39] speculates that an narrow enhancement in the $p_{\psi(1S)}$ distribution at about 0.5 GeV would possibly be a signature of the triple-charm decay $B \to \psi(1S)D\pi$. A faint hint of a bump is exhibited in measurements of CLEO [35] and Belle [40], although the statistics make this far from clear. Recent results from BaBar [41] show no evidence of the bump. The results from this thesis should help to settle the matter.

1.6.4 Review: Polarization Definitions

Before summarizing the calculations for ψ polarization, we will define the helicity angle θ and the polarization parameters α and Γ_L/Γ.

Helicity Angle

Given the decay chain $B \to \psi X$, $\psi \to \ell^+\ell^-$, the helicity angle θ is defined as the angle formed by a daughter lepton in the ψ rest frame, and the direction of the ψ in the B rest frame.

If the ψ is produced in a helicity ± 1 state (transverse polarization), then the resulting angular distribution of daughter leptons is given by

$$\frac{1}{\Gamma}\frac{d\Gamma(\cos\theta)}{d\cos\theta}\left[\psi \to \ell^+\ell^-(\theta)\right] = \frac{3}{8}(1+\cos^2\theta) \tag{1.45}$$

whereas the distribution of leptons from a helicity-zero (longitudinally polarized) ψ is

$$\frac{1}{\Gamma}\frac{d\Gamma(\cos\theta)}{d\cos\theta}\left[\psi \to \ell^+\ell^-(\theta)\right] = \frac{3}{4}(1-\cos^2\theta) \tag{1.46}$$

Here Γ is the decay width. The normalization factors are chosen so that integrating the above equations in $\cos(\theta)$ over $[-1,+1]$ yields 1. Also, note that because the $\cos^2\theta$

term cannot distinguish between the angles θ and $(\theta + \pi)$, experimentally either the ℓ^+ or ℓ^- may be used to determine the helicity angle.

Average Polarization

The average polarization of a population of ψ is quantified as follows. Let N_0, N_+, and N_- to be the number of ψ produced with helicity 0, +1, and −1. Experimentally, we cannot distinguish helicity +1 from helicity −1; what we can measure is the sum of the two,

$$N_\pm \equiv N_+ + N_- \tag{1.47}$$

$$N \equiv N_0 + N_\pm \tag{1.48}$$

The angular distribution for all N events is:

$$\sum_N \frac{1}{\Gamma}\frac{d\Gamma}{d\cos(\theta)}\left[\psi \to \ell^+\ell^-(\theta)\right] = \frac{3}{8}\left[(N_\pm + 2N_0) + (N_\pm - 2N_0)\cos^2\theta\right] \tag{1.49}$$

$$= \frac{3}{8}(N_\pm + 2N_0)\left(1 + \frac{N_\pm - 2N_0}{N_\pm + 2N_0}\cos^2\theta\right) \tag{1.50}$$

$$= \frac{3}{2(\alpha+3)}(1 + \alpha\cos^2\theta)N \tag{1.51}$$

where the polarization parameter α is defined as

$$\alpha \equiv \frac{N_+ + N_- - 2N_0}{N_+ + N_- + 2N_0} = \frac{N_\pm - 2N_0}{N_\pm + 2N_0} \tag{1.52}$$

To determine the polarization and size of the entire sample, it is sufficient to measure either the pair of variables (N_0, $N_\pm$), or the pair (N, α). Transforming from the (N_0, $N_\pm$) basis to the (N, α) basis is described by the above equations (1.48) and (1.52). The inverse transformation, from (N, α) to (N_0, $N_\pm$), can be derived from these two equations:

$$N_0 = \left(\frac{1-\alpha}{\alpha+3}\right)N \tag{1.53}$$

$$N_\pm = 2\left(\frac{\alpha+1}{\alpha+3}\right)N \tag{1.54}$$

An alternate parameterization of polarization for the ψ population is the fraction of ψ which are longitudinally polarized.

$$\frac{\Gamma_L}{\Gamma} = \frac{N_0}{N} = \frac{N_0}{N_\pm + N_0} \tag{1.55}$$

The relations between the two parameterizations are

$$\frac{\Gamma_L}{\Gamma} = \frac{1-\alpha}{3+\alpha} \tag{1.56}$$

$$\text{whence} \quad \Delta\frac{\Gamma_L}{\Gamma} = \frac{4}{(3+\alpha)^2}(\Delta\alpha) \tag{1.57}$$

$$\text{and} \quad \alpha = \frac{1-3\left(\frac{\Gamma_L}{\Gamma}\right)}{1+\left(\frac{\Gamma_L}{\Gamma}\right)} \tag{1.58}$$

$$\text{whence} \quad \Delta\alpha = \frac{4}{(3+\frac{\Gamma_L}{\Gamma})^2}\left(\Delta\frac{\Gamma_L}{\Gamma}\right) \tag{1.59}$$

Table 1.2 summarizes the above discussion.

α	Γ_L/Γ	Polarization	Helicity State(s)	Angular Distribution
$+1$	0	Transverse	$+,-$	$1+\cos^2\theta$
0	$+1/3$	Unpolarized	$+,-,0$	1
-1	$+1$	Longitudinal	0	$1-\cos^2\theta$

Table 1.2: Parameters describing the average polarization of a population of ψ. The cases of complete spin alignment (transverse, longitudinal) and random spin alignment (unpolarized) are shown.

1.6.5 Polarization

In the CSM, in the limit $m_d = m_u = m_s = 0$, the decay rates Γ for the three ψ helicity states are given by [42]

$$\Gamma_-/\Gamma_0 = 2\left(\frac{m_{\psi(1S)}}{m_b}\right)^2 \tag{1.60}$$

$$\Gamma_+ = 0 \tag{1.61}$$

$$\Rightarrow \alpha = \frac{\Gamma_+ + \Gamma_- - 2\Gamma_0}{\Gamma_+ + \Gamma_- + 2\Gamma_0} = \frac{m^2_{\psi(1S)} - m_b^2}{m^2_{\psi(1S)} + m_b^2} \tag{1.62}$$

Equation 1.61 comes about because in the decay $b \to \psi q$, the massless q quark must have helicity $+1/2$; this forces the helicity states of the spin-1 ψ to be either 0 or -1. Table 1.3 lists predictions for $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$ for several values of m_b.

m_b (GeV)	4.1	4.4	4.7	5.0	5.3
$\alpha_{\psi(1S)}$	-0.27	-0.34	-0.39	-0.45	-0.49
$\alpha_{\psi(2S)}$	-0.11	-0.18	-0.24	-0.30	-0.35

Table 1.3: Predictions for α in the CSM.

The effect of B meson effects on the b quark was examined by Palmer *et al.* [43] (0,4,1), which combine CSM predictions with two different B meson models (the inclusive parton and ACCMM models), resulting in predictions for $\alpha_{\psi(1S)}$ which are listed in Table 1.4. The variation in α gives a qualitative estimate of the effect of B hadron effects on the polarization.

$\epsilon_p = 0.004$	$\epsilon_p = 0.006$	$\epsilon_p = 0.008$	$\epsilon_p = 0.010$
-0.40	-0.38	-0.37	-0.36
$p_f = 0.3$	$p_f = 0.4$	$p_f = 0.5$	$p_f = 0.55$
-0.42	-0.41	-0.39	-0.38

Table 1.4: Predictions for α, from [43]. ϵ_p is the parameter of the Peterson [44] distribution function. p_f is the Fermi momentum.

Fleming *et al.* [45] (1,4,0), calculating short-distance coefficients to LO, obtain

$$\alpha = \frac{-0.39\langle\mathcal{O}_1(^3S_1)\rangle - 17\langle\mathcal{O}_8(^3S_1)\rangle + 52\langle\mathcal{O}_8(^3P_0)\rangle/m_c^2}{\langle\mathcal{O}_1(^3S_1)\rangle + 44\langle\mathcal{O}_8(^3S_1)\rangle + 211\langle\mathcal{O}_8(^3P_0)\rangle/m_c^2 + 61\langle\mathcal{O}_8(^1S_0)\rangle} \tag{1.63}$$

which, for then-current ranges of LDME's, lead to the predictions in Table 1.5.

m_b (GeV)	4.1	4.4	4.7	5.0	5.3
$\alpha_{\psi(1S)}$	[-0.12, -0.01]	[-0.28, +0.05]	[-0.33, +0.05]	[-0.41, -0.01]	[-0.48, -0.10]

Table 1.5: Predictions for α from NRQCD calculations. m_b is the b quark mass.

Ma [46] (1,8,0) carries out the LO calculation to 4 additional LDME's and finds that the correction obtained, relative to the above results, is up to 70 $\sim$ 80% for

some matrix elements, concluding that any (1,4,0) predictions such as Table 1.5 are therefore unreliable.

Prospects None of the calculations have yet to express α to NLO in α_S. Once such calculations are available, our measurement of the polarization parameter α is expected to provide useful constraints on the NRQCD LDME's. In the meantime, the CEM, although already disallowed by measurements of $B \to \chi$ production (see Section 1.3.2), makes a clean prediction of $\alpha = 0$, against which we may compare our results.

Chapter 2

Experimental Apparatus

2.1 CESR

2.1.1 Historical Background: 1977-1980

The Cornell Electron Storage Ring, or CESR, began operation in 1979. It is capable of delivering electron-positron collisions at a center-of-mass energy of between 9 and 11 GeV. Collisions occur at a precisely determined location, around which the multipurpose detector CLEO I was built. In 1977, the year when construction of CESR began, a fixed-target experiment at Fermilab [47] reported an observation of a narrow resonance in the hadronic cross-section, the $\Upsilon(1S)$, which was immediately interpreted as a quarkonium state made of a new, third-generation quark: beauty (or bottom), b. Three years later, the first two publications [1] to emerge from the CESR/CLEO partnership reported observations of the Υ resonances, including a first observation of $\Upsilon(4S)$. Figure 2.2 shows the results of these energy scans.

Even with statistics that now seem miniscule, it was clear that the measured width of the $\Upsilon(4S)$ was substantially larger than the widths of the lower resonances, signifying that this resonance had a large enough energy to allow decays that were disfavored in the others. Furthermore, by examining the the Fox-Wolfram moment `R2`, a dimensionless, rotationally invariant measure of the degree to which the decay products uniformly cover 4π, the decay products of hadronic events created at the

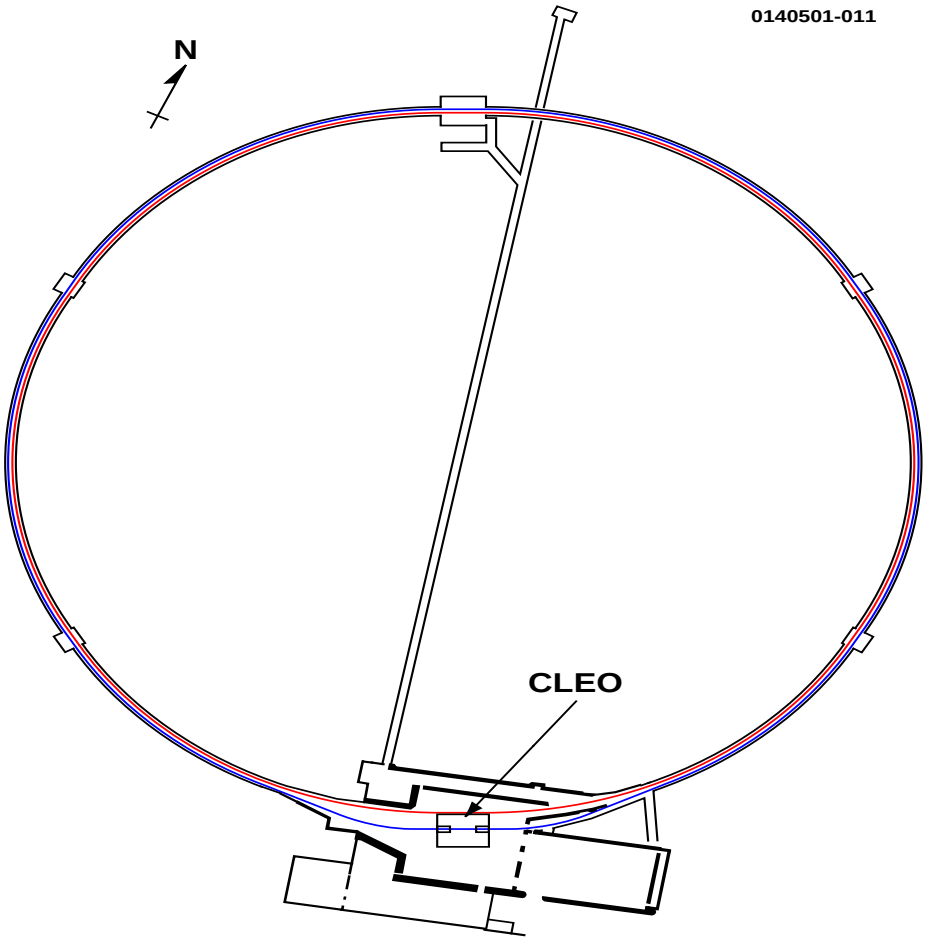

Figure 2.1: Aerial view diagram of CESR and CLEO, to scale. CESR is represented by the line which passes through CLEO. Also shown are the outer walls of Wilson Lab (the building which houses CLEO) and the cross tunnel providing a shortcut to the North area, with its experimental hall (unused since 1989). Not shown are the Cornell athletic fields and the several stories of soil which separate them from the tunnel.

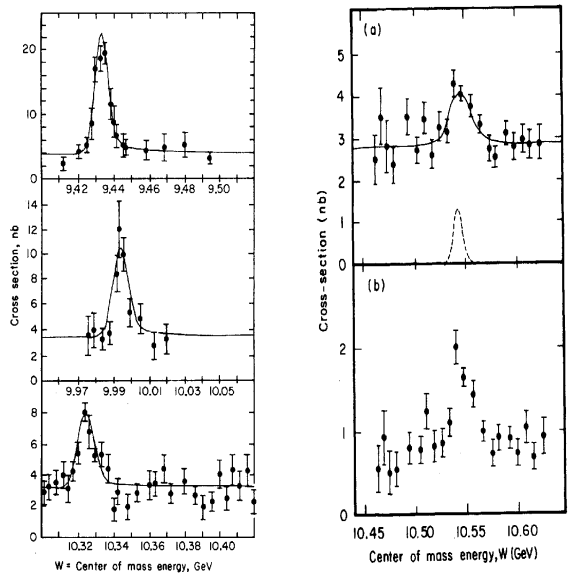

Figure 2.2: Energy scan results from the first two CLEO publications, circa 1980 [1]. Left: The $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ resonances. Right, above: first observation of $\Upsilon(4S)$. Right, below: same as above, keeping only "spherical" events with R2 < 0.3. Horizontal axes: E_{CM}, in GeV. Vertical axes: cross-section, in nb.

$\Upsilon(4S)$ were observed to be, on average, more "spherical" than events observed when the beam energy was below the resonance. These observations established the $\Upsilon(4S)$ as the lowest-energy beautyonium resonance with sufficient energy to decay into two mesons, each containing a beauty quark (or antiquark).

The benefits of running at the $\Upsilon(4S)$ for studying beauty physics were immediately clear. The reaction $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ created "clean" events in which all decay products were attributable to a B meson. Furthermore, energy and momentum conservation impose very powerful constraints on the reconstruction of exclusive decays, and increase the power of the R2 variable to distinguish between $B\overline{B}$ events and "continuum" events, i.e. $e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$, or $c\overline{c}$.

2.1.2 CESR Operation

Figure 2.3 is a diagram showing the main components of the CESR/CLEO system. During a data-taking run, the positron beam is filled first. Electrons are emitted

from a cathode, formed into bunches, and are accelerated by the linac; halfway down, at 150 MeV, they encounter a fixed tungsten target. In the shower that ensues, the small fraction of the particles that are positrons are magnetically selected and injected into the synchrotron. After about 4,000 revolutions around the ring, the positrons are accelerated to their final energy of 5.29 GeV, then transferred via the west transfer line into CESR. This entire process is repeated 60 times per second until the positron beam reaches the desired current. Filling the positron beam from scratch takes on the order of 5 minutes.

The electron beam is then injected by the same method except without the fixed tungsten target; electrons traverse the full length of the linac and are injected into the synchrotron, and subsequently CESR, in the opposite direction as the positrons. Injection of the electron beam requires less than a minute.

Synchrotron radiation is a major effect. The energy lost to synchrotron radiation per $e^{\pm}$ per revolution is about 1.2 MeV, or 0.02% of its energy. The radiated energy is replenished with 4 RF cavities located around the ring. (In 1997 the first of two 500 MHz superconducting cavities was installed; the second was installed in 1999, after CLEO II was decommissioned.) Synchrotron radiation, while undesirable to the high-energy physicist due to electricity expense and physics backgrounds, is highly valuable for X-ray diffraction studies. The research facility known as CHESS, the Cornell High-Energy Synchrotron Source, maintains two target areas at the kinks in the southern part of the ring (see Figure 2.1) where the mean energy per photon is the greatest.

Collisions at the IP occur as soon as the electron and positron beams are fully injected into CESR. The beams collide for about an hour, at which point the electron and positron beams are replenished with additional current. Figure 2.4 shows a recent example of a "perfect day" of running, in which the luminosities stay consistently high ($1 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) and there is no downtime due to lost beams.

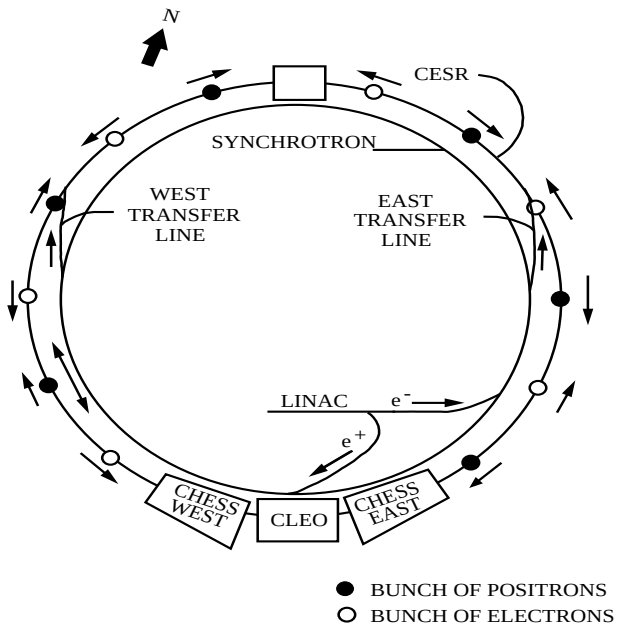

Figure 2.3: Main components of CESR and CLEO.

2.1.3 CESR Luminosity

The b quark in B mesons most often decays in a cascade of weak decays to c and subsequently d or s, along with the creation of two virtual W's which typically become charged pions, kaons, or lepton-neutrino pairs. The c quark, which hadronizes as a D or charmonium state will only be measured in certain signature decay modes with branching fractions on the order of 1-10%. Therefore, any detailed investigation of beauty physics requires a large number of decays to study. This number is directly proportional to the integrated luminosity delivered to the detector. In the 20 years since CESR started running at the $\Upsilon(4S)$, a myriad of improvements in beam techniques and hardware upgrades has led to an exponential increase in the instantaneous luminosity (see Figure 2.5.) The long-term success of CESR has enabled CLEO to make precision measurements of B decays, such as the ones presented in this thesis.

Until 1994, the positron and electron currents in CESR were both distributed in 9 "bunches" spaced equally around the ring. In 1994 the crossing angle scheme [48]

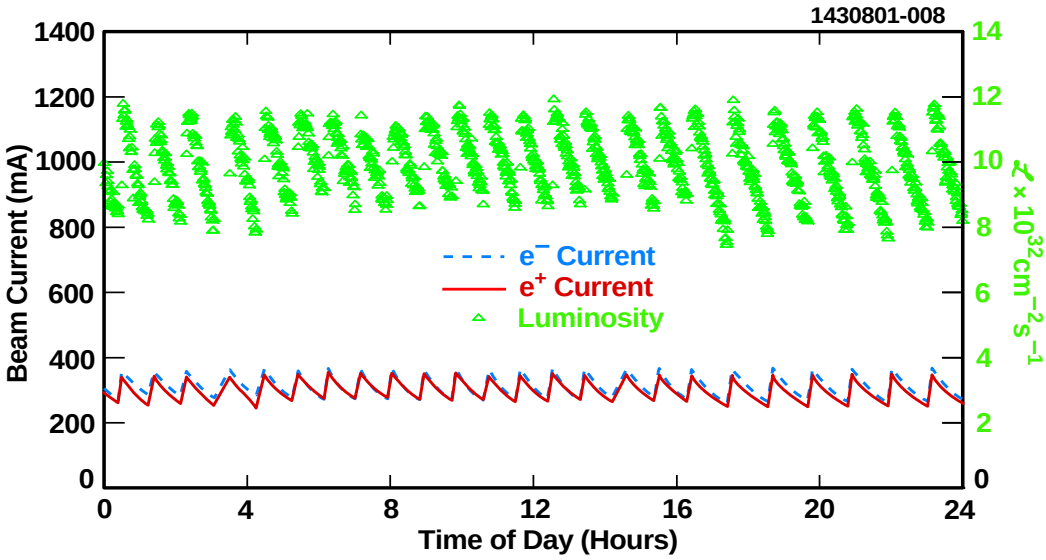

Figure 2.4: 24 hours of data-taking.

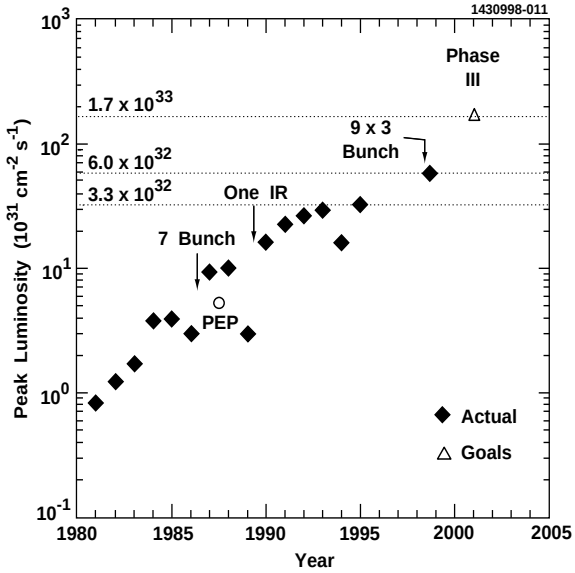

Figure 2.5: Peak instantaneous luminosity of CESR, 1980-2000.

was implemented. This required the installation of electrostatic separators, which manipulate the horizontal position of both beams; as shown in Figure 2.6, the separators modify the orbits of the beam, thus avoiding parasitic crossing points. Furthermore, during the final approach to the IP, the beams were steered in at an angle of 2 mrad off of horizontal. Consequently, the $\Upsilon(4S)$ rest frame became slightly boosted ($\beta = 0.004$) relative to the lab frame. This enabled an upgrade of the 9 bunches around the ring to 9 "trains" of bunches, each train consisting of 5 neighboring buckets, 14 ns apart. Each of the 45 buckets in the bunch trains may be filled with particles, thus increasing the maximum sustainable beam current. CESR first demonstrated this "multi-bunch running" with 9 trains, 2 bunches per train (or, "9×2" running) with an inter-bunch spacing of 42 ns. Soon, records for beam current and instantaneous luminosity were broken. Later, CESR moved up to a 9×3 bunch configuration. Figure 2.6 shows the time elapsed between the beam-crossing signal (or, to be precise, the train-crossing signal) and hits from charged particles on the CLEO time-of-flight system. Clear peaks are seen for each of the 21 filled buckets. Track-fitting requires knowledge of which bunch caused a given event; the author has written a "bunch-finder," described later, which uses CLEO time-of-flight information for this purpose.

The distribution of points in space at which electron-positron collisions occur is directly related to the profile of the beams. At CLEO the distribution of electrons and positrons in a given bucket is roughly Gaussian in all dimensions. The beam size[1] at the interaction point was recently measured by Cinabro *et al.* [49] by relating the beam parameters to the dimensions of the luminous region, as determined from tracks in the CLEO data. They observe $\sigma_y = 6.9 \pm 2.8\ \mu$m, $\sigma_x = 0.5$ mm, and $\sigma_z = 1.93 \pm 0.02$ cm. Here, z is the direction of the positron beam, y is up, and x is radially outward, i.e. away from the center of the ring. These measurements agree well with lattice calculations, modified by previously observed effects such as the dynamic beta effect [50] and beam-beam focusing. Given these beam size measurements, the electron and positron bunches may be visualized as long, thin, and very flat ribbons.

[1] given by $\sqrt{\epsilon\beta^*}$, where ϵ is the beam emittance and β^* is the betatron function

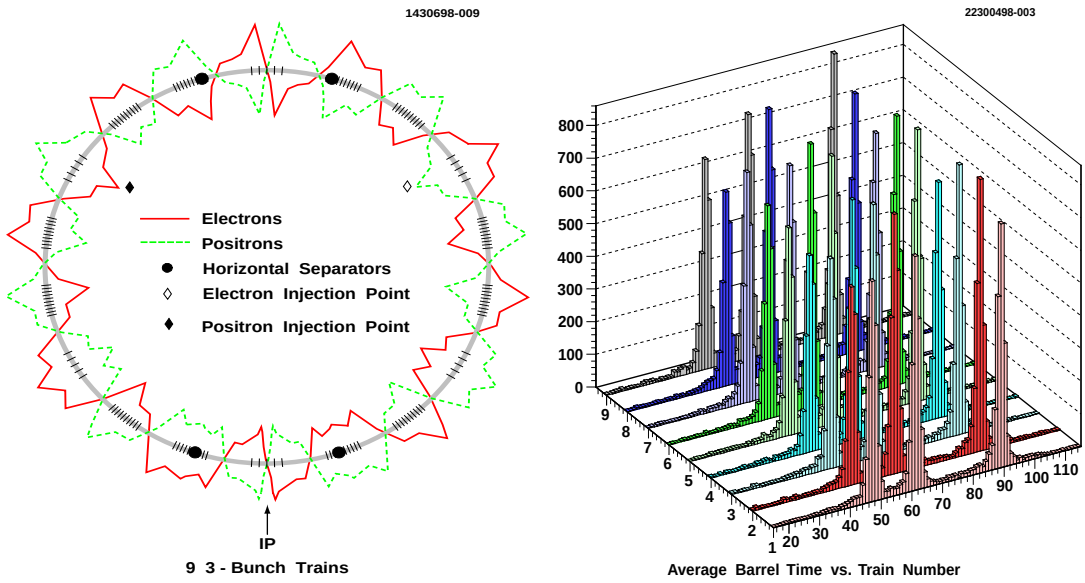

Figure 2.6: Left: Manipulation of the horizontal beam position in CESR to enable a crossing angle, and therefore multibunch trains, at the interaction point. Right: Observation of 9 train, 3 bunch running by the CLEO time-of-flight scintillators. Vertical axis: events. Left axis: train number. Right axis: t_{TOF} in nanoseconds.

2.2 CLEO

2.2.1 Overview

The data analyzed in this thesis was taken from the CLEO II (1989-1995) and CLEO II.V (1995-1999) experiments. CLEO II was designed to measure both charged and neutral particles with high efficiency and accuracy, covering over 95% of 4π. Figures 2.7, 2.8, and 2.9 point out the major components of the detector. In 1995 the innermost detector (PTL, or Precision Tracking Layer) was removed and a 3-layer Silicon Vertex Detector was installed in its place, marking the beginning of the CLEO II.V years. Table 2.1 shows the integrated luminosity observed by the two experiments.

The sections that follow describe in overview the components of CLEO II(.V), starting from the elements closest to the interaction point and proceeding radially outward. For more information, Ref. [51] describes CLEO II in great detail.

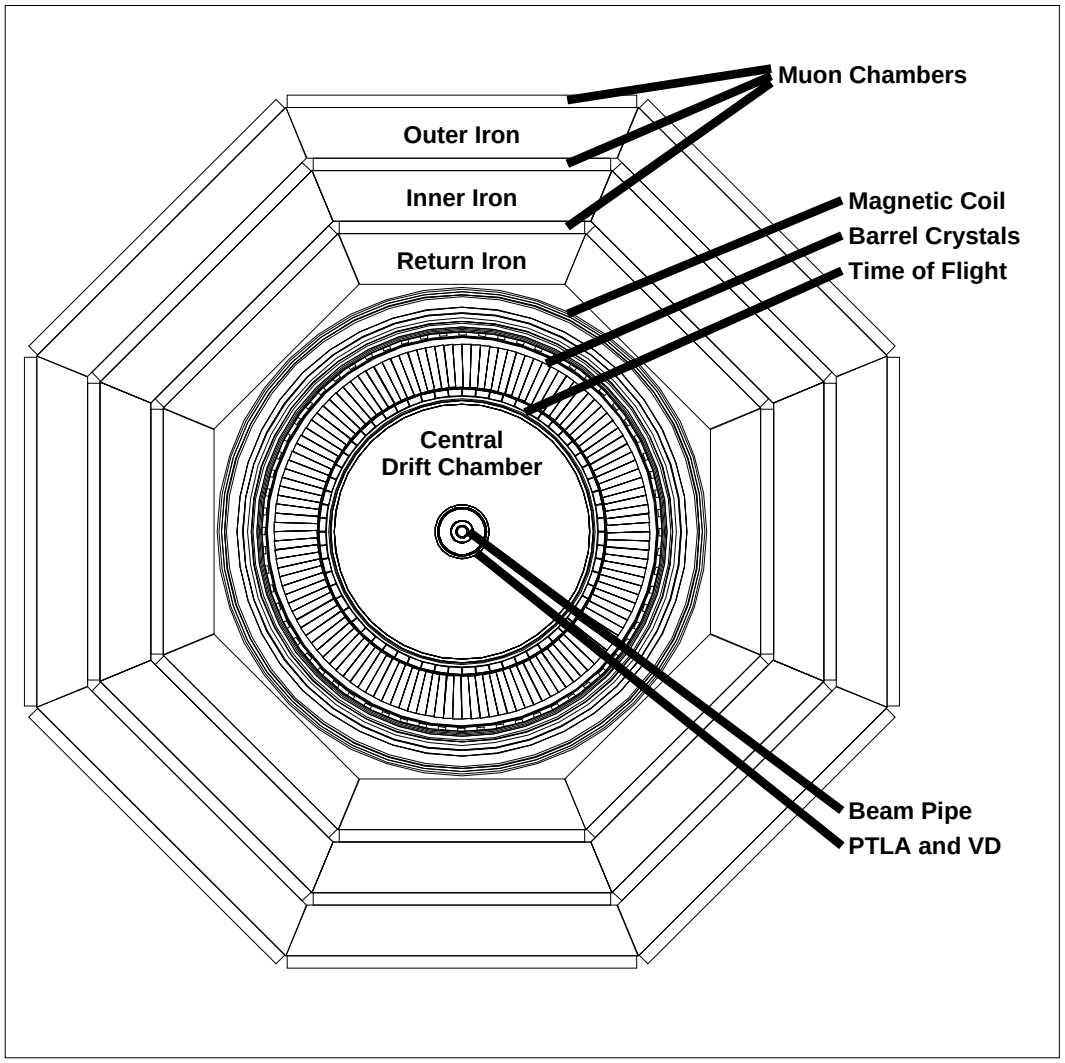

Figure 2.7: CLEO II detector, end view. The SVX replaces the PTLA in CLEO II.V.

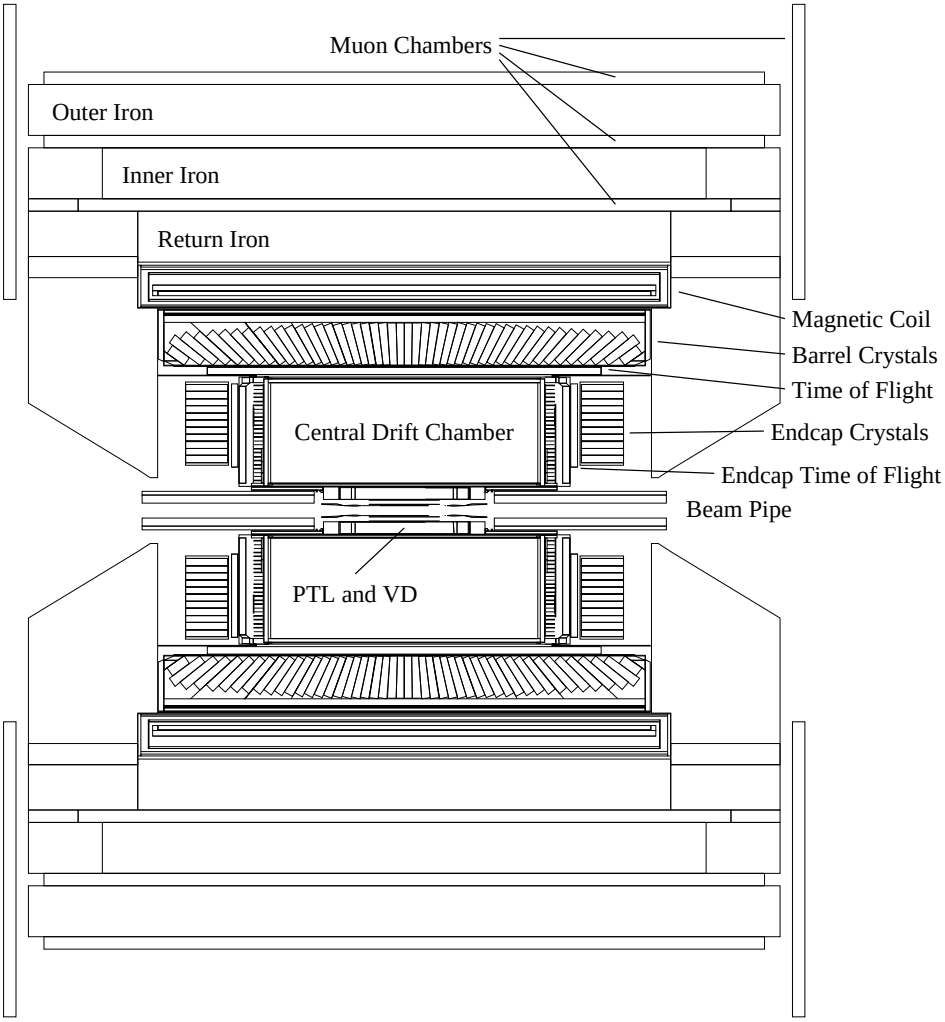

Figure 2.8: CLEO II detector, cutaway view. The SVX replaces the PTL in CLEO II.V.

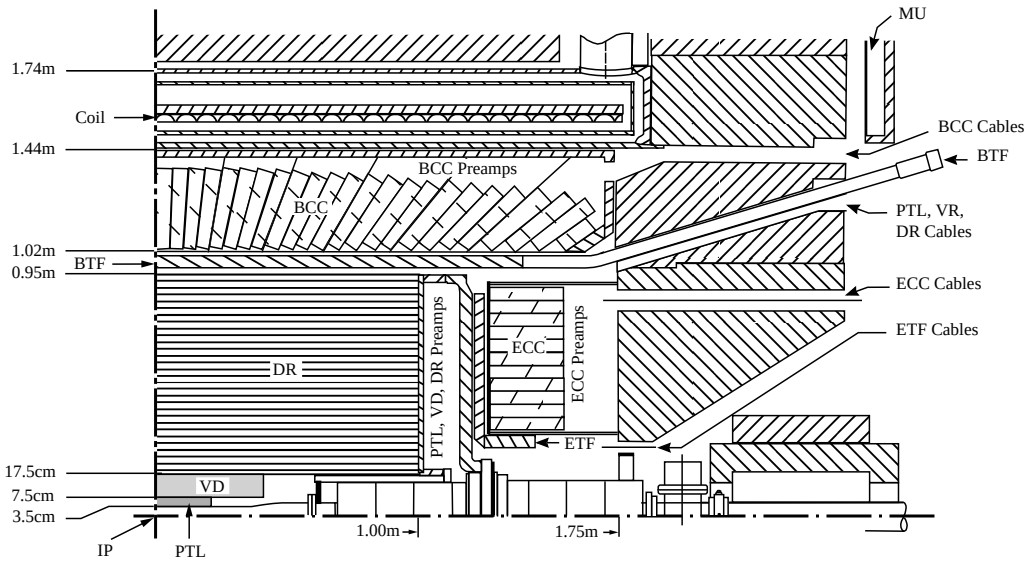

Figure 2.9: CLEO II quarter-detector cutaway view. The SVX replaces the PTL in CLEO II.V.

2.2.2 Tracking

In CLEO II, three concentric proportional wire chambers, all in a 1.5 T magnetic field, are used together to measure the 3-vector momentum of charged tracks. The r-ϕ components of momentum are measured from the curvature of tracks. The z component of momentum is measured from cathode pads attached to the cylindrical surfaces of the outer two detectors, and stereo angle wires in the outermost detector (again, z is the direction of the positron beam). The main detector also provides particle ID information by measuring dE/dx. CLEO II.V uses the two outer wire chambers of CLEO II, but replaced the innermost wire chamber with a Silicon Vertex Detector.

PTL; Beampipe (CLEO II)

The Precision Tracking Layer (PTL) directly surrounds the beampipe and consists of 6 sets of 64 tubes, the dimensions and arrangement of which are depicted in Figure 2.10. The distance between the inner and outer walls is just over an inch. The tubes

Detector	$\int \mathcal{L}dt$ (fb^{-1}) On-Resonance	Continuum	$B\overline{B}$ Events (10^6)
CLEO II	3.1	1.6	3.3
CLEO II.V	6.0	2.8	6.4
Total	9.1	4.4	9.7

Table 2.1: CLEO data sets used in this analysis. "On-Resonance" refers to data taken at a center-of-mass energy of 10.580 GeV, at which $e^+e^- \to \Upsilon(4S)$ is a possible decay. The center-of-mass energy is lowered for "Continuum" running, in which $\Upsilon(4S)$ production does not occur.

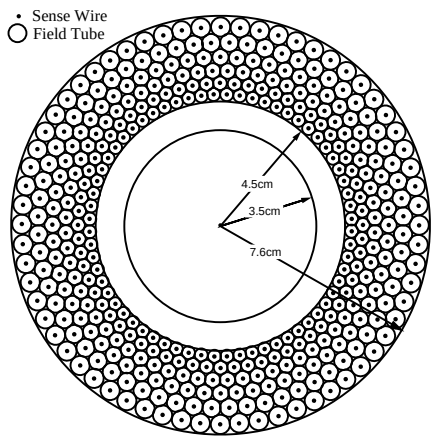

Figure 2.10: Precision Tracking Layer (PT). CLEO II only.

are made of aluminized mylar and are electrically grounded, while the axial sense wires, centered within each tube, are held at a positive high voltage. The tubes were filled with argon-ethane (50-50) until 1992, when a switch was made to dimethyl ether for its lower drift velocity. The half-cell shift in ϕ of adjacent layers not only allows for maximally efficient use of space, but also avoids possible ambiguity in ϕ when fitting a track to a series of isochrones. This detector only measures the transverse component of momentum, as all wires are parallel to the beamline.

The 35mm radius beampipe for CLEO II is made of 0.5 mm thick beryllium. A 25 μm silver coating on the inside mitigates backgrounds from synchrotron radiation.

SVX; Beampipe (CLEO II.V)

Replacing the PTL with the CLEO II.V SVX, a three-layer double-sided silicon detector, resulted in a twofold improvement in the $r\phi$ impact parameter resolution, and a tenfold improvement for z (see Figure 2.11.) This gave CLEO the ability to reconstruct secondary decay vertices of, for example, boosted charm mesons.

The beampipe in CLEO II.V had much more stringent design constraints than that of CLEO II. The readout electronics of the SVX generated about 100 W of heat, and radiation studies of the CAMEX readout chips showed that extra shielding was required. The CLEO II.V beampipe has two beryllium walls of 0.25 mm thickness with a 0.5 mm gap through which cooling water flows. The radiation dose to the SVX is reduced by the use of masks and an elaborate sliding-shield mechanism.

The change in IR-region materials between CLEO II and CLEO II.V affects the probability of multiple scattering and external Bremsstrahlung emission. Since this analysis depends on the Monte Carlo simulation for detection efficiencies, when generating Monte Carlo events we have generated them in the correct dataset ratios, not only between CLEO II and CLEO II.V but also in the datasets within each.

VD

The intermediate Vertex Detector extends from 8.1 cm to 16.4 cm radially. 800 sense wires are arranged in 10 layers; 2272 field wires form hexagonal cells around the sense wires, as shown in Figure 2.12. The wires are parallel to the beamline; however, the sense wires, which are made of nickel plated Cr-alloy, have sufficient resistivity to enable a z measurement by comparing the measured currents at both sides of the wire. Cathode strips, attached to the cylindrical surfaces of the drift chamber, attract and detect induced charges from electron avalanches on the sense wires nearest to them. These strips are segmented in z, as shown in Figure 2.12. The gas used is a 50-50 argon-ethane mixture.

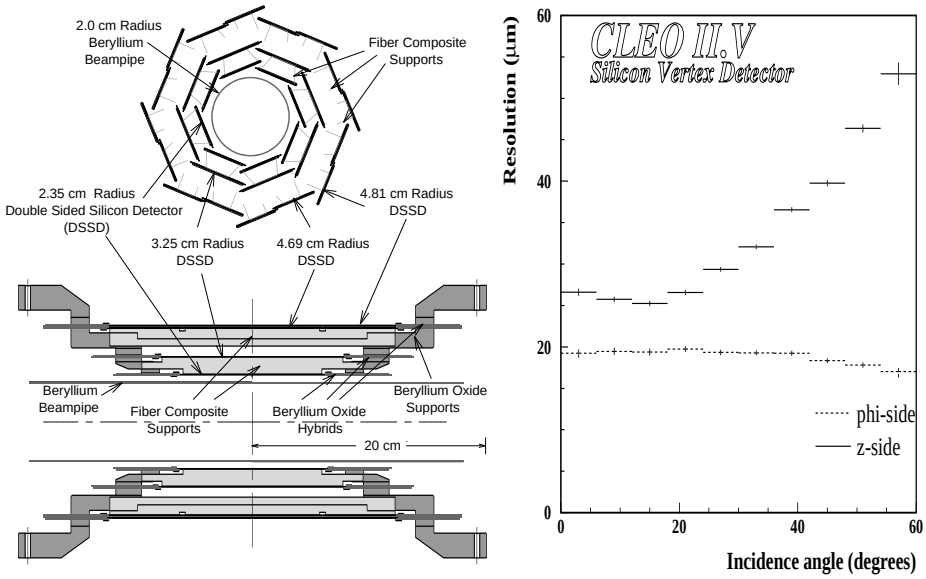

Figure 2.11: Left: The CLEO II.V Silicon Vertex Detector (SVX). Right: SVX resolution for tracks, as a function of the angle between the track and the plane perpendicular to the beamline.

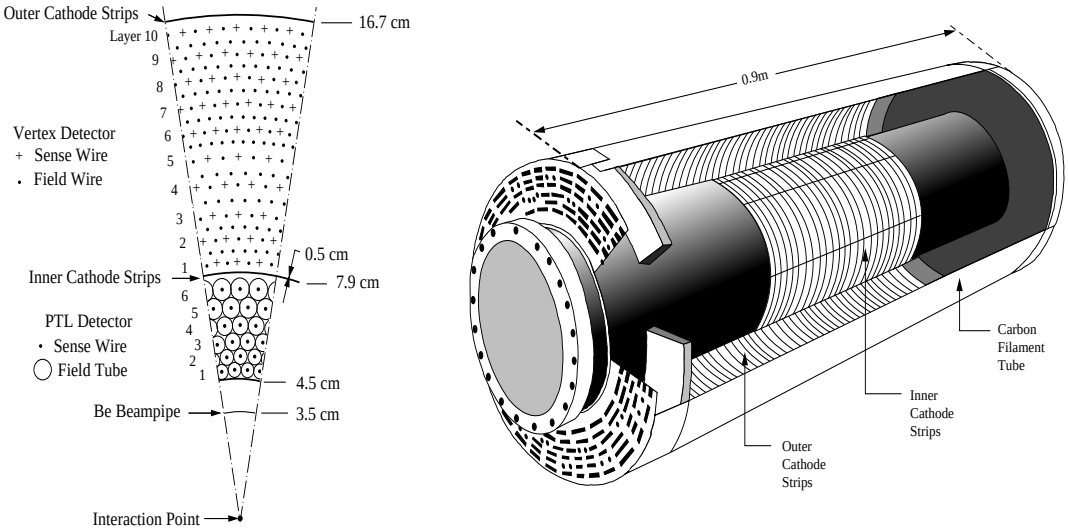

Figure 2.12: The inner two tracking layers of CLEO II: PTL and VD.

DR

The Drift Chamber extends from 19 cm to 91 cm, and uses 12,240 sense wires and 36,240 field wires arranged in a 51-layer pattern. The DR uses cathode pads in a similar manner as the VD (see Figure 2.13). Eleven of the 51 layers include a stereo angle of between 4 and 7 degrees (the angle increases with radius), resulting in p_z information.

In CLEO II an 50-50 argon-ethane gas mixture was used. In CLEO II.V this changed to 60-40 helium-propane. The modification led to a reduction in multiple scattering and a simplification of the the isochrone shape at the outer part of the cell, which in turn led to an increase in the charge collection efficiency.

When reconstructing leptons from the decay $B \to \psi \to \ell\ell$, we consider leptons ranging from about $0.6 - 2.4$ GeV. Results for tracking parameter resolutions [52] [53], for tracks generated in this momentum region, generated perpendicularly to the beampipe, are listed in Table 2.2:

Tracking Error	CLEO II	CLEO II.V
$\frac{\delta p}{p}$ (%)	$0.5 - 0.7$	$0.5 - 0.6$
σ_ϕ (mrad)	$0.6 - 2.0$	$0.7 - 2.5$
σ_θ (mrad)	$2.5 - 4.1$	$1.3 - 2.8$
σ_{xy} (μm)	$50 - 120$	$25 - 70$
σ_z (μm)	1000	$40 - 80$

Table 2.2: Errors on track parameters for the momentum region $0.6 - 2.4$ GeV, for tracks perpendicular to the beamline.

2.2.3 Time of Flight Counters

Surrounding the tracking devices are fast plastic scintillators which serve as event triggers and particle ID devices. The barrel and endcap counters are depicted at the top and left of Figure 2.15; 64 long thin scintillators are arranged in a barrel shape around the DR, and 28 wedge-shaped scintillators cover each endcap region of the detector. Photomultiplier tubes are connected to the light pipes.

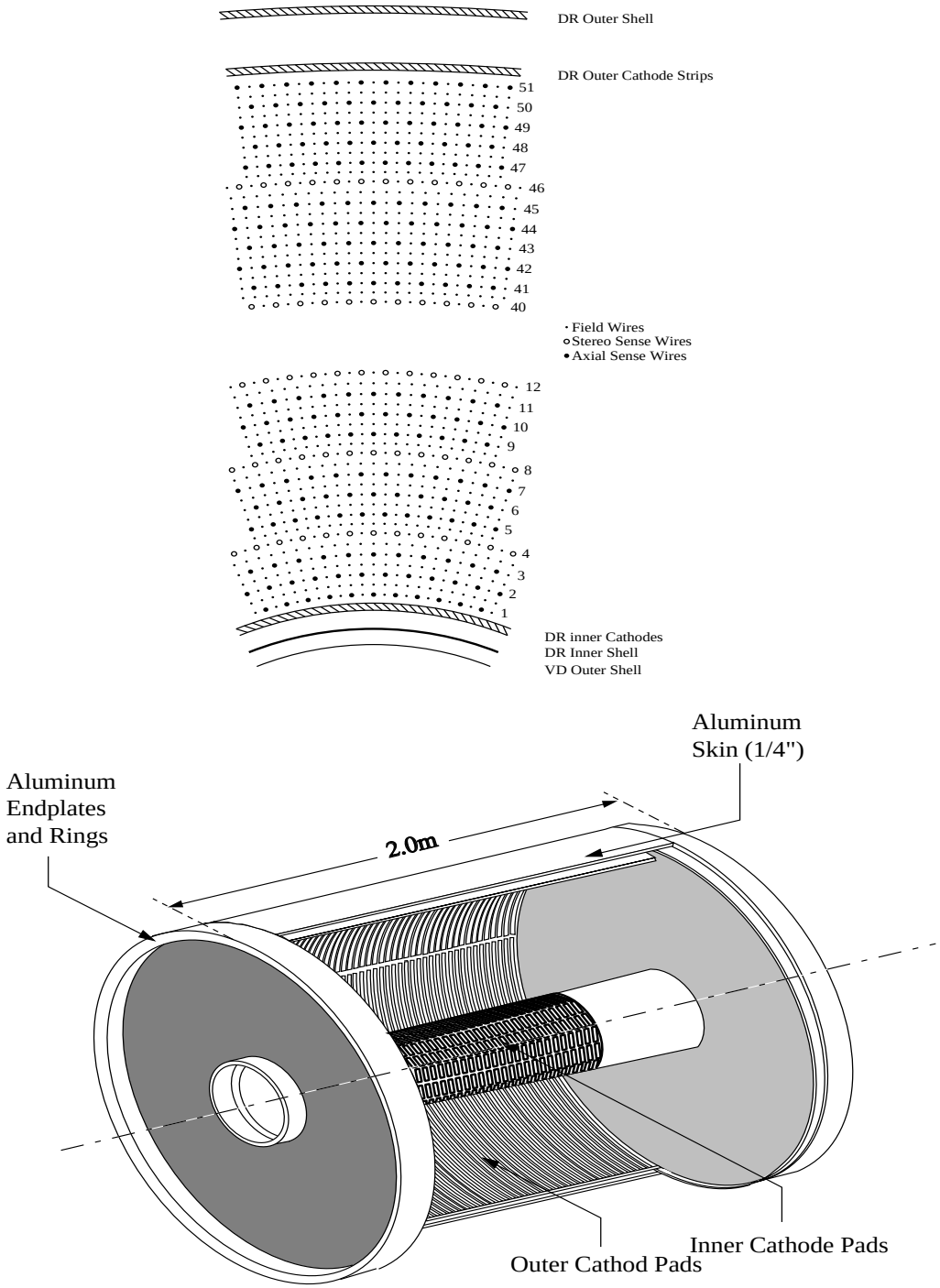

Figure 2.13: Main Drift Chamber (DR) wire configuration and cathode strip locations.

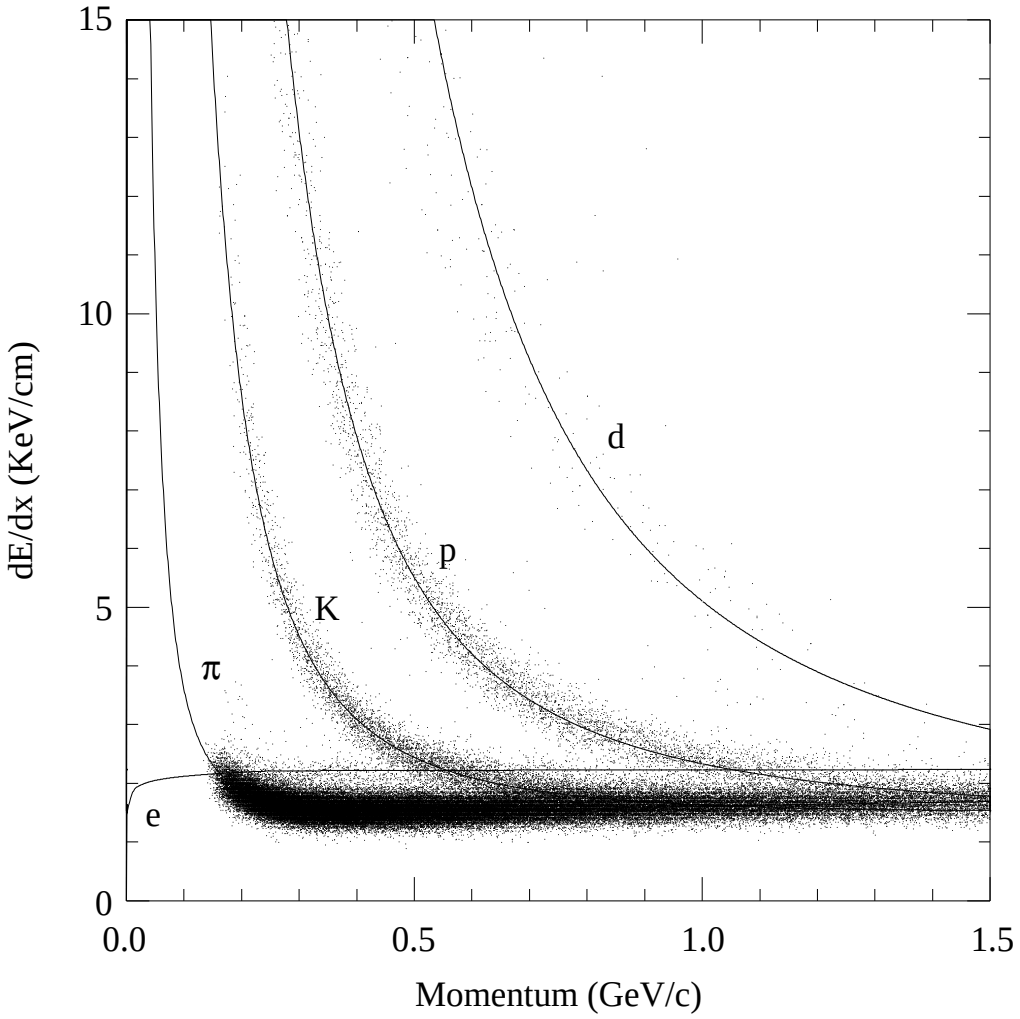

Figure 2.14: Measured dE/dx and predicted curves for each particle species.

For each scintillator, the readout electronics measure the time elapsed between the CESR beam crossing signal and a detected light signature. In the case of the barrel scintillators, the difference in timing between the east and the west side provide information in where the scintillation occurs in z. Since, for a given track momentum, the velocity of the track depends on its mass, the Time-of-Flight measurement can provide particle ID information as long as the TOF hit corresponds to a unique track. Use of the TOF as particle ID gives results that are shown in Figure 2.15.

2.2.4 Time of Flight Bunchfinder

As discussed previously, part of CESR's luminosity upgrade program was to introduce a beam crossing angle, thus enabling "multibunch" running. In this scenario, CESR fills nine "trains" of electrons and positrons, each train containing five buckets spaced 14 ns apart, with an actual bunch in any or all of the train buckets. Only one crossing signal is given for the entire train. This raises a problem for the tracking algorithm, which must know which bunch the event came from in order to establish a correct base time. Otherwise, the track fitter attempts to fit to isochrones that are too large or small, resulting in poor fits.

A contribution by the author to CLEO II was to write, calibrate, and maintain the "Time-of-Flight Bunchfinder," or TFBN, an algorithm which utilizes information from each TOF hit in a given event to determine the bunch from which the event occurred. The raw material for making the decision is shown in the top of Figure 2.16, which shows the aggregate distributions for many events of double-bunch data.

Given an event, TFBN loops through every TOF hit. Based on the pulse height of the hit, endcap/barrel, east-west timing correlation information (for barrel hits), and event trigger type, an appropriate lookup table is called which then assigns appropriate weight to the hit based on whether the measurement of the time supports any or all of the allowed bunch answers. The algorithm not only reports a bunch prediction with 100% efficiency, but also reports the amount of TOF evidence and the solidarity of that evidence as well. In accuracy studies of the algorithm, an accuracy rate of 99.45% with 100% efficiency was found for hadronic events. Bhabha and μ

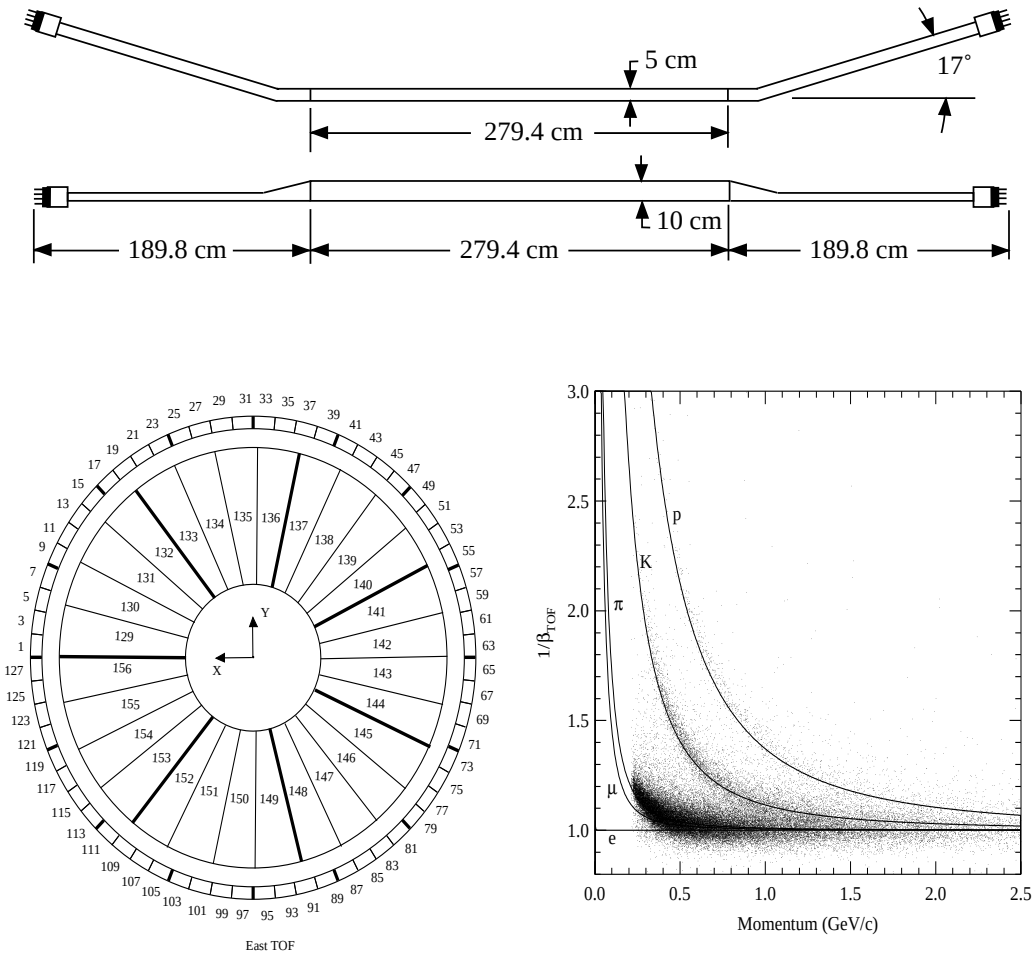

Figure 2.15: Time-of-Flight system. Top: A Barrel counter, showing scintillator and photomultiplier tubes. Bottom, Left: End view of endcap counters, barrel counters, and the counter numbering system. Bottom, right: Time of Flight for particle ID: Measured $1/\beta$ and predicted curves for each particle species.

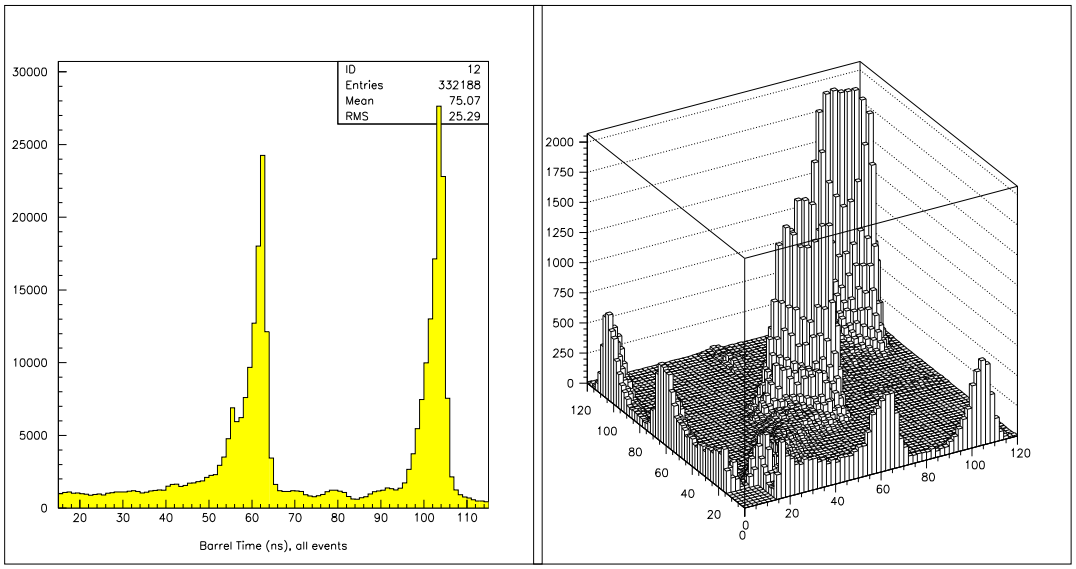

Figure 2.16: Top left: timing distribution of all TOF counters (horizontal scale is in ns), all pulseheights, many runs of double-bunch data. Top right: Raw timing measurements, double-bunch data. x and y axes are times reported by the east and west PMTs of a barrel scintillator. z axis is the number of events.

pair events, used for calibration, yielded 99.74% and 99.63% accuracy, respectively, with high efficiency. CLEO II.V relied exclusively on the TOF bunchfinder for on-line data acquisition. The bunchfinder was used until the TOF system was removed from CLEO II to make room for new RICH particle ID hardware for CLEO III.

2.2.5 Crystal Calorimeter

The crystal calorimeter (CC) is an array of 7800 thallium doped CsI crystals. Each crystal is approximately 5 cm $\times$ 5 cm $\times$ 30 cm, but in order to gain maximum coverage of 4π, 25 different shapes of crystals were fabricated. Figure 2.17 shows barrel and endcap crystal configurations.

Incident $e^{\pm}$ or γ on a crystal will pair-produce and undergo Bremsstrahlung; the resulting shower of numerous, low-energy $e^{\pm}$ and γ produces a flash of light in the crystal, the intensity of which is related to the energy of the original particle. Showers from $e^{\pm}$ or γ will typically be localized, i.e. not spread over a large region of crystals. The CC measures these particles with a high degree of precision in energy and angle.

Incident $\mu^{\pm}$ will pass through the crystal, leaving only ionization of about 200 MeV in the crystal. This fact is used in our ψ analysis to increase the detection efficiency in the $\psi \to \mu\mu$ mode. Hadrons, too, will leave ionization in the crystal; however, since the length of the crystals is approximately the same as the hadronic interaction length, the probability of a hadronic interaction is significant. Hadronic interactions will often cause the production of other hadrons, in unpredictable directions; often these secondary hadrons also deposit energy in nearby crystals. In general, hadronic showers cover a larger area of crystals than do electromagnetic showers.

Our analysis uses shower reconstruction in searching for $\chi_{c1} \to \psi(1S)\gamma$. From Monte Carlo, the spectrum of γ in this decay ranges from $240 - 600$ MeV. The energy and angular resolution of the CC is given generally in [51]; for light in the wavelength range we are interested in, the resolutions are listed in Table 2.3. The relatively poor performance of the endcap is because of extra detector material (preamps, DR endplate) separating the endcap crystals from the IP (see Figure 2.9).

Tracking Error	Barrel	Endcap
$\frac{\delta E}{E}$ (%)	$2.3 - 2.9$	$2.9 - 3.6$
σ_ϕ (mrad)	$5.5 - 7.6$	$12.1 - 14.9$
σ_θ (mrad)	$4.4 - 6.1$	$7.4 - 8.5$

Table 2.3: Errors on shower parameters for γ in the energy range $240 - 600$ MeV.

2.2.6 Magnet

The superconducting niobium solenoid carries a current of 3300 A and requires liquid helium for cooling. The 1.5 T magnetic field produced is uniform to within 0.2% in over 95% of the drift chamber volume. An iron yoke, 1 foot thick, is used to complete the magnetic field loop.

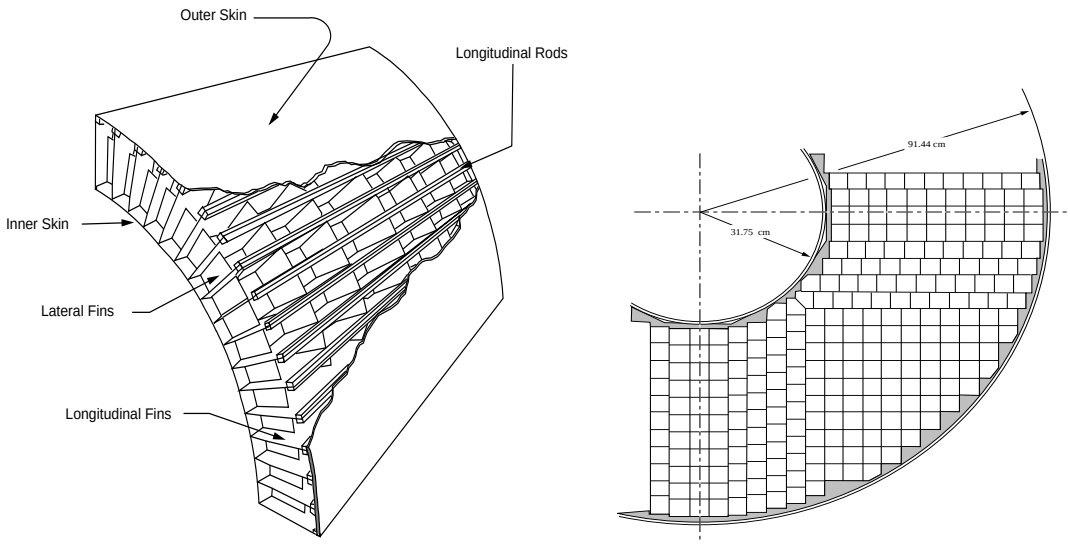

Figure 2.17: Crystal Calorimeter: Barrel and Endcap.

2.2.7 Muon Detectors

Of the observable particles, muons have the unique ability to penetrate through several feet of iron with significant probability. By placing iron between the interaction point and charged particle detectors, we hope that the charged particles that are detected are indeed muons. CLEO has three sets of "superlayer" detectors, which are embedded behind 3, 5, and 7 hadronic interaction lengths of iron (assuming a $\theta = \pi/2$ particle.) Each superlayer is made of three layers of proportional drift cells, shown in Figure 2.18. Each of the 8 counters includes an anode wire in the center, operates with the same gas as the DR (argon-ethane in CLEO II, helium-propane in CLEO II.V), and is coated with graphite on 3 sides to provide an electrical ground. The coordinate perpendicular to the wires is measured by the use of external copper pickup strips. Because of space limitations, the muon detector coverage, 85% of 4π, is less than that of the other CLEO detectors.

Selection of muon candidates at CLEO using the muon detectors is practically limited to one of three cuts. The CLEO variable `DPTHMU`, computed for every charged track, is the number of hadronic interaction lengths of steel that the track punched

through. Cuts of `DPTHMU` > 3, 5, or 7 are the standard cuts, in decreasing order of efficiency and hadronic punch-through rate. Our analysis combines pairs of oppositely charged muons and fits the resulting invariant mass peak at $M(J/\psi)$; hadrons that punch through will not peak in this invariant mass and therefore are not a concern. Maximizing efficiency, however, is crucial in a precision measurement like this one. From the lower plots of Figure 2.18 we see that even the loose requirement of `DPTHMU` $>$ 3 leads to a loss of all muons softer than 0.9 GeV in the barrel and 1.5 GeV in the endcap. The cuts used in the analysis require `DPTHMU` > 3 in the barrel region but loosen the endcap requirements in order to increase the detection efficiency.

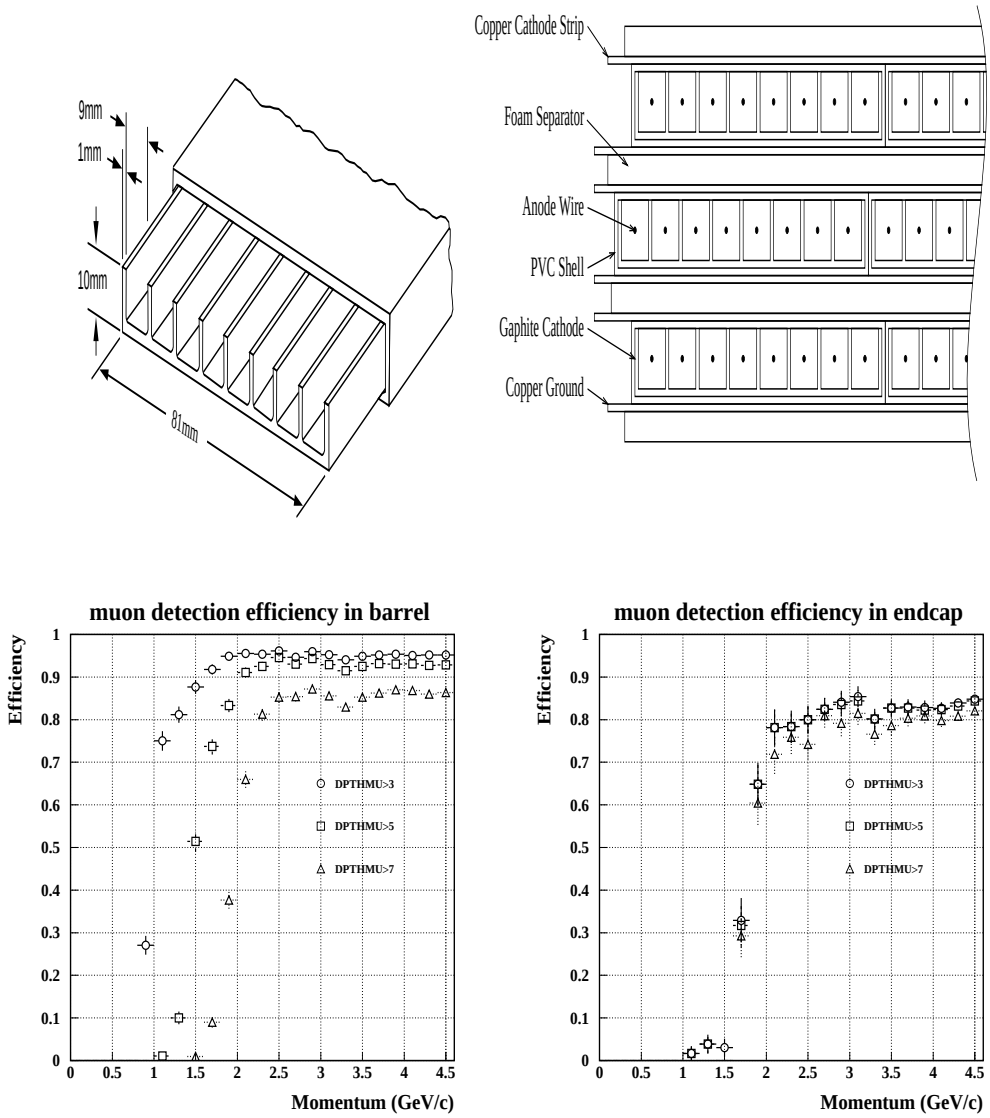

Figure 2.18: Muon detectors at CLEO. Upper Left: cross section of a single detector layer. Upper Right: a superlayer (3 layers). Lower Left: efficiencies of the three standard CLEO cuts for muons, in the barrel region. Lower Right: endcap region.

Chapter 3

Experimental Techniques for Charmonium Measurement

3.1 Obstacles to Theoretically Relevant Measurements

The theoretical framework that was discussed in Chapter 1 makes predictions for the inclusive branching fraction, momentum distribution, and polarization of ψ which are produced directly from B mesons. Measurement of these very basic properties of ψ mesons might appear at first to be a simple matter. However, there are subtleties which must be accounted for in order to obtain measurements that measure exactly what theory predicts. These potential pitfalls are discussed in this section.

3.1.1 Reference Frames: B vs. Laboratory

There are two inertial reference frames relevant to our measurement. One is the reference frame in which the B meson (which subsequently decays to ψX) has zero momentum. We refer to this as the "B frame." The second is the CLEO lab frame, or "$\Upsilon(4S)$ frame," in which the the $\Upsilon(4S)$ (which subsequently decays into $B\overline{B}$) is

created with zero momentum[1]. We note that the relation between these two reference frames is different for every event.

Ideally, we wish to measure the momentum of the ψ in the B frame, as this is the relevant frame for theoretical calculations[2]. We now catalog some ways in which attempts might be made to determine the direction of this boost. We cannot reconstruct the entire B decay chain as this is an inclusive measurement. Reconstructing the other B in the event would lead to an unacceptable loss in efficiency due to the small branching fraction of exclusive B decay modes. If we could reconstruct the decay vertex of the ψ as well as that of the interaction point, this would give us the direction of the initially produced B. However, $\beta\gamma$ of the B is approximately 0.06, so that $\beta\gamma c\tau(B^+) = 30\mu\text{m}$, which is significantly less than the best-case track impact parameter resolution of the CLEO II.V SVX [54]: $\delta = \delta_{r\phi} \oplus \delta_z \geq 19\mu\text{m} \oplus 42\mu\text{m} = 46\mu\text{m}$ (see Figure 2.11).

Since it is the best we can do, our measurements of p_ψ and $\cos\theta$ (the helicity angle of the ψ, as defined in Sec. 1.6.4) are made in the lab frame. However, by definition, measurement of the helicity angle θ, requires knowledge of p_ψ^B (or more precisely, the direction of $\vec{p}_\psi^B$)[3]. Since we can only measure p_ψ^{lab}, our calculation of θ, which substitutes p_ψ^{lab} for p_ψ^B, is incorrect. We will use θ^{lab} to denote the (smeared) helicity angle thus obtained. We will refer to this phenomenon, which causes $p_\psi^{lab} \neq p_\psi^B$ and $\theta^{lab} \neq \theta$, as "B-frame smearing." The effect of B-frame smearing on the quantities p_ψ and θ are shown in Figure 3.1.

Theoretical predictions for the ψ momentum distribution are made in the in the b frame, but our measurement is in the $\Upsilon(4S)$ frame. The transformation between these frames must be done in order to make comparisons between theory and experiment.

[1]For most of the data, this is not exactly true, due to the CESR crossing angle scheme, as discussed in Section 2.1.3. However, the boost of the $\Upsilon(4S)$ ($\beta \approx 0.004$) is small relative to the boost of the B mesons which the $\Upsilon(4S)$ decays into ($\beta \approx 0.06$), making the former boost negligible.

[2]Actually, the frame in which the b *quark* is at rest is the ideal reference frame; however, properties of the b are not directly measurable by experiment. The b frame and B frame are not trivially related, as discussed in Section 1.5. Therefore the normal procedure is for the phenomenologists to adjust their b-frame predictions to the B frame, instead of experimentalists attempting to extract b-frame results from B frame measurements.

[3]When it is important to be explicit about the reference frame, we will use p_ψ^B and p_ψ^{lab} to refer to the ψ momentum in the B frame or lab frame, respectively.

One approach is to "smear" the b-frame model predictions into the B or $\Upsilon(4S)$ frames. This is precisely what is done in the momentum spectrum predictions [37] [38] that were summarized in Section 1.6.3.

The other approach is to attempt to recover the p_ψ^B distribution from the measured p_ψ^{lab} distribution. We are indeed able to perform this deconvolution, but only when the bin size is significantly larger than the smearing width. Because of the coarse binning required to do this (see Fig. 3.1, upper right plot), we will not present results of it. Our final results for the momentum spectrum are presented in the $\Upsilon(4S)$ frame.

As far as the polarization measurement is concerned, the lower right plot of Figure 3.1 is much less of a problem than it might first appear to be. This is because our method of measuring the polarization, described later, compensates for the θ mismeasurement in a simple and elegant manner which does not require that we attempt to deconvolve θ from θ^{lab}.

For the remainder of this thesis, references to momentum and the helicity angle θ will refer to measurements made in the *lab* frame, unless explicitly specified otherwise.

3.1.2 Mismeasurement due to Bremsstrahlung and Detector Resolution

There is no such thing as a perfect detector; every measurement of every quantity will have an error associated with it that is intrinsic to the process of measurement. Figure 3.2 shows Monte Carlo simulations of the reconstructed $\psi(1S)$ momentum compared against the the generated momentum, where both momenta are in the lab frame. We will use the phrase "measurement smearing" to describe the difference between the generated and reconstructed ψ momenta.

The difference in resolution between the electron and muon decay modes is primarily due to Bremsstrahlung. In the di-electron mode, relative to the di-muon mode, there is a greater probability that we will fail to detect and add Bremsstrahlung photons to the ψ candidate's four-vector, thus leading to the familiar radiative tail as well as a widening of the signal peak, in invariant mass. If, for example, we require $\psi(1S) \rightarrow e^+e^-$ candidates to have an invariant mass within 36 MeV (2σ) of the

nominal mass, thus excluding events in the radiative tail, the momentum resolution σ improves from 38 MeV (see Figure 3.2, lower left plot) to 15 MeV.

The final step of our ψ momentum measurement, which will be discussed in detail in Section 4.4, is to recover the generated ψ distribution from the measurement-smeared ψ distribution. Our procedure for measuring the polarization automatically takes care of measurement smearing in the same manner as it does for B-frame smearing.

3.1.3 Intermediate Parents

We are interested in measuring only those ψ which are produced *directly* from B mesons. This requirement necessitates extra efforts in the case of $\psi(1S)$, but not for $\psi(2S)$.

Table 3.1 lists every mode in which a $\psi(1S)$ can be be produced from a beauty meson B.[4] There are several modes in which the $\psi(1S)$ is produced via an "intermediate parent," which we define as a particle, produced by the B, which in turn produces the $\psi(1S)$; e.g. χ_{c0}, χ_{c1}, χ_{c2}, and $\psi(2S)$. The last section of the table lists the two modes in which we detect the $\psi(1S)$.

The goal is to obtain measurements of momentum distribution and polarization for those $\psi(1S)$ that are the *direct* daughters of the B. To do this we must first examine the sample of $\psi(1S)$ without regard to whether or not there is an intermediate parent, i.e. the "*inclusive*" $\psi(1S)$. We must then correct this inclusive sample for those $\psi(1S)$ (the *"feed-down"*) which have intermediate parents. Of the several $\psi(1S)$ feed-down modes, we explicitly reconstruct two: $\psi(2S) \to \psi(1S)\pi^+\pi^-$ and $\chi_{c1} \to \psi(1S)\gamma$. In the case of $\psi(2S) \to \psi(1S)\pi^0\pi^0$, we assume that the $\psi(1S)$ from this decay will have the same momentum distribution and polarization as the $\psi(1S)$ from $\psi(2S) \to \psi(1S)\pi^+\pi^-$. We will discuss this assumption in further detail in subsequent sections. As for the four decay chains involving an intermediate $\psi(2S)$ with relatively small branching fractions (we will refer to these as the "miscellaneous" feed-down modes), we use a Monte Carlo simulation of these decays to obtain the momentum distribution

[4] "B" denotes the admixture of B^0, $\bar{B}^0$, B^+, and B^- produced from decays of the $\Upsilon(4S)$).

Mode	$\mathcal{B}$ (10^{-3})	Measure?
$B \to \psi(1S)X$	11.5 ± 0.6	Yes
$B \to \psi(1S)(\text{direct})\ X$	8.0 ± 0.8	No - correct inclusive
$B \to \chi_{c1}X \to (\psi(1S)\gamma)X$	1.15 ± 0.20	Yes
$B \to \psi(2S)X \to (\psi(1S)\pi^+\pi^-)X$	1.22 ± 0.20	Yes
$B \to \psi(2S)X \to (\psi(1S)\pi^0\pi^0)X$	0.71 ± 0.12	No - use $\psi(1S)\pi^+\pi^-$
$B \to \psi(2S)X \to (\psi(1S)\eta)X$	0.08 ± 0.02	No - MC simulation
$B \to \psi(2S)X \to (\chi_{c2}\gamma)X \to (\psi(1S)\gamma\gamma)X$	0.037 ± 0.007	No - MC simulation
$B \to \psi(2S)X \to (\psi(1S)\pi^0)X$	0.003 ± 0.001	No - MC simulation
$B \to \psi(2S)X \to (\chi_{c0}\gamma)X \to (\psi(1S)\gamma\gamma)X$	0.002 ± 0.001	No - MC simulation
Sum of the above:	11.2 ± 0.9	
Reconstruction Modes:		
$\psi(1S) \to e^+e^-$	59.3 ± 1.0	Yes
$\psi(1S) \to \mu^+\mu^-$	58.8 ± 1.0	Yes

Table 3.1: $B \to X \to \psi(1S)$: Every non-upper-limit decay chain from the 2001 PDG [10]. The last column shows whether or not we directly measure the contribution of $\psi(1S)$ from a particular channel. If the answer is "No," then the method by which we deduce the momentum distribution of $\psi(1S)$ from that particular channel is also listed.

of the $\psi(1S)$ created from these decays.

Table 3.2 shows the known decay chains which produce a $\psi(2S)$ from $\Upsilon(4S)$. In contrast to the case with $\psi(1S)$, the situation here is much simpler as there there is no "feed-down." We reconstruct the $\psi(2S)$ via the $\psi(1S)\pi^+\pi^-$ mode as well as the dilepton modes.

Mode	BF (10^{-3})	Measure?
$B \to \psi(2S)X$	3.5 ± 0.5	Yes
Reconstruction Modes:		
$\psi(2S) \to \psi(1S)\pi^+\pi^- \to \ell^+\ell^-\pi^+\pi^-$	41.1 ± 3.3	Yes
$\psi(2S) \to e^+e^-$	7.9 ± 0.5	Yes
$\psi(2S) \to \mu^+\mu^-$	12 ± 4	Yes

Table 3.2: $B \to X \to \psi(2S)$: Every non-upper-limit decay chain from the 2001 PDG [10].

3.2 Charmonium Reconstruction: Overview

We detect both the $\psi(1S)$ and $\psi(2S)$ in the $\psi \to \ell\ell$ dilepton channel, which offers a clean experimental signature for both the invariant mass and helicity of ψ candidates. The $\psi(1S)$ and $\psi(2S)$ mesons decay primarily into hadronic states; however, the knowledge of the branching fractions involved is relatively poor, the signal-to-background level for multibody hadronic final states is high, and the helicity information of the parent ψ is obscured.

Determining the number of ψ which come directly from B decays requires several steps. We first reconstruct the ψ candidates in the on-resonance data, obtaining a ψ yield from fitting invariant mass peaks. We do the same procedure for off-resonance data. Subtracting the latter, properly scaled, from the former, we obtain a yield of ψ which occur from $\Upsilon(4S)$ decays. By assuming that $\Upsilon(4S) \to B\bar{B} = 100\%$, as is standard for "admixture" analyses [10], we conclude that every ψ from $\Upsilon(4S)$ also came from a B. In this vein we note that a recent BaBar measurement [55] reports $\mathcal{B}(\Upsilon(4S) \to \psi X) < 4.3 \times 10^{-4}$ at 90% C.L.

In the case of $\psi(1S)$, we combine the reconstructed $\psi(1S)$ candidates with other particles in the event to determine if the $\psi(1S)$ had an intermediate parent. We add the four-momenta of γ candidates to that of the $\psi(1S)$ candidates to look for the intermediate parent χ_{c1}; similarly, we add the four-momenta of observed $\pi^+\pi^-$ candidates to the $\psi(1S)$ four-momentum to look for the intermediate parent $\psi(2S)$. We note that this process will not identify 100% of the $\psi(1S)$ with intermediate parents due to inefficiency in finding the extra particles. We also note that we must have branching fraction information for the intermediate parent particles. These issues are taken care of in the analysis.

This leads to three sets of $\psi(1S)$ candidates. First, is the "inclusive" set, in which no attempt is made to ascertain the immediate parent of the ψ candidate. The other two sets are subsets of the "inclusive" set in which the ψ candidate is also an daughter-of-an-intermediate-state candidate.

- "$\psi(1S)$ inclusive": $\psi(1S) \to \ell^+\ell^-$

- "$\psi(1S)$ from $\psi(2S)$": $\psi(2S) \to \psi(1S)\pi^+\pi^- \to \ell^+\ell^-\pi^+\pi^-$
- "$\psi(1S)$ from χ_{c1}": $\psi(1S)\gamma \to \ell^+\ell^-\gamma$

As for the $\psi(2S)$, there is no feed-down to worry about; however, we reconstruct it in two different final states:

- "$\psi(2S)$ via dileptons": $\psi(2S) \to \ell^+\ell^-$
- "$\psi(2S)$ via $\psi\pi\pi$": $\psi(2S) \to \psi(1S)\pi^+\pi^- \to \ell^+\ell^-\pi^+\pi^-$

The "$\psi(2S)$ via $\psi\pi\pi$" mode is the identical set of events as the "$\psi(1S)$ from $\psi(2S)$" mode. The reason for the semantic difference is that when we say "$\psi(1S)$ from $\psi(2S)$," we are interested in the properties (momentum, polarization) of the $\psi(1S)$; adding particles to this $\psi(1S)$ to form a $\psi(2S)$ is done only to see if the $\psi(1S)$ had an intermediate parent. In contrast, when we say "$\psi(2S)$ via $\psi\pi\pi$," we are interested in the properties of the $\psi(2S)$. In this case the polarization of the $\psi(2S)$ cannot be easily determined from examining the decay products $\psi\pi\pi$. So, this mode is not included in the $\psi(2S)$ polarization measurement.

3.3 Bremsstrahlung Recovery

In the decay $\psi \to \ell^+\ell^-$, either or both of the final state particles may emit Bremsstrahlung radiation. "Internal" Bremsstrahlung is due to the acceleration experienced by an electron when it is first created, while "external" Bremsstrahlung comes about due to the electron's passage through matter in the detector. The ψ decays we reconstruct are therefore better described as $\psi \to \ell^+(\gamma)\ell^-(\gamma)$. If the four-momenta of Bremsstrahlung photons is ignored, then the invariant mass of the $\ell^+\ell^-$ will be lower than the nominal mass M_ψ. The problem is much worse for the electron mode than it is for the muon mode, because the lower mass of the electron increases the probably of Bremsstrahlung emission.

The rest of this Section is a summary of the findings from a study of Bremsstrahlung photon recovery at CLEO; more details may be found in Ref. [56].

At CLEO, most of the matter encountered by the daughter leptons is concentrated in the beampipe and SVX, both of which are located near the interaction point (IP). The spatial difference is small enough that internal and external Bremsstrahlung may be treated identically.

In the case of $\psi \to \ell^+(\gamma)\ell^-(\gamma)$, almost all Bremsstrahlung photons are emitted within a narrow (5 degree) cone around the momentum vector (at the IP) of the emitting electron. By searching for showers in the crystal calorimeter within this cone that are not are not associated with charged tracks, we identify and add up to one shower per electron track to the ψ candidate four-momentum. The increase in efficiency by doing so is on the order of 25%.

In the case of $\psi \to \mu^+(\gamma)\mu^-(\gamma)$, the probability of Bremsstrahlung photon emission is lower; when photons are emitted, they are far less collinear with the IP muon momentum. Attempting the above technique with a correspondingly larger cone results in a large increase in the number of wrongly identified photons. Photon recovery is therefore not used for the $\psi \to \mu + \mu^-$ case.

3.4 Selection Criteria

The general motivation behind the ψ selection criteria is to maximize the ψ detection efficiency. Lepton identification cuts are loose, and no part of the detector geometry is excluded. These cuts are similar to those used in other CLEO charmonium analyses [56].

3.4.1 Dataset

We use the full datasets collected by the CLEO II and CLEO II.V detectors. Most CESR running occurs at a center-of-mass beam energy of $m_{\Upsilon(4S)}$, or 10.580 GeV; at this energy, hadrons are created via $e^+e^- \to b\overline{b}$ in addition to $e^+e^- \to q\overline{q}$, where $q \in \{u, d, c, s\}$. One-third of the luminosity is devoted to "continuum" data-taking, where the beam energy is tuned to about 10.550 GeV, thus excluding production of the $\Upsilon(4S)$. Many B analyses have significant backgrounds from $e^+e^- \to q\overline{q}$ events; the

purpose of obtaining the continuum dataset is to subtract the contribution (properly scaled for luminosity and energy scale) of $e^+e^- \to q\overline{q}$ from the on-resonance data.

For CLEO II data, the integrated luminosity of the data is 3.137 fb^{-1} on-resonance and 1.608 fb^{-1} continuum; for CLEO II.V, 6.029 fb^{-1} on-resonance and 2.944 fb^{-1} continuum are analyzed.

3.4.2 Events

Events must be deemed to be a hadronic event by the CLEO event classifier. Over 99.8% efficient for $B\overline{B}$ events [57], it requires that at least 3 charged tracks are observed, that the detected energy of all tracks and showers must be greater than 20% of the beam energy, and that either at least 5 charged tracks are observed, or the total energy deposited in the crystal calorimeter is more than 15% of the beam energy.

Additionally, we require that the ratio of the second to zeroth Fox-Wolfram moments [58] `R2GL`[5] be less than 0.5 (thus selecting events in which the energy is more spherically distributed), and that at least 4 tracks are observed.

3.4.3 Tracks

Many of the tracking cuts are fairly typical for CLEO analyses. The common criteria for all tracks, along with specific requirements for $e^\pm$ and $\pi^\pm$ candidates, are listed in Table 3.3.

Electron ID information is reported for each track in the continuous variable `R2ELEC`. Based on information from the crystal calorimeter, dE/dx, tracking, and time-of-flight systems, `R2ELEC` is designed to give a log-likelihood probability that the track in question is an electron. Of the tracks that satisfy the requirement `R2ELEC` $>$ 3, for example, approximately one track out of every 10^3 will not be an electron. Our requirement of `R2ELEC` > 0 is therefore loose relative to the cut used in most CLEO analyses (`R2ELEC` > 3), and is designed to maximize the detection efficiency

[5]Uppercase `TYPE` font will be used to denote standard CLEO analysis variables. Lowercase `type` font will denote variables defined specifically for this analysis.

of electrons. We can afford to allow a higher percentage of "fake" electrons, since we reconstruct ψ candidates by combining *pairs* of oppositely charged electrons. ψ candidates reconstructed from one or two hadrons that happen to pass the electron-ID cuts will not peak in m_ψ. The same logic leads to our choosing loose muon-ID requirements in the interest of maximizing detection efficiency.

The selection criteria which categorize the muon candidates into three "tiers" are listed in Table 3.4. The primary variable for muon identification is `DPTHMU`, which is the number of nuclear interaction lengths of iron traversed by a track before being detected by one of the muon detectors. The first layer of iron, beyond which lie the first set of muon detectors, is 3 interaction lengths thick, so the loosest cut possible for muon detection is 3. Tracks which pass this cut are called "first-tier" muon candidates.

As the tracks become more parallel to the beamline, they must traverse more material before even reaching the iron, as shown in Figures 2.8 and 2.9. This effectively increases the minimum momentum that muon candidates must have to punch through the iron. We therefore define "second-tier" muons as being charged tracks that go in the direction of the endcap muon chamber, do not have enough momentum to be likely to punch through, and leave a minimum ionizing shower in the crystal calorimeter. Similarly, "third-tier" muons are tracks leaving a minimum ionizing shower in the calorimeter and point to the extreme endcap region where there is no muon chamber coverage.

3.4.4 Calorimeter Showers

Showers in the crystal calorimeter are used for three purposes: 1) identify Bremsstrahlung photons, 2) identify showers consistent with a minimum ionizing particle (muon), and 3) identify photons from the $\chi_{c1} \to \psi(1S)\gamma$ decay. Requirements for each type of shower are listed in Table 3.5.

3.4.5 $\psi \to \ell^+\ell^-$

In the $\psi \to e^+(\gamma)e^-(\gamma)$ mode, the γ candidates must be Bremsstrahlung Showers as defined in Table 3.5. Furthermore, they must lie within a 5-degree cone of either of the e tracks. If more than one shower is within this cone, then the shower lying in the smaller cone is chosen. No one shower may be added twice, even if it meets the above conditions for both electron tracks. The net result is that 0, 1, or 2 Bremsstrahlung photons are added to ψ candidate four-momentum.

In the $\psi \to \mu^+\mu^-$ mode, at least one of the muon candidates must be a first-tier muon candidate as defined in Table 3.4. Both must be either first-, second-, or third-tier muon candidates.

3.4.6 $\psi(2S) \to \psi(1S)\pi^+\pi^-$

We impose a cut on the invariant mass of the $\psi(1S)$ candidate -50 MeV $<$ $M(\ell^+(\gamma)\ell^-(\gamma)) - M_{\psi(1S)} < +25$ MeV. This includes the signal peak in the distribution of $M(\ell^+(\gamma)\ell^-(\gamma))$ but does not include the radiative tail. Excluding the latter results in a loss of efficiency, but is offset by the more accurate measurement of the ψ four-momentum and an increase in the signal-to-background ratio.

The distribution of $M(\pi^+\pi^-)$ does not follow three-body phase space but peaks towards higher $M(\pi^+\pi^-)$, as observed by the MARK III collaboration in 1992 [59]. We impose a cut of $M(\pi^+\pi^-) > 0.45$ GeV, which has a signal efficiency of about 85%, and significantly reduces combinatoric background from pairs of soft pions. Figure 3.3 contrasts the 3-body phase space distribution to the MARK III observed distribution.

When searching for a signal in this decay mode we plot the mass difference between the $\psi(2S)$ and $\psi(1S)$ candidates: $(M_{\psi(2S)} - M_{\psi(1S)})$ – or, to be more precise, $(M_{\ell^+(\gamma)\ell^-(\gamma)\pi^+\pi^-} - M_{\ell^+(\gamma)\ell^-(\gamma)})$. Doing so partially discounts any mismeasurement of the $\psi(1S)$ candidate and therefore narrows the signal.

3.4.7 $\chi_{c1} \to \psi(1S)\gamma$

This mode is treated similarly to the $\psi(2S) \to \psi(1S)\pi^+\pi^-$ search mode: we accept $\psi(1S)$ candidates within [−50 MeV, +25 MeV] of the nominal $\psi(1S)$ mass, and plot the mass difference $(M_{\chi_{c1}} - M_{\psi(1S)}) = (M_{\ell^+(\gamma)\ell^-(\gamma)\gamma} - M_{\ell^+(\gamma)\ell^-(\gamma)})$ to search for χ_{c1} candidates.

The photon candidate must satisfy the "Isolated Shower" cuts of Table 3.5, and must be detected in the "Good Barrel" region of the detector. Finally, the shower must pass a π^0 veto; namely, the two-photon invariant mass of the candidate shower, combined with any other "Isolated Shower" in the same event, must not be within $[-3\sigma, +2\sigma]$ of the nominal π^0 mass.

3.5 Data Yields

The results of applying the selection criteria to the combined CLEO II and CLEO II.V dataset are shown in Figure 3.4 (on-resonance), Figure 3.5 (continuum), and Table 3.6. Raw (uncorrected) yields are obtained by fitting the mass plots with a signal lineshape and cubic polynomial background. The details of the fitting procedure will be described later.

The purpose of this exercise is to obtain an estimate of the yield in the entire dataset; this indicates how well we might be able to measure the momentum distribution and polarization, both of which will require us to divide the dataset into many partitions, and then fit for the yield of ψ in each partition.

For the momentum distribution, we will partition the dataset into bins of p_ψ and obtain a yield from fitting the mass peak in each partition. The momentum distribution is then a plot of the yields in each partition. Similarly, to measure the polarization, we will simply define the partitions differently – namely, in $\cos\theta$ – and again fit the mass peaks in each partition. Plotting the yield as a function of $\cos\theta$ results in an angular distribution. By comparing the observed distribution to simulations of longitudinally polarized ψ and transversely polarized ψ, we obtain the average polarization of ψ in the data.

In the continuum data, the yields of $\psi(2S)$ and feed-down $\psi(1S)$ are consistent with zero, so we will proceed assuming that their effect in this analysis is negligible. In the inclusive $\psi(1S)$ mode, we observe a small but significant signal. The analyses will use these events to correct for the contribution of non-$b\bar{b}$ events in the on-resonance data.

3.6 Signal Monte Carlo

In order to estimate the detection efficiency and determine signal lineshapes as a function of ψ momentum and helicity, we rely on a Monte Carlo simulation of $\Upsilon(4S)$ decays and their observation by the CLEO detector. Two programs are used to generate simulated data files, which are of the same file format as the real data, and are processed through the same reconstruction code as the data.

The first program is called `QQ`; it uses a random number generator to select a complete decay chain of the $\Upsilon(4S)$, down to the stable particles that are detected. A lookup table contains the definitions of the particles, their spins and masses, branching fractions, and probability of internal Bremsstrahlung, all of which define the kinematics of the decays. Fragmentation processes cannot be handled with a lookup table, so for this the CERN program JETSET [60] is used. The end result is a list of particles with explicitly listed 4-momenta, creation coordinates, and decay coordinates.

This particle list is an input to the second program, `CLEOG`, which simulates the CLEO detector response to particles and generates simulated data files of the same format as the real data. The GEANT [61] package from CERN is used for the detector response, modelling a large number of physics processes such as dE/dx loss, external Bremsstrahlung, electromagnetic showers in the crystals, multiple scattering of tracks, penetration of charged particles through iron, etc. Detector noise is also modeled, as are the decays of "long-lived" particles such as muons and kaons.

Over the approximately ten years in which CLEO II and CLEO II.V collected data, the detector went through many changes, major and minor. Some changes had important implications for the detector response. For example, the introduction of the SVX added significant amounts of material to the detector, thus increasing the

probabilities of Bremsstrahlung and multiple scattering near the IP. Other examples of changes were the changing of the drift chamber gas to helium-propane (implications for tracking and dE/dx) and the switch to multibunch running (introducing a tiny boost to the $\Upsilon(4S)$ in the lab frame). All of these changes, and many more, have been incorporated into the CLEO version of GEANT. For this analysis, we have taken care to generate CLEO II and CLEO II.V events in the same ratio as the luminosities of these data sets. In fact, we generate events in the same ratios of the "sub-datasets" within CLEO II and CLEO II.V, known within CLEO as the `4s2-G` (II) and `4sH-T` (II.V) datasets.

In order to obtain information about the specific decays that we reconstruct, we generate "Signal" Monte Carlo in these decay modes. This is done by "loading the dice" used by `QQ` to force the "produced" $\Upsilon(4S)$ to decay in a manner in which we prescribe. There are eight sets of signal Monte Carlo that are crucial to this analysis: 4 for each of the search modes, times 2 for each spin-alignment state of the ψ. All of the signal Monte Carlo begins with the decay $e^+e^- \to \Upsilon(4S) \to B\bar{B}$. We allow either the B or the $\bar{B}$ (50-50 probability) to decay via the standard `QQ` lookup table. The other B in the event decays is forced to decay in one of four ways, enumerated below:

- "$\psi(1S)$ direct": $B \to \psi(1S)X$.
- "$\psi(1S)$ from $\psi(2S)$, or $\psi(2S)$ via $\psi\pi\pi$": $B \to \psi(2S)X$, $\psi(2S) \to \psi(1S)\pi + \pi^-$.
- "$\psi(1S)$ from χ_{c1}": $B \to \chi_{c1}X$, $\chi_{c1} \to \psi(1S)\gamma$.
- "$\psi(2S)$ via dileptons": $B \to \psi(2S)X$.

For each of these decay modes, we generate two separate sets of Monte Carlo events, based on the helicity state of the ψ:

- $H = 0$ (or $\alpha = -1$)
- $H = \pm 1$ (or $\alpha = +1$)

In the dilepton decays, the $\psi(1S)$ and $\psi(2S)$ decay to both lepton species equally. In the above, "X" refers to any number of states, whose composition and relative probability are given by the standard `QQ` lookup table. (If X is a spin-one particle, for example, K^*, then its helicity is defined with the correct correlation, with respect to the helicity of the ψ.)

In the case of $\psi(2S) \to \psi(1S)\pi^+\pi^-$, the CLEO Monte Carlo generates the events in 3-body phase space. However, it is known from Mark III measurements [59] that the observed distribution of $M(\pi^+\pi^-)$ peaks strongly towards higher $M(\pi^+\pi^-)$, relative to phase space. The phenomenological formula for the observed distribution is given in [62] and is used to "pare down" the CLEO Monte Carlo to the observed distribution via the rejection method. (See Figure 3.3.)

Variable	Criteria	Description of Cut/Variable
		All Tracks
DREDGE	$= 0$	cuts tracks with noise hits
ZESCAPE	$= 0$	require tracks with sufficient z information
KINCD	$\in \{0, -2\}$	ensure well-measured tracks
TRKMAN	≥ 0	ensure 1-to-1 mapping of logical to physical tracks
|DBCD|	< 0.003	max. $r - \phi$ impact parameter (m), with respect to primary vertex
|Z0CD-ZVPTX|	< 0.025	z-distance between track and primary vertex (m); applies only if primary vertex in z is determined by 2 or more tracks with an error of less than 0.010 m
|Z0CD|	< 0.05	z-distance between track and expected IP (m); applies only if the above cut does not apply
ABSMOM	< 5.3	track momentum (GeV)
		Electron Candidates
R2ELEC	> 0.0	CLEO electron ID variable; uses calorimeter, dE/dx, tracking, time-of-flight information; about $90 - 95\%$ efficient for $e^{\pm}$ from ψ
ABSMOM	> 0.6	track momentum (GeV)
num_good_tng	5	Number of tracks in event satisfying the TRKMAN and KINCD cuts listed above
		Pion Candidates
|DBCD|	< 0.005	maximum $r - \phi$ impact parameter (m)
R2ELEC	< 3.0	CLEO electron ID variable, as above; rejects tracks with an extremely high probability of being an $e^{\pm}$
DPTHMU	< 5.0	# of nuclear interaction lengths of iron traversed by a track before being detected by the muon detectors
|SGPIDI|	< 3.0	normalized difference between measured dE/dx and that expected for a pion (in σ); applies only if the quality of dE/dx info is good

Table 3.3: Selection criteria for charged tracks.

Variable	Criteria	Description of Variable
Tier-1 Muon Candidates		
DPTHMU	> 3.0	number of nuclear interaction lengths of iron traversed by a track before being detected by the CLEO muon detectors
Tier-2 Muon Candidates		
|CZCD|	$\in [0.707, 0.85]$	cosine of the angle formed between the track and the beamline
ABSMOM	$\in [1.0, 1.8]$	track momentum (GeV)
muon_shower	$= 1$	Track is matched with a minimum-ionizing shower in the calorimeter
Tier-3 Muon Candidates		
|CZCD|	> 0.85	cosine of the angle formed between the track and the beamline
muon_shower	$= 1$	Track is matched with a minimum-ionizing shower in the calorimeter

Table 3.4: Selection criteria for muon candidates.

Variable	Criteria	Description of Variable
		Bremsstrahlung Showers
EBUMP	> 0.010	shower energy (GeV)
NBCREG	$= 1$	number of showers in a connected region
ANGBT	> 20.0	angle between shower and closest track (degrees)
		"Muon" Showers
EBUMP	$\in [0.160, 0.300]$	shower energy (GeV)
NBCREG	$= 1$	number of showers in a connected region
E925	$>$ C92501	energy deposited in a 3×3 array of crystals, divided by the energy in the 5×5 array that surrounds it; C92501 is a cut value with $\sim 99\%$ efficiency for true photons
		Isolated Showers
EBUMP	> 0.030	shower energy (GeV), for $\lvert\cos\theta\rvert > 0.707$ region only ("Good Barrel")
	> 0.050	for $\lvert\cos\theta\rvert < 0.707$ region only ("Bad Barrel","Endcap")
IBSTOP	$= 0$	reject possible shower fragments
NBCREG	$= 1$	number of showers in a connected region
E925U	$> 0.98 \cdot$ C92501	"unfolded" E925
ANGBT	> 20.0	angle between shower and closest track (degrees)
RMO	> 0.50	ratio of second moments in θ and ϕ
CRMASS	< 0.09	mass of connected region

Table 3.5: Selection criteria for calorimeter showers. Regarding E925: the cone angle defined by the IP and a 3×3 block of crystals is typically ~ 4 degrees; for a 5×5 block, ~ 7 degrees.

Search Mode	Decay Chain	Observed Yield, Data	
		On-Resonance	Continuum
"$\psi(1S)$ inclusive"	$\psi(1S) \to \ell^+\ell^-$	16478 ± 166	142 ± 31
"$\psi(1S)$ from χ_{c1}"	$\chi_{c1} \to \psi(1S)\gamma$	689 ± 46	2 ± 7
"$\psi(1S)$ from $\psi(2S)$"	$\psi(2S) \to \psi(1S)\pi^+\pi^-$	703 ± 37	8 ± 5
"$\psi(2S)$ via $\psi\pi\pi$"	$\psi(2S) \to \ell^+\ell^-\pi^+\pi^-$	same as above	same as above
"$\psi(2S)$ via dileptons"	$\psi(2S) \to \ell^+\ell^-$	639 ± 48	15 ± 18

Table 3.6: Yields for each measured decay mode, CLEO II and CLEO II.V data. The yields are not corrected for detection efficiency and include both electron and muon modes.

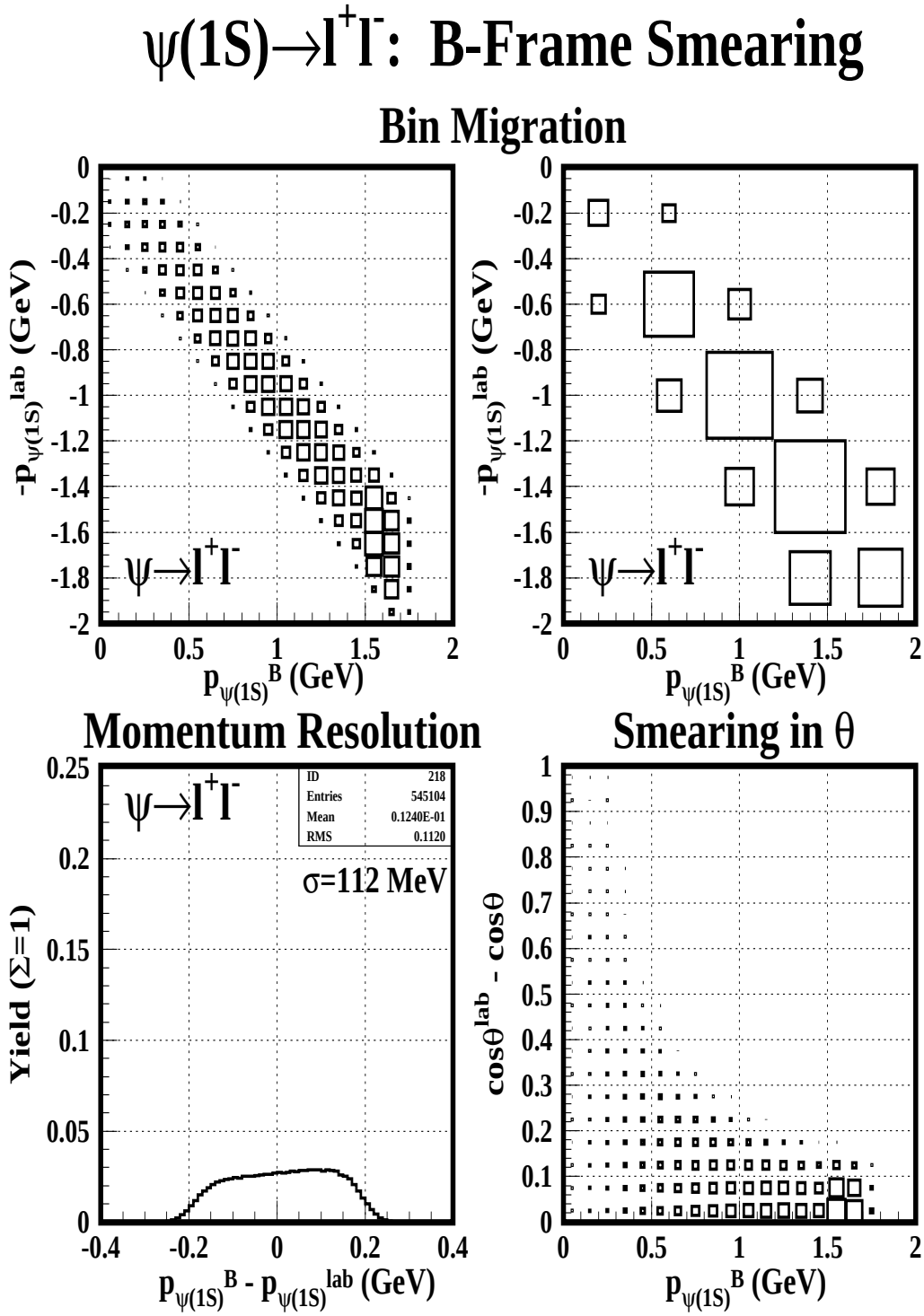

Figure 3.1: Effects of B-frame Smearing. Monte Carlo events, $\psi(1S) \to \ell\ell$, generator-level momenta. Top: Box plots of B-frame versus lab-frame momentum. The vertical axis shows $-p$ so that the similarity between the plot and the identity matrix is evident. The left plot uses the standard bin width of 100 MeV; the right plot uses a coarse bin width (400 MeV) which reduces the bin migration to the 1-bin level. Bottom Left: Momentum resolution. The normalization is such that the sum over all bins yields 1. Bottom Right: Smearing in $\cos\theta_{\psi(1S)}$ as a function of $p_\psi(1S)$. (a-11-11)

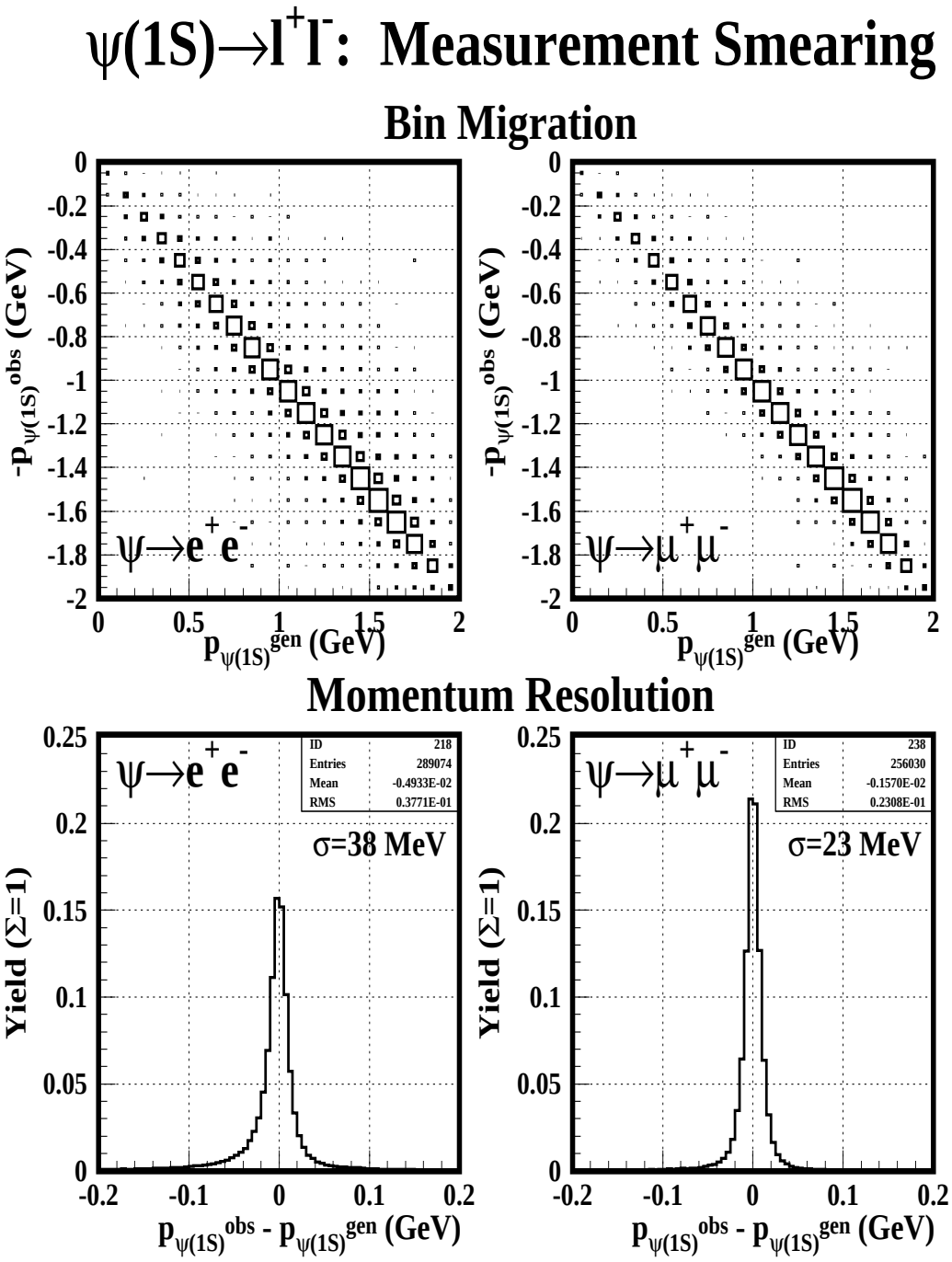

Figure 3.2: Effects of Measurement Smearing. Tagged signal Monte Carlo events, $\psi(1S) \to \ell\ell$. Top: Box plots of generated versus observed momentum. The vertical axis shows $-p$ so that the similarity between the plot and the identity matrix is evident. Bottom: Momentum resolution. The normalization is such that the sum over all bins yields 1.

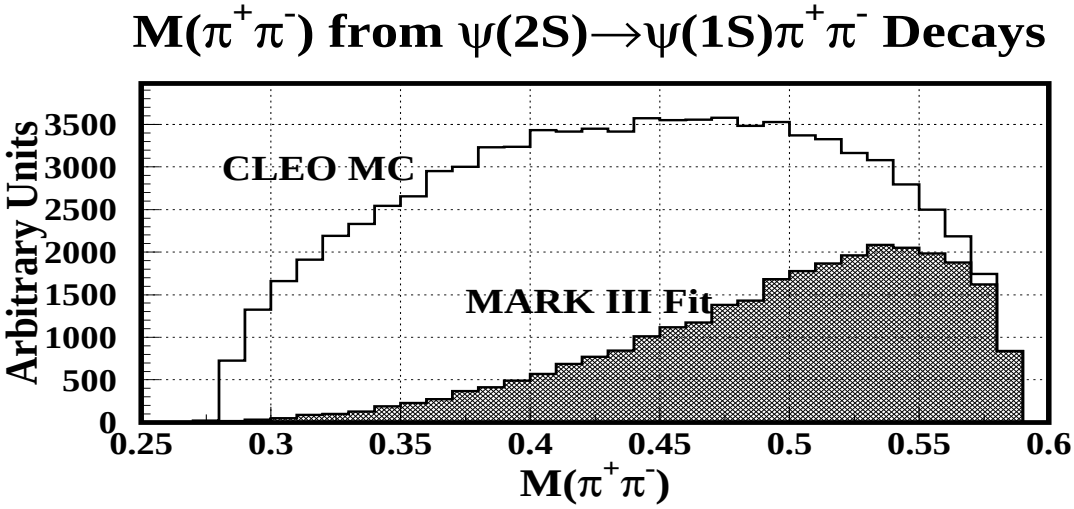

Figure 3.3: $M(\pi^+\pi^-)$ in the decay $\psi(2S) \to \psi(1S)\pi^+\pi^-$. The upper curve shows the output of CLEO Monte Carlo, which is 3-body phase space. The lower curve is given by a formula in Ref. [62], which matches the observed spectrum from MARK III [59]. Monte Carlo events are originally generated with the distribution shown by the upper curve; we use an event rejection method to obtain the distribution shown by the lower curve.

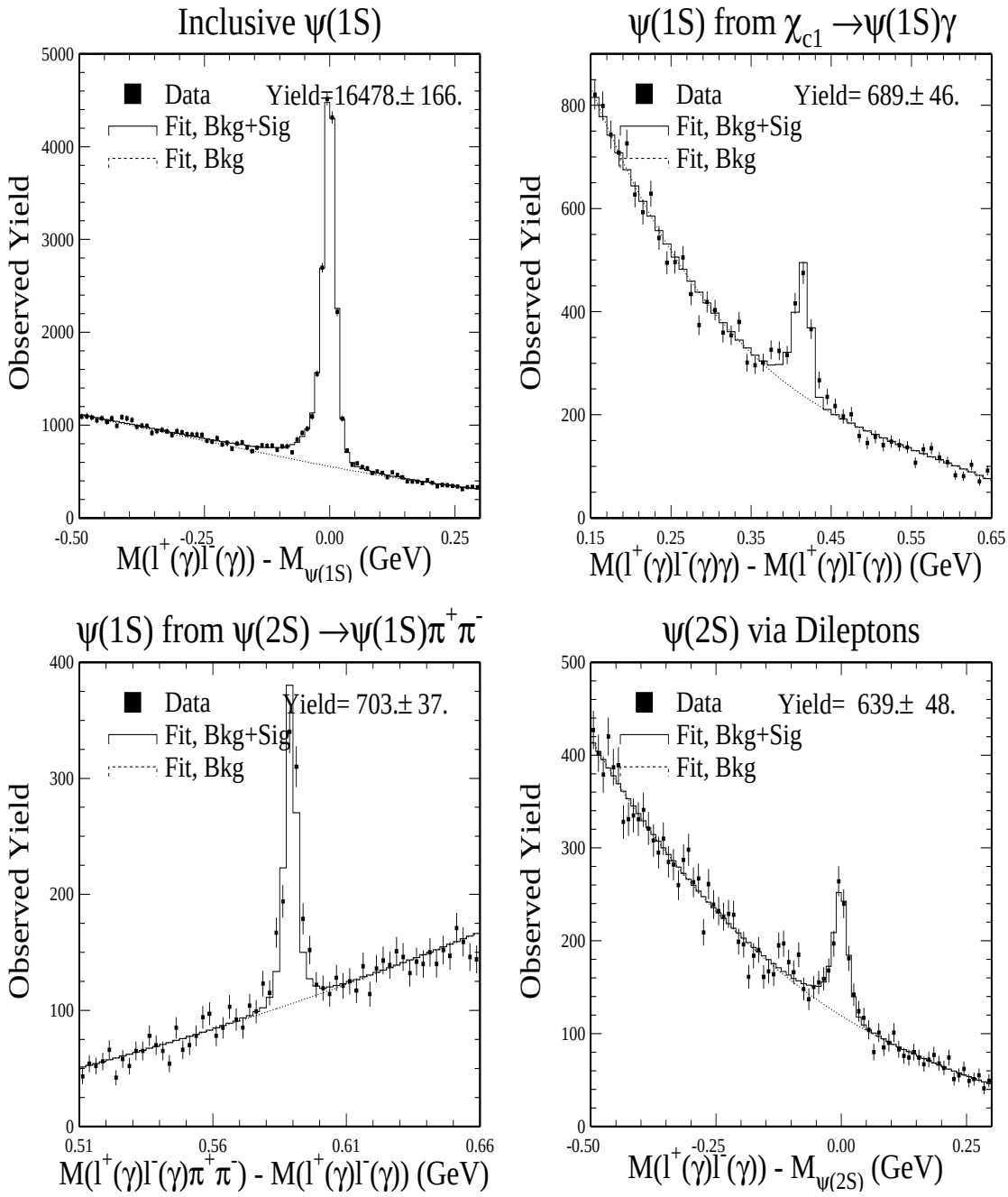

Figure 3.4: The four search modes for the sum of CLEO II and CLEO II.V on-resonance data. All ψ helicity angles; both ψ dilepton decay modes (e and μ). $p_\psi(1S) < 2.0$ GeV. $p_\psi(2S) < 1.6$ GeV.

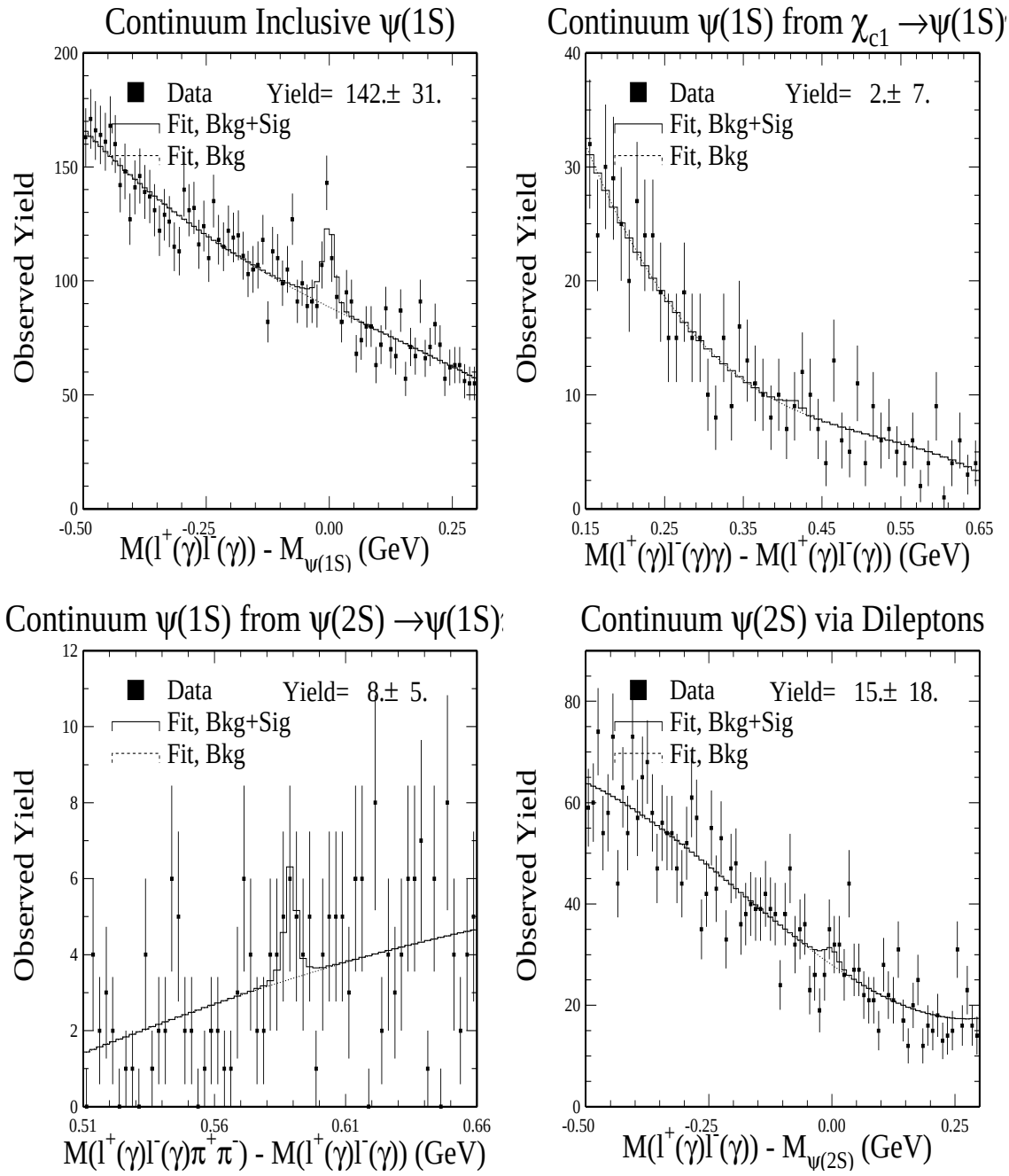

Figure 3.5: The four search modes for the sum of CLEO II and CLEO II.V continuum data. All ψ helicity angles; both ψ dilepton decay modes (e and μ). $p_\psi(1S) < 2.0$ GeV. $p_\psi(2S) < 1.6$ GeV.

Chapter 4

Measurements of ψ Momentum Distributions and Inclusive Branching Fractions

4.1 Analysis

4.1.1 Analysis Overview

The basic steps in measuring the ψ momentum spectrum are to partition the data into momentum bins, fit the mass peaks in each partition, then correct each yield for detection efficiency.

First, we obtain signal lineshapes from tagged Monte Carlo, which are used to parametrize the signal in fitting both the Monte Carlo and the data. Including the systematic error studies and the polarization measurement, these analyses required on the order of 10^4 fits of mass peaks. An automated system of plot-fitting was developed which mechanized the process while monitoring the quality of each individual fit.

In the $\psi(1S)$ mode, we must subtract the momentum distributions from the two feed-down modes as well as the distribution of continuum $\psi(1S)$. In the case of $\psi(1S)$ from $\psi(2S)$, we only measure one of the modes in which this occurs; Monte Carlo and isospin arguments give us the momentum distributions of the modes that we do

not measure.

The efficiency is measured as a function of ψ momentum; the data is corrected bin-by-bin for efficiency before being added together to give us the inclusive yield, which becomes the inclusive branching fraction.

In order to estimate systematic errors, we repeat the momentum spectrum measurement 19 times for $\psi(1S)$ and 16 times for $\psi(2S)$, each time activating a "knob" which modifies an assumption or procedure. Measuring the deviation from the standard procedure caused by the knobs gives a measure of systematic error, both separately for each bin in momentum, as well as for the inclusive yield. Since our measurements rely on the Monte Carlo for absolute efficiencies, many of the knobs test the Monte Carlo at this level.

The efficiency is defined in such a way as to enable a later disentanglement of the measurement smearing from the observed momentum distribution. We present a general result which states that any detector measurement matrix may be expressed as a product of two detectors, one which measures perfectly but inefficiently and the other which measures inaccurately but with perfect efficiency.

Our final results are inclusive branching fractions for the $B \to \psi(1S)X$, $B \to \psi(2S)X$, and $B \to \chi_{c1}X$ modes, and momentum distributions of $\psi(1S)$ and $\psi(2S)$ produced directly from B decays. The momentum distributions are corrected for feed-down, continuum $\psi(1S)$, and measurement smearing.

Finally, we find that the CLEO Monte Carlo generates momentum distributions that are differ substantially from the data in the modes involving $\psi(2S)$ or χ_{c1}. We employ a "trim" program to force the MC distributions to match those of the data, since incorrect momentum distributions in the MC could bias the polarization measurement, described in the next chapter.

4.1.2 Signal Lineshapes

In order to accurately measure the number of events in the signal region, we require a parametrization of the signal lineshapes. We define 36 ($4 \times 3 \times 3$) partitions, and obtain a signal lineshape for each from the signal Monte Carlo. The partition

divisions are defined as follows:

- Search mode (4):
  - "$\psi(1S)$ inclusive"
  - "$\psi(1S)$ from χ_{c1}"
  - "$\psi(1S)$ from $\psi(2S)$" (or "$\psi(2S)$ via $\psi\pi\pi$")
  - "$\psi(2S)$ via dileptons"
- Dilepton mode lepton species (3):
  - electrons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$
  - muons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow \mu^+\mu^-$
  - both: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$ or $\mu^+\mu^-$
- p_ψ (measured momentum) (3):
  - "low": 0.0 GeV$\leq p_{\psi(1S)} <$ 0.8 GeV, 0.0 GeV$\leq p_{\psi(2S)} <$ 0.7 GeV
  - "mid": 0.8 GeV$\leq p_{\psi(1S)} <$ 1.4 GeV, 0.7 GeV$\leq p_{\psi(2S)} <$ 1.1 GeV
  - "high": 1.4 GeV$\leq p_{\psi(1S)} <$ 2.0 GeV, 1.1 GeV$\leq p_{\psi(2S)} <$ 1.6 GeV

Even though every event in the signal MC contains a signal decay chain, tracks and showers from the decay of the other B in the event will appear as combinatoric background in the invariant mass plots. In order to remove these backgrounds from the signal lineshapes, we use CLEO tagging routines to link measured tracks and showers to particles as generated in the Monte Carlo. The following final state particles were checked by the tagger: the two charged tracks used to form $\psi(1S)$ and $\psi(2S)$ dilepton candidates, all four charged tracks in the final state of $\psi(2S) \rightarrow \psi(1S)\pi^+\pi^-$, and the photon candidate in $\chi_{c1} \rightarrow \psi(1S)\gamma$. Showers that are added to electron candidates as Bremsstrahlung photons, however, are not checked by the tagger. The performance of the tagger is illustrated by Figure 4.1, which shows the invariant-mass distributions of signal Monte Carlo events of each of the four search modes, separated into background and signal components.

Signal lineshapes were taken from an equally-weighted mix of the helicity 0 and helicity ± 1 sets, making the overall polarization of the sample $\alpha = -1/3$. Any inaccuracy in this assumption will have little effect on the lineshapes, as they have little dependence on the polarization of the ψ.

We propagate the statistical errors on the lineshapes into the statistical errors in the mass fits. This is technically incorrect because errors due to lack of Monte Carlo statistics should be counted as systematic error. However, the statistical error of the lineshapes is nearly negligible when compared to the statistical errors of the mass fits in the data; in the end it makes no difference on our results.

4.1.3 Invariant Mass Fits

We now divide the data *and* signal Monte Carlo datasets into partitions and fit each invariant mass distribution to obtain ψ yields. There are a total of 522 $[2 \times 3 \times (20 + 19 + 16 + 16 + 16))]$ mass plots to consider. The partition divisions are defined as follows:

- Data source (2):
  - Signal MC ($\alpha = -1/3$; same source as lineshapes)
  - CLEO II and II.V on-resonant data
- Dilepton mode lepton species (3):
  - electrons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$
  - muons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow \mu^+\mu^-$
  - both: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$ or $\mu^+\mu^-$
- Search mode (5):
  - "$\psi(1S)$ inclusive"
  - "$\psi(1S)$ from χ_{c1}"
  - "$\psi(1S)$ from $\psi(2S)$"

- "$\psi(2S)$ via $\psi\pi\pi$"
- "$\psi(2S)$ via dileptons"

- p_ψ (measured momentum) ($16 \sim 20$ per search mode):
  - "$\psi(1S)$ inclusive": 20 bins, 0.0 to 2.0 GeV
  - "$\psi(1S)$ from χ_{c1}": 19 bins, 0.0 to 1.9 GeV
  - "$\psi(1S)$ from $\psi(2S)$": 16 bins, 0.0 to 1.6 GeV
  - "$\psi(2S)$ via $\psi\pi\pi$": 16 bins, 0.0 to 1.6 GeV
  - "$\psi(2S)$ via dileptons": 16 bins, 0.0 to 1.6 GeV
  - 100 MeV binsize

There are 40 additional invariant mass plots to measure yields of inclusive $\psi(1S)$ in the continuum data. We partition the continuum data into CLEO II and CLEO II.V divisions; each division has 20 bins in $p_{\psi(1S)}$. We combine the electron and muon modes in these fits.

The yield of $\psi(1S)$ or $\psi(2S)$ in each partition is obtained using a binned maximum likelihood fit. Mn_fit is the fitting package used. The signal is modeled by the signal lineshapes obtained in the last section. The background is modeled with a cubic polynomial. In all except 16 of the plots, no constraints are placed on any of the fit variables, and the fits converge with a physically valid (positive) yield. The exceptions are partitions with extremely low statistics and no visually apparent signal. These fits return a negative yield. They are then re-fit with the constraint that the yield must be positive. Errors in the yield are from MIGRAD, except in the positive-constraint cases, in which case MINOS errors are used.

The fitter returns a "fit status" of 3 to signify that the fit has converged; we check this variable for all fits. Mn_fit also returns the "Likelihood," or χ^2 in the Baker-Cousins definition [10]. The χ^2 value is plotted for every fit, in the histograms shown in Figure 4.2; the mean χ^2 for each category of fits is summarized in Table 4.1. We expect that the average χ^2 should be approximately equal to the number of degrees

of freedom ($d.o.f.$) of the fit. This is indeed the case for all search modes in both data and Monte Carlo.

As a cross-check, we calculate $Y = \sum_i Y_i$, where Y_i is the yield obtained from fitting the ith momentum bin of the combined e-μ partitions. The statistical error on the sum is related to the statistical error of each bin, by $\Delta Y = \sqrt{\sum_i (\Delta Y_i)^2}$. We obtain values for Y and ΔY which are listed in Table 4.1. We may compare these values with the yields which were obtained by fitting the entire dataset (Figure 3.4). We find good agreement between the two methods.

We therefore conclude that the fits obtained are reasonable.

	Fit Quality			Yield Over All p_ψ	
search mode	$d.o.f.$	Data $\overline{\chi^2}$	MC $\overline{\chi^2}$	Summed	Figure 3.4
"$\psi(1S)$ inclusive"	75	76.1	73.3	16448 ± 165	16478 ± 166
"$\psi(1S)$ from χ_{c1}"	45	48.0	43.7	687 ± 46	689 ± 46
"$\psi(1S)$ from $\psi(2S)$"	55	58.3	57.6	670 ± 37	703 ± 37
"$\psi(2S)$ via $\psi\pi\pi$"	55	56.7	60.7	669 ± 36	703 ± 37
"$\psi(2S)$ via dileptons"	75	76.9	67.6	640 ± 48	639 ± 48
"continuum $\psi(1S)$ incl."	75	75.1	n/a	170 ± 30	142 ± 31

Table 4.1: Properties of the Invariant Mass Fits. Mean Values of χ^2 for each ψ search mode. $d.o.f.$ = degrees of freedom = number of bins − number of fit parameters.

4.1.4 Yields, MC (Efficiency)

To calculate the detection efficiency of ψ in a given momentum range, we first isolate all the signal Monte Carlo events which contain only 1 ψ in the event (to avoid complexities with the tagger); the generated value of p_ψ must be in the momentum range under consideration. The denominator of the efficiency ratio is the number of events in this group. This group of events is then run through the detector simulation. The yield of the invariant mass peak formed by the reconstructed events is the numerator of the efficiency ratio. The numerator does not include any cuts on the reconstructed ψ momentum. In both the numerator and denominator of the efficiency calculation, we cut on the generator-level momentum in the CLEO lab frame. This

differs from the more typical definition of efficiency, in which the numerator cuts on the *observed* momentum. The reason for our adoption of this particular definition of momentum will be discussed in Section 4.4. The efficiency results are shown in the upper right and center right plots of Figures 4.3, 4.4, 4.5, 4.6, and 4.7 for each of the search modes.

Figure 4.3 shows data for the inclusive $\psi(1S)$ search mode, but uses the signal Monte Carlo events of directly produced $\psi(1S)$ to estimate the detection efficiency. This is valid because for a given bin of $\psi(1S)$ momentum, whether or not a $\psi(1S)$ has an intermediate parent has little bearing on the efficiency of finding the $\psi(1S)$. This is justified by Figure 4.8, which shows that the ratio of $\epsilon_{\psi(1S)(direct)}$ to $\epsilon_{\psi(1S)(feed-down)}$ is consistent with one for all $p_{\psi(1S)}$ and both feed-down modes.

The ψ detection efficiency depends somewhat on polarization. A Monte Carlo study of this effect can be seen in Figure 4.9, which shows the ratio of purely transverse $\psi(1S)$ efficiency to purely longitudinal $\psi(1S)$ efficiency, as a function of $p_{\psi(1S)}$. The ratios are consistent with 1, with the exception of high momentum, in which transverse ψ, with a greater occurrence of forward-backward decays, produce more soft leptons than do longitudinal ψ. By assuming that $\alpha = -1/3$ as we have for this measurement, we are vulnerable to bias to the extent that the actual polarization differs from $-1/3$. We consider the magnitude of this effect later in the systematic error study.

4.1.5 Yields, Data

The left top and left middle plots of Figures 4.3, 4.4, 4.5, 4.6, and 4.7 show the efficiency-corrected yields from data for each of the search modes.

In order to obtain the momentum distribution of $\psi(1S)$ created directly from B decays, we must subtract the "feed-down" distributions (Figures 4.4 and 4.5) from the inclusive $\psi(1S)$ distribution (Figure 4.3). However, as discussed in Section 3.1.3, there are also several feed-down modes which we do not attempt to measure. We now construct the momentum distributions from these unmeasured decays.

Assuming that the $\psi(1S)$ momentum distributions for $\psi(2S) \to \psi(1S)\pi^+\pi^-$ and $\psi(2S) \to \psi(1S)\pi^0\pi^0$ are identical, we can multiply the efficiency-corrected yield of

$\psi(2S) \to \psi(1S)\pi^+\pi^-$ by $(1 + \frac{20.3}{34.8}) = 1.583$ to obtain the efficiency-corrected yield of $\psi(2S) \to \psi(1S)\pi\pi$. The ratio is taken from the 2001 PDG, and the uncertainty on the branching fractions is handled later in the systematic error study.

Likewise, using the Monte Carlo simulation, we generate (pre-detector simulation) momentum distributions of the "miscellaneous" unmeasured modes $\psi(2S) \to \psi(1S)\eta$, $\psi(2S) \to \psi(1S)\pi^0$, $\psi(2S) \to \chi_{c2}\gamma \to \psi(1S)\gamma\gamma$, and $\psi(2S) \to \chi_{c0}\gamma \to \psi(1S)\gamma\gamma$. The relative normalization of the modes is given by the 2001 PDG. The Monte Carlo distributions for the miscellaneous modes are then scaled so that the momentum-integrated, pre-detector yield, over the momentum-integrated, efficiency-corrected yield of $\psi(2S) \to \psi(1S)\pi^+\pi^-$, is equal to the PDG 2001 ratio of branching fractions.

Having measured (or obtained from Monte Carlo) the momentum distributions of all $B \to \psi(2S) \to \psi(1S)X$ decay chains listed in the 2001 PDG, we obtain the momentum distribution of $\psi(1S)$ from these decays, Figure 4.10.

Since the PDG lists only one mode in which χ_{c1} decays into a $\psi(1S)$, we assume that this mode, $\chi_{c1} \to \psi(1S)\gamma$, accounts for 100% of $\psi(1S)$ produced from χ_{c1}.

By subtracting the efficiency-corrected feed-down momentum distributions from the inclusive $\psi(1S)$ distribution, we obtain the distribution of directly produced $\psi(1S)$. Figure 4.11 summarizes the measurement for $\psi(1S)$, showing the momentum distribution of inclusively produced $\psi(1S)$ and all sub-components.

There is no feed-down to worry about in the $\psi(2S)$ mode. The momentum distribution of directly produced $\psi(2S)$ is obtained by adding the statistically independent samples of $\psi(2S) \to \psi(1S)\pi^+\pi^-$ (Figure 4.6) and $\pi \to \ell^+\ell^-$ (Figure 4.7).

4.1.6 Normalization; Inclusive Branching Fraction

The inclusive branching fractions $\mathcal{B}(B \to \psi X)$ are related to the efficiency-corrected yields we have obtained, by the following:

$$N_{data} = 2N_{B\overline{B}}\ \mathcal{B}(B \to \psi X)\ \sum_{i=1}^{m} \mathcal{B}(\psi \to f_i) \tag{4.1}$$

where N_{data} is the efficiency-corrected yield of ψ observed in the data, $2N_{B\overline{B}}$ is the total number of B mesons produced (2×9699702), $\mathcal{B}(B \to \psi X)$ is the inclusive

branching fraction, and $\sum_{i=1}^{m} \mathcal{B}(\psi \rightarrow f_i)$ is the fraction of ψ which we attempt to reconstruct; f_i is the final state of the i-th reconstruction mode, and m is the number of decay modes included in the measurement of N.

When we wish to combine results from all reconstructed decay modes to provide the best statistical result for $\mathcal{B}(B \rightarrow \psi X)$, we set m equal to the total number of modes: 2 for $\psi(1S)$ and 4 for $\psi(2S)$. However, by setting $m = 1$, we may also determine results for $\mathcal{B}(B \rightarrow \psi X)$ implied by a single decay mode. This provides a useful cross-check as we can compare measurements of different ψ modes which have different branching fractions.

Equation (4.1) enables us to convert the normalization of the vertical axis of our momentum distributions from efficiency-corrected yield per bin, N, to inclusive branching fraction per bin in momentum, $d\mathcal{B}/dp$. Doing so gives us the plots on the lower right corners of Figures 4.3, 4.4, 4.5, 4.6, and 4.7.

Figure 4.12 shows summary plots of the momentum distribution of directly produced $\psi(1S)$ and $\psi(2S)$. The upper plots are our results; the lower plots show the results of the previous CLEO measurement [35]. In all cases, only statistical errors are shown.

A summary of the inclusive branching fraction measurements is given in Table 4.2 (again, with no systematic errors at this point.) We see excellent agreement between the e and μ modes of ψ reconstruction for $\mathcal{B}(B \rightarrow \psi(1S)X)$. When using PDG 2001 values for the reconstruction mode branching fractions, the four measured decay modes of $\psi(2S)$ also agree reasonably well, with the exception of the $\psi(2S) \rightarrow \mu^+\mu^-$ mode. The significantly lower value of $\mathcal{B}(B \rightarrow \psi(2S)X)$ for this mode occurs primarily because of the very high PDG 2001 value (but with a very large uncertainty) for $\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)$.

The PDG has only one experimental measurement on which to base its evaluation of $\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)$, a 1975 SPEAR measurement [63]. Our efficiency-corrected yields, as well as recent results from BaBar [41] [64] [65], are consistent with $\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-) = \mathcal{B}(\psi(2S) \rightarrow e^+e^-)$ at the 20% level, as opposed to at the 50% level as the central values of the PDG 2001 values suggest. For our final result, we

will therefore assume lepton universality, i.e. $\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = \mathcal{B}(\psi(2S) \to e^+e^-)$ (neglecting phase space differences), with a 20% uncertainty. As a result of making this assumption, the four independent measurements of $\mathcal{B}(B \to \psi(2S)X)$ are significantly more consistent with each other than if we had used the PDG 2001 value for $\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$. The results are listed in Table 4.2.

Reconstructed Modes	$10^3\times$	Data	
Search Mode ($\psi \to f_i$)	$\sum \mathcal{B}(\psi \to f_i)$	N_{data}	$\mathcal{B}(B \to \psi X) \times 10^3$
$\psi(1S) \to e^+e^-$	59.3 ± 1.0	13214 ± 196	$11.49 \pm 0.17 \pm 0.19$
$\psi(1S) \to \mu^+\mu^-$	58.8 ± 1.0	13054 ± 190	$11.44 \pm 0.17 \pm 0.19$
$\psi(1S) \to \ell^+\ell^-$	118.1 ± 1.4	26231 ± 272	$11.45 \pm 0.12 \pm 0.14$
$\psi(1S) \to \ell^+\ell^-$ (continuum-sub.)	118.1 ± 1.4	25682 ± 289	$11.21 \pm 0.13 \pm 0.13$
$\psi(2S) \to e^+e^-$	7.9 ± 0.5	433 ± 48	$2.82 \pm 0.32 \pm 0.18$
$\psi(2S) \to \mu^+\mu^-$ (lepton univ.)	7.9 ± 1.6	510 ± 49	$3.33 \pm 0.32 \pm 0.67$
$\psi(2S) \to \ell^+\ell^-$ (lepton univ.)	15.8 ± 1.7	946 ± 71	$3.09 \pm 0.23 \pm 0.32$
$\psi(2S) \to \psi(1S)[\to e^+e^-]\pi^+\pi^-$	20.6 ± 1.7	1291 ± 98	$3.22 \pm 0.24 \pm 0.27$
$\psi(2S) \to \psi(1S)[\to \mu^+\mu^-]\pi^+\pi^-$	20.5 ± 1.7	1240 ± 96	$3.12 \pm 0.24 \pm 0.26$
$\psi(2S) \to \psi(1S)[\to \ell^+\ell^-]\pi^+\pi^-$	41.1 ± 3.3	2538 ± 137	$3.18 \pm 0.17 \pm 0.26$
All $\psi(2S)$ modes (lepton univ.)	61 ± 5.2	3484 ± 154	$3.16 \pm 0.14 \pm 0.16$
$\psi(2S) \to \mu^+\mu^-$ (PDG 2001)	12 ± 4	510 ± 49	$2.19 \pm 0.21 \pm 0.73$
$\psi(2S) \to \ell^+\ell^-$ (PDG 2001)	19.9 ± 4.0	946 ± 71	$2.45 \pm 0.18 \pm 0.50$
All $\psi(2S)$ modes (PDG 2001)	61 ± 5.2	3484 ± 154	$2.94 \pm 0.13 \pm 0.25$

Table 4.2: $B \to \psi X$ results. The errors on $\mathcal{B}$ are statistical and scale error. "Scale error" refers to the uncertainty on the branching fractions of the modes in which we reconstruct the ψ, i.e., the error in the second column of the table. The bottom third of the table shows the results of using the PDG 2001 value for $\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$ instead of assuming lepton universality in $\psi(2S)$ dilepton decay to 20% as is done in the standard procedure.

knob	knob description / Measurement:	Momentum $\psi(1S)$	Momentum $\psi(2S)$	Polarization $\psi(1S)$	Polarization $\psi(2S)$
	a: unmeasured $\psi(2S) \to \psi(1S)X$ decays				
a1	$\mathcal{B}(\psi(2S) \to \psi(1S)\pi^0\pi^0)$	X	-	X	-
a2	$\mathcal{B}(\psi(2S) \to \psi(1S)X_{misc})$	X	-	-	-
a3	$\alpha(X_{misc}) = +1$	-	-	X	-
a4	$\alpha(X_{misc}) = -1$	-	-	X	-
	b: invariant mass fit procedure				
b1	background: quadratic polynomial	X	X	X	X
b2	background: fix bkgd., get shape from data	X	X	X	X
b3	signal: allow horizontal offset	X	X	X	X
b4	signal: sum of Gaussians, get shape from data	X	X	X	X
	c: low-level MC efficiency accuracy				
c1	ϵ(track): cut 2.6% ($<$ 250 MeV), 0.7% ($>$ 250 MeV)	X	X	X	X
c2	absolute $\epsilon(\gamma)$: cut 2%	X	X	X	X
c3	non-GB $\epsilon(\gamma)$: cut 2% in non-GB	X	X	X	X
c4	absolute $\epsilon(\mu)$: cut 1% ($<$ 2 GeV), 0.5% ($>$ 2 GeV)	X	X	X	X
c5	non-GB $\epsilon(\mu)$: cut 2.5% in non-GB	X	X	X	X
c6	absolute $\epsilon(e)$: cut 2%	X	X	X	X
c7	non-GB $\epsilon(e)$: cut 5% in non-GB	X	X	X	X
	d: high-level MC efficiency accuracy				
d1	$\alpha_{MC} = +1$	X	X	-	-
d2	$\alpha_{MC} = -1$	X	X	-	-
d3	$M(\pi^+\pi^-)$ cut in $\psi(2S) \to \psi(1S)\pi^+\pi^-$	X	X	X	-
d4	`R2GL` cut	X	X	-	-
	e: luminosity-related				
e1	continuum scale factor	X	-	X	-
e2	$n_{B\overline{B}}$	X	X	-	-
	f: feed-down correction, α extraction procedure				
	first, obtain angular distribution of direct ψ				
STD	then, fit directly for α and N	-	-	-	-
f1	then, fit for N_0 and N_1; compute α	-	-	X	X
f2	then, fit for N_0 and N_1; get α from toy MC	-	-	X	X
	first, fit for N_0 and N_1, for inclusive and feed-down angular distributions				
f3	then, compute α	-	-	X	-
f4	then, get α from toy MC	-	-	X	-

Table 4.3: Knobs for the Systematic Error Study, as discussed in Section 4.2 (for the momentum distribution and inclusive branching fraction analyses) and Section 5.2 (for the polarization analysis). "X" means that the measurement [momentum or polarization for $\psi(1S)$ or $\psi(2S)$] was repeated with the knob "activated;" "-" denotes that the knob was not applicable to the measurement. "STD" denotes "standard procedure," which is listed for reference.

4.2 Systematic Error Study

4.2.1 Overview

Table 4.3 shows the list of of "knobs" used to estimate the systematic error. We re-measure the momentum distributions and inclusive branching fractions, each time activating exactly one of the knobs. Many of the knobs (a1-a2, c1-c7, d3, e1-e2) change some parameter used in the analysis. We assume that the small changes produced by wiggling these parameters are symmetric; therefore we need only change each parameter in one direction to determine the deviation produced by changing the parameter.

We now provide detailed descriptions of each of the knobs used for the momentum distribution measurement. The reader may wish to refer to Tables 4.6, 4.7, and 4.8 to see the effect of each knob as they are discussed. Table 4.6 lists the deviation caused by activating each knob on the inclusive branching fractions of each search mode. Tables 4.7 and 4.8 show the deviation caused by the knobs on the yield of directly produced $\psi(1S)$ and $\psi(2S)$, as a function of p_ψ.

4.2.2 Uncertainty in Unmeasured Modes

Knobs a1 and a2: Branching Fractions of Unmeasured Modes

In the $\psi(1S)$ momentum spectrum measurement, we made use of PDG 2001 branching fractions in order to estimate the momentum distributions of several $\psi(2S)$ feed-down modes which we do not directly measure. In this section we consider the effect that the errors on these PDG results have on the momentum distribution measurement.

The first PDG figure to consider is $\dfrac{\mathcal{B}(\psi(2S) \to \psi(1S)\pi^0\pi^0)}{\mathcal{B}(\psi(2S) \to \psi(1S)\pi^+\pi^-)} = 1.6532 \pm 0.0698$. We used this central value to adjust the yield of $\psi(2S) \to \psi(1S)\pi^+\pi^-$ upwards, assuming that the $\psi(1S)$ distribution is identical in both modes. Knob a1 adjusts this branching fraction by $+1\sigma$, using 1.7230 instead.

Second, there were several modes with small branching fractions (the "miscel-

laneous feed-down modes" of Section 3.1.3) which we did not measure; the Monte Carlo simulated provided the ψ momentum distributions for these decay modes. Using PDG 2001 results, we obtain

$$\frac{\mathcal{B}(\psi(2S)\to\psi(1S)\eta)+\mathcal{B}(\psi(2S)\to\chi_{c2}\gamma\to\psi(1S)\gamma\gamma)+\mathcal{B}(\psi(2S)\to\psi(1S)\pi^0)+\mathcal{B}(\psi(2S)\to\chi_{c0}\gamma\to\psi(1S)\gamma\gamma)}{\mathcal{B}(\psi(2S)\to\psi(1S)\pi^+\pi^-)}$$

$= 0.0984 \pm 0.0237$. Knob a2 uses the value 0.1221 instead, which is $+1\sigma$ from the PDG central value.

– **Discussion, a Knobs**

For the inclusive branching fraction measurements, the a knobs will affect only the "$\psi(1S)$ from $\psi(2S)$" feed-down mode and, by subtraction, the direct $\psi(1S)$ mode. The momentum spectrum of direct $\psi(1S)$ is also affected; the bins that are the most susceptible to systematic error are the bins which have the highest yields in the "$\psi(1S)$ from $\psi(2S)$" feed-down mode. These knobs are the dominant systematic error in determining $\mathcal{B}[B \to \psi(2S)X \to \psi(1S)X]$, but are minor in the direct $\psi(1S)$ measurements.

4.2.3 Invariant Mass Fit Procedure

Knobs b1 and b2: Invariant Mass Background

When fitting the invariant mass plots, we parametrized the background as a cubic polynomial ($dx^3 + cx^2 + bx + a$, where x is the invariant mass variable in GeV), of which all 4 terms (a, b, c, and d) were allowed to float in each fit.

Knob b1 replaces the floating cubic polynomial with a floating quadratic polynomial. The lack of a cubic term caused the fitter to never model the background shape as a "wave" or sideways "S", as was occasionally observed in cubic background fits of very low statistics. Distributions of the fit likelihoods were made and the results are shown in Table 4.4. Relative to the number of degrees of freedom, the mean χ^2 of the fits stay nearly constant.

Knob b2 retains the cubic polynomial parametrization, but the shape is fixed, so that the overall normalization is the only degree of freedom. This is accomplished by

floating the cubic term d but fixing the ratios $\frac{c}{d}$, $\frac{b}{d}$, and $\frac{a}{d}$. The values of these ratios are taken from fits of the data, integrated over momentum. We obtain 15 different background shapes: 5 for each of the search modes, and 3 for each lepton possibility (e, μ, or both). With only one degree of freedom, the susceptibility of the fitter to wrongly accommodate statistical fluctuations in some background bins is eliminated. The values of $\chi^2 - d.o.f.$ (Table 4.4) show that in both dilepton modes, fits of the data are worse than in our standard procedure, suggesting a significant variation of background shape as a function of p_ψ.

Knob:	Std. Procedure			b1			b2		
Fit Background:	Float Cubic			Float Quadratic			Fixed Cubic		
	d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$	
search mode		Data	MC		Data	MC		Data	MC
"$\psi(1S)$ inclusive"	75	76.1	73.3	76	76.7	76.4	78	82.5	81.5
"$\psi(1S)$ from χ_{c1}"	45	48.0	43.7	46	50.2	53.9	48	51.9	49.4
"$\psi(1S)$ from $\psi(2S)$"	55	58.3	57.6	56	59.3	57.8	58	60.6	67.5
"$\psi(2S)$ via $\psi\pi\pi$"	55	56.7	60.7	56	57.3	61.3	58	59.3	71.6
"$\psi(2S)$ via dileptons"	75	76.9	67.6	76	77.5	69.6	78	92.9	71.0

Table 4.4: Mean χ^2 for invariant mass fits, using alternate parametrizations of the background (knobs b1 and b2).

Knobs b3 and b4: Invariant Mass Signal Lineshape

Now we vary the other parameterization used in the invariant mass fits, namely, that of the signal. Our "standard" procedure is to use signal lineshapes obtained from tagged signal Monte Carlo.

Knob b3 parametrizes the signal as a "smoothed" histogram taken from the tagged signal Monte Carlo. Using a smoothed histogram allows us to add an additional degree of freedom, a horizontal offset. This parametrization therefore has the advantage of correctly accommodating any shifts in invariant mass between the data and Monte Carlo. In order to see how much this potential shift might be, we first did 4 preliminary fits, one for each of the search modes. In each of these fits we fit the entire

unpartitioned data with the Monte Carlo tagged signal lineshape. The amounts by which the Monte Carlo lineshape had to be shifted to optimally fit the data were on the sub-MeV level, and are listed in Table 4.5.

Using smoothed histograms in repeating the analysis, we found that obviously unreasonable values of the horizontal offset parameter were obtained for some fits of low-statistics partitions. Therefore the value of the horizontal offset was constrained to be within $\pm 3\sigma$ of the central value of the offset obtained from fitting the unpartitioned data as above. As usual, distributions of the mean χ^2 were obtained from the fits and the means are listed in Table 4.5. The mean χ^2 was lower than our standard procedure in many cases, as can be seen by comparing to the standard procedure results in Table 4.4.

Knob b4 models the signal as a sum of Gaussians, the parameters of which are obtained as much as possible from fits of the data, integrated over ψ momentum. In all but one search mode, the signal/background ratio of the fits is insufficient to obtain sensible results for the parameters by fitting the data alone. Each mode requires a unique procedure to extract the signal lineshape.

For $\psi(1S) \to \ell\ell$, we parametrize the signal as a sum of two asymmetric ("bifurcated") Gaussians. The means of the two Gaussians are set to be equal, but the four widths and relative normalization are allowed to float freely. Fitting the unpartitioned data with this signal shape and a cubic polynomial background, we obtain 3 signal lineshapes for $\ell = e, \mu$, or both. To check the lineshapes, we attempt to fit the tagged signal Monte Carlo with them, allowing the overall width to float. The fits do not converge, because the lineshapes obtained from the data do not include the long radiative tail in invariant mass. The fits of the momentum-integrated data incorporate the radiative mass tails into the background[1]. We make a second attempt to fit the Monte Carlo with the data lineshape by adding a floating cubic polynomial.

[1]The procedure for Knob b4 redefines the signal as the peak in invariant mass, ignoring the radiative mass tail. Since our final results are the ratio between data and Monte Carlo (efficiency) distributions, each of which uses this new definition of signal lineshape, the final results should not be affected by this redefinition. In that sense, Knob b4 tests the accuracy of the $\psi \to \ell\ell$ Monte Carlo in two important ways – first, in the shape of the mass peak (without mass tail), and second, in the size of the radiative tail relative to the mass peak.

These fits converge, and the ratio of Monte Carlo to data signal widths is found to be consistent with one.

Now we proceed to obtain signal lineshapes for the related modes $\psi(2S) \to \ell\ell$. We find it impossible to obtain a sensible fit for this lineshape directly from the momentum-integrated data, given the poor signal-to-background ratio. However, we know that $\psi(2S) \to \ell\ell$ decays should result in an invariant mass distribution that is very similar to that of $\psi(1S) \to \ell\ell$. Indeed, when we reproduce the above signal Monte Carlo test, namely, fit the $\psi(2S) \to \ell\ell$ signal Monte Carlo using the lineshape obtained from $\psi(1S) \to \ell\ell$ data plus a cubic polynomial background, the χ^2 is consistent with a good fit. However, we observe that the overall width of $\psi(2S) \to \ell\ell$ is about 20% larger than that of $\psi(1S) \to \ell\ell$. Based on these observations, we set the $\psi(2S) \to \ell\ell$ lineshape to be the same as $\psi(1S) \to \ell\ell$, but with an overall width, w_{pt}^{data}, defined as follows: $w_{\psi(2S)}^{data} = w_{\psi(1S)}^{data} \frac{w_{\psi(2S)}^{MC}}{w_{\psi(1S)}^{MC}}$. In other words, the overall width is that of the $\psi(1S)$ data, multiplied by the ratio of $\psi(2S)$ to $\psi(1S)$ overall widths from the Monte Carlo.

We experienced difficulties in attempting to fit the the "$\psi(1S)$ from χ_{c1}" and "$\psi(1S)$ from $\psi(2S)$" with freely floating signal lineshape parameters, due to the low signal-to-background ratio in these modes. We therefore obtained some of the parameters from the Monte Carlo. For these modes we parametrize the signal with the `mn_fit` "Sum of Two Bifurcated Gaussians", which differs from the parametrization described above (and is somewhat misleading) in that this parametrization has only 3 degrees of freedom for the width parameters (both Gaussians must have left-right widths in the same ratio). We fit the tagged signal Monte Carlo to this signal parametrization plus a floating cubic polynomial, allowing all variables to float. Then, we fit to the data, fixing all of the width parameters to those of the MC (except for overall width). The combination of MC-fitted shape and data-fitted width gives us the lineshapes for these modes in each of the leptonic ψ decay channels. We find that the Monte Carlo tends to significantly underestimate the width of the signal lineshapes in the feed-down modes, which can be seen in Figure 3.4.

Having obtained the lineshapes for each of 12 partitions (4 search modes $\times$ 3 bins in ℓ), we repeat the momentum distribution measurement. The χ^2 results are shown

in Table 4.5. We observe that the χ^2 improves relative to our standard procedure for most data modes; however, the fits to the Monte Carlo simulation are not nearly as good. Many of these fits showed that the signal lineshapes from data had subtly different shapes and/or different widths than the Monte Carlo signals, leading to degraded fit quality. However, the fits are not bad in the sense that they return obviously incorrect results for the signal yield.

Knob:	b3 (Smooth Hist.)				b4 (Gaussian)		
	Offset (MeV)	d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$	
search mode			Data	MC		Data	MC
"$\psi(1S)$ inclusive"	-0.05 ± 0.15	76	74.8	73.4	75	75.8	187.7
"$\psi(1S)$ from χ_{c1}"	$+0.09 \pm 0.96$	46	47.6	42.7	45	48.1	172.7
"$\psi(1S)$ from $\psi(2S)$"	$+0.51 \pm 0.16$	56	57.4	58.5	55	58.1	188.5
"$\psi(2S)$ via $\psi\pi\pi$"	$+0.51 \pm 0.16$	56	55.6	60.6	55	56.1	193.6
"$\psi(2S)$ via dileptons"	-0.37 ± 0.16	76	76.1	76.2	75	78.2	198.2

Table 4.5: Mean χ^2 for invariant mass fits of each search mode, using alternative parametrizations of the signal lineshape. The column labeled "Offset" is from a fit of the entire unpartitioned dataset, both e and μ channels combined, and is used to constrain the maximum offset of the momentum-partitioned fits. These can be compared to the corresponding results for the standard procedure, shown in Table 4.4.

– **Discussion, b Knobs**

To both the signal and background parametrization we have assigned two knobs which change the fitting procedure. In both cases, we have a minor change (knobs b1, b3) and a much more drastic change (knobs b2, b4). The more ambitious changes to the procedure make use the data to fix the signal (b4) or background (b2) shapes; however, in order to get sufficient statistics in the data, we are forced to use the momentum-integrated data. Therefore these methods have the disadvantage that they do not allow for the signal or background shape to change as a function of p_ψ. Since for each of the signal and background parametrizations we have one conservative and one aggressive change to the procedure, we define the systematic error due to

signal (background) definition to be equal to the average deviation from the standard procedure caused by the two knobs.

In the feed-down modes, knob b4 gives the largest deviation. This is expected given our observation that the Monte Carlo and data differ significantly in the width of feed-down signal lineshapes. For the $\psi \to \ell\ell$ modes, knob b2 results in the largest deviation from the standard procedure.

4.2.4 Monte Carlo Low-Level Accuracy

Knobs c1-c7: Low-Level MC Tests

In our measurements of inclusive branching fraction and momentum distribution, we obtain the absolute detection efficiency from the Monte Carlo simulation. Since we normalize to the number of produced B mesons any difference in detection efficiency between the data and Monte Carlo simulation will bias our results. For each of the 7 "low-level" knobs, our strategy is to artificially decrease the detection efficiency by using a random number generator to randomly rejecting some percentage of the tracks, showers, muons, or electrons. The specific percentage cut out is taken from CLEO "conventional wisdom" in the form of benchmark systematic studies, cited below. Since our final result are ratios of data yields to Monte Carlo yields, we must carry out this artificial rejection on either the data or the Monte Carlo but not both. To the extent that the data has sufficient statistics, the choice of whether to cut events from the data or the Monte Carlo is arbitrary. We apply the knobs on the data, as described below:

Tracks (c1): Based on an embedding study on recompress/CLEO II.V data [66], we randomly reject 0.7% of the charged tracks; in the case of very soft tracks (< 250 MeV), we reject 2.6%. We expect that this knob will result in a 1.4 % effect for "inclusive $\psi(1S)$", "$\psi(1S)$ from χ_{c1}", and "inclusive $\psi(2S)$" modes, all of which have 2-track final states; and a 2.8 % effect for the "$\psi(1S)$ from $\psi(2S)$"/"$\psi(2S)$ via $\psi(1S)$" mode, with 4 tracks in the final state.

Showers (c2, c3): A CLEO analysis of $\frac{\Gamma(\eta \to 3\pi^0)}{\Gamma(\eta \to \gamma\gamma)}$ [67] finds that data and MC agree in the efficiency of finding single showers to 1.6%. This finding applies to

the entire barrel region for CLEO II and CLEO II.V. Knob (c2) rejects 2% of all observed showers. We conservatively assume that showers in the non-Good Barrel region are less well simulated, due to extra detector material, by defining knob (c3) to reject 2% of showers, but only in the non-Good Barrel region. Crystal calorimeter measurements will primarily affect the "$\psi(1S)$ from χ_{c1}" mode; however, all modes will be affected subtly because showers are also added as Bremsstrahlung in $\psi \to ee$ decays.

Lepton ID, Fake Rate: Fake rate and efficiency are the two quantities that should match between MC and data. In our analysis, the correct simulation of the true fake rate is not important, as we combine pairs of leptons to form ψ; ψ which are reconstructed with particles other than true leptons will not peak in $M(\psi)$. We assign no systematic error due to incorrectly modeled fake rates in the MC.

Muon ID, Efficiency (c4, c5): Comparative studies of MC and the first half of CLEO II data [68] [69] indicate that CLEOG simulates muon efficiency correctly to 0.5%. Knob (c4) rejects 0.5% of muons with momenta greater than 2.0 GeV; below 2.0 GeV, the efficiency function drops steeply and so we reject 1.0% of muons. Knob (c5) accounts for the possibility that endcap muons (such as the "Tier 2" and "Tier 3" muons used in our analysis) are modeled less accurately; this knob rejects 2.5% of muons not in the Good Barrel region. As most of the muons from ψ are softer than 2 GeV, we expect that knob (c4) will lead to a 1% effect in all modes (1% for each muon in $\psi \to \mu\mu$; but the electron mode is unaffected, leading to a 1% effect in $\psi \to \ell\ell$).

Electron ID, Efficiency (c6, c7): The software package that calculates the electron-ID variable `R2ELEC` [70] [71] fundamentally treats data differently than MC. For the data, `R2ELEC` is calculated based on physics quantities such as dE/dx and calorimeter information. For the MC, `R2ELEC` ignores the simulated physics information, as it is deemed untrustworthy for the purposes of electron ID. Instead, it considers the momentum, angle formed with the beamline, charge, and true identity[2] of the track, and returns an "expected" value of `R2ELEC` from a lookup table based

[2] i.e. whether the track is truly an electron or not, which is known for Monte Carlo

on these quantities. The lookup table is based on a study of radiative Bhabha events $e^+e^- \to e^+e^-\gamma$, a physics process which provides a clean sample of true electrons. Therefore, in principle, the MC should match the data exactly.

Our electron ID cut, `R2ELEC > 0`, is loose compared to most analyses; this results in high lepton efficiencies, which gives the Monte Carlo less room in which to be incorrect. On the other hand, unlike most analyses, we make no cuts on the detector region that the electron is detected in. We accept Good Barrel, Bad Barrel, Good Endcap, and Bad Endcap electrons in order to maximize yield. Nevertheless, most $\psi(1S) \to e^+e^-$ events (77%) are made of two Good Barrel electrons, insulating us somewhat from poor MC modelling of non-GB electrons.

To determine appropriately scaled knobs to test the effect of incorrectly modeled electron ID on our results, we need to know by what extent the MC is known to deviate from data for our particular cuts. We carried out a study [72] of hadronic events with embedded electrons from data radiative Bhabha events, using techniques from previous similar studies [73] [74]. We found that the degree to which the MC and data agree depends a great deal upon the cuts made to define an acceptable radiative Bhabha event. Variations on these cuts led to uncertainties of 2% of Good Barrel electrons and about 5% for non-Good Barrel electrons.

Based on this study we define knob (c6) to reject 2% of all electron candidates, and (c7) to reject 5% of electron candidates not detected in the Good Barrel region. We expect that knob (c6) will lead to a 2% effect over all search modes (2% for each electron in $\psi \to ee$; but the muon mode is unaffected, leading to a 2% effect in $\psi \to \ell\ell$).

– Discussion, c Knobs

The uncertainties in lepton ID are the major contributors to systematic error in the inclusive $\psi(1S)$ measurement.

As mentioned in the paragraph labeled "Knobs c1-c7," we applied the knobs to the data, instead of the Monte Carlo. Had we applied them to the MC, we would expect that knob c1 affects the 2-track final states by 1.4% and the 4-track final states

by 2.8%; knob c2 would affect the "$\psi(1S)$ from χ_{c1}" mode by 2%; knob c4 would affect all modes by 1%, and knob c6 would affect all knobs by 2%.

These predicted deviations may be compared to the deviations obtained from applying the knobs to the data, in Table 4.6. We find that statistical errors do smear the deviations away from the values that we expected. However, particularly when adding the deviations in quadrature, we see that the deviations obtained are in reasonable agreement with the predictions. The final results would change little if we had applied the knobs to the Monte Carlo instead.

4.2.5 Monte Carlo High-Level Accuracy

Knobs d1 and d2: Monte Carlo ψ Helicity

In the standard procedure, our Monte Carlo sample has an average ψ polarization of $-1/3$ (see Section 4.1.2). This assumption leads to a potential bias because that as discussed earlier, the detection efficiency of high-momentum $\psi \to \mu\mu$ events is sensitive to the ψ polarization (see Figure 4.9).

To gauge the dependence of our result to having possibly used Monte Carlo of the wrong helicity, we repeated the entire momentum measurement using Monte Carlo with both extreme polarizations, i.e. with all ψ longitudinal or transverse. The systematic error is taken to be the average deviation that results from applying these knobs.

Knob d3: $M(\pi^+\pi^-)$ Cut Efficiency

To decrease combinatoric background in $\psi(2S) \to \psi(1S)\pi^+\pi^-$, we require $M(\pi^+\pi^-) > 0.45$ GeV (see Figure 3.3). From the Signal Monte Carlo, we have $\epsilon_{MC} = (81.7 \pm 0.2)\%$. To test this, we obtained a $M(\pi^+\pi^-)$ distribution from data in a manner analogous to finding the p_ψ distribution: the data is partitioned into 4 bins in $M(\pi^+\pi^-)$ and invariant mass peaks are fit for each partition. Doing so, we obtain $\epsilon_{data} = (90.3 \pm 3.1)\%$. However, if we change the signal shape parameterization to a Gaussian sum, as described in Section 4.2.3, we obtain $\epsilon_{data}^{gaussian} = (85.7 \pm 3.6)\%$.

Based on these results we assign an additional systematic error of 5% on the $M(\pi^+\pi^-)$ cut efficiency. Knob d3 accomplishes this by scaling the final $\psi(2S) \to \psi(1S)\pi^+\pi^-$ distribution by 1.05.

This turns out to be the largest systematic error for $\mathcal{B}[B \to \psi(2S)X \to \psi(1S)X]$.

Knob d4: `R2GL` cut

Figure 4.13 compares the distributions of `R2GL`, the ratio of Fox-Wolfram moments, between Monte Carlo and data. The plot is generated by fitting for the yield of $\psi(1S)$ in bins of `R2GL`. In the data, we fit for inclusive $\psi(1S)$, and subtract the scaled contribution from the continuum. The Monte Carlo for this exercise is a mixture of signal Monte Carlo: 73% direct $B \to \psi(1S)X$, 17% $\psi(1S)$ from $\psi(2S)$, and 11% $\psi(1S)$ from χ_{c1}. The exact ratios are given by the observed yields of these modes in the data (see Table 4.6). Furthermore, $\alpha = -1/3$ and there is a 50-50 mix of electron and muon modes. The resulting distributions in `R2GL` are shown in Figure 4.13.

We observe excellent agreement between the data and Monte Carlo. The cut `R2GL`< 0.5 is found to have an efficiency of $(99.69 \pm 0.80)\%$ in the data and $(99.52 \pm 0.20)\%$ in the Monte Carlo. We take the statistical uncertainty of 0.80% as a systematic error for this cut.

Knob d4 increases the yield of data by 0.80% in all modes.

4.2.6 Luminosity-Related Uncertainty

Knob e1: Continuum scale factor

Knob e1 increases the continuum scaling factors for both CLEO II and CLEO II.V by 1%. This has a negligible effect on the final results.

Knob e2: $n_{B\overline{B}}$

Knob e2 increases the yield of data by 2%, which is the uncertainty in $n_{B\overline{B}}$. The value of 2% is taken from past CLEO analyses. This knob affects all search modes and is one of the major systematic errors in measuring $\mathcal{B}[B \to \psi(1S)X]$.

4.3 Results with Systematic Errors

4.3.1 Inclusive Branching Fraction

For each of the 18 knobs for $\psi(1S)$, and 15 knobs for $\psi(2S)$, we obtain efficiency-corrected momentum distributions. By summing over the momentum bins, we obtain total yields. The yields for the standard procedure, as well as the percentage deviation from those yields obtained from activating the knobs, are shown in Table 4.6.

When determining a systematic error, we can either quote a number which attempts to represent a "1σ error", or a larger number which represents an "entire range" of values. We take the former approach to obtain a final systematic error, $\Delta\mathcal{B}_{Syst}$, by combining the knob deviations, Δ_{knob}, as follows:

$$\begin{aligned}
\Delta\mathcal{B}^2_{Syst} &= \Delta^2_{a1} + \Delta^2_{a2} \\
&+[(|\Delta_{b1}| + |\Delta_{b2}|)/2]^2 + [(|\Delta_{b3}| + |\Delta_{b4}|)/2]^2 \\
&+\Delta^2_{c1} + \Delta^2_{c2} + \Delta^2_{c3} + \Delta^2_{c4} + \Delta^2_{c5} + \Delta^2_{c6} + \Delta^2_{c7} \\
&+[(|\Delta_{d1}| + |\Delta_{d2}|)/2]^2 + \Delta^2_{d3} + \Delta^2_{d4} \\
&\Delta^2_{e1} + \Delta^2_{e2}
\end{aligned} \tag{4.2}$$

In 3 cases (knobs b1 and b2 for background fit, knobs b3 and b4 for signal fit, and knobs d1 and d2 for α_{MC}), we consider pairs of knobs to be two measurements of the same systematic source, and so take the average deviation of the knob pair before adding in quadrature.

Using the prescription given in Table 4.2, we normalize the yields obtained by the number of $B\overline{B}$ generated to extract final results for inclusive branching fractions, listing statistical, systematic, and scale errors (uncertainty in the branching fractions of the modes in which we reconstruct ψ), in that order.

$$
\begin{aligned}
\mathcal{B}[B \to \psi(1S)X] &= (1.121 \pm 0.013 \pm 0.040 \pm 0.013)\% \\
\mathcal{B}[B \to \psi(1S)(direct)X] &= (0.813 \pm 0.018 \pm 0.036 \pm 0.010)\% \\
\mathcal{B}[B \to \chi_{c1}X \to \psi(1S)X] &= (0.119 \pm 0.008 \pm 0.009 \pm 0.001)\% \\
\mathcal{B}[B \to \psi(2S)X \to \psi(1S)X] &= (0.189 \pm 0.010 \pm 0.018 \pm 0.002)\% \\
\mathcal{B}[B \to \psi(2S)X] &= (0.316 \pm 0.014 \pm 0.023 \pm 0.016)\%
\end{aligned}
$$

4.3.2 Momentum Distributions

The procedure used above to determine the systematic error of the inclusive branching fractions may be directly applied to each bin of the momentum distributions. Doing so results in systematic error estimations for every momentum bin in $p_\psi(1S)$ and $p_\psi(2S)$. Since the momentum distributions of direct $\psi(1S)$ and direct $\psi(2S)$ are of the greatest interest, we show systematic errors for these distributions only in Tables 4.7 and 4.8.

Knob	$B \to \psi(1S)$ Inclusive	From χ_{c1}	From $\psi(2S)$	Direct	$B \to \psi(2S)$ $\to \ell\ell, \psi(1S)\pi^+\pi^-$
$\mathcal{B} \times 100$	1.121	0.119	0.189	0.813	0.316
$\Delta\mathcal{B}_{stat}$	1.12%	6.73%	5.45%	2.23%	4.43%
a1	0	0	+4.16%	-0.97%	0
a2	0	0	+1.41%	-0.33%	0
b1	-0.25%	-4.40%	+0.57%	+0.17%	+0.60%
b2	-1.24%	+2.85%	+3.32%	-2.89%	-2.93%
b3	-0.41%	+1.59%	+1.96%	-1.25%	+1.72%
b4	+0.33%	+8.32%	+7.11%	-2.41%	+4.43%
c1	-0.91%	-0.07%	-2.48%	-0.67%	-2.29%
c2	-0.38%	-2.00%	+0.01%	-0.23%	+0.02%
c3	-0.38%	-0.37%	+0.01%	-0.47%	+0.02%
c4	-1.04%	-1.69%	-0.31%	-1.11%	-0.67%
c5	-0.36%	-0.82%	-0.05%	-0.37%	-0.19%
c6	-1.93%	-0.85%	-3.26%	-1.78%	-2.87%
c7	-0.65%	+0.14%	-1.49%	-0.57%	-0.77%
d1	+1.23%	+1.05%	+0.85%	+1.34%	+0.69%
d2	-0.77%	-1.00%	-0.50%	-0.80%	+4.49%
d3	0	0	-5.00%	+1.16%	-3.64%
d4	+0.80%	+0.80%	+0.80%	+0.80%	+0.80%
e1	-0.02%	0	0	-0.03%	0
e2	+2.00%	+2.00%	+2.00%	+2.00%	+2.00%
Σa	0	0	4.39%	1.02%	0
Σb	0.83%	6.14%	4.93%	2.39%	3.54%
Σc	2.55%	2.90%	4.37%	2.36%	3.82%
Σd	1.28%	1.30%	5.11%	1.77%	4.54%
Σe	2.00%	2.00%	2.00%	2.00%	2.00%
$\Delta\mathcal{B}_{syst}$	3.58%	7.20%	9.63%	4.41%	7.19%

Table 4.6: $\mathcal{B}(B \to \psi X)$ Systematic Error Results. The first row lists continuum-corrected inclusive branching fractions ($\mathcal{B}$) from the standard procedure, using both electron and muon decays of ψ. The statistical error on $\mathcal{B}$ is shown in the second row. The following rows show the percentage change in $\mathcal{B}$ from each knob. The final six rows list the "subtotal" systematic errors for each category of knobs, and the final systematic error, as defined in Equation 4.2.

	Lower Limit of Momentum Bin (GeV)									
$p_{\psi(1S)}$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\frac{d\mathcal{B}}{dp}(\frac{\%}{GeV})$	0.01	0.03	0.10	0.19	0.22	0.25	0.33	0.41	0.52	0.43
$\Delta\frac{d\mathcal{B}}{dp}_{Stat}$	36.3%	64.2%	25.8%	17.7%	18.4%	19.1%	14.5%	11.8%	9.3%	11.6%
a1	+0.0%	-3.5%	-2.4%	-2.6%	-2.5%	-3.8%	-2.7%	-2.0%	-1.3%	-2.0%
a2	-0.1%	-0.6%	-0.4%	-0.4%	-0.5%	-0.6%	-0.5%	-0.7%	-0.6%	-0.8%
b1	+5.1%	-1.6%	-0.3%	+0.3%	-2.5%	-0.1%	-0.6%	-1.1%	+1.3%	-0.5%
b2	+2.4%	-26.2%	+2.0%	-8.5%	-3.7%	-1.3%	-3.1%	-3.7%	+0.7%	-0.9%
b3	-2.4%	-15.9%	-6.2%	-3.4%	-1.5%	-6.0%	-0.7%	-2.7%	-0.2%	-0.5%
b4	+2.8%	+25.9%	-9.0%	-8.5%	-6.3%	-12.6%	-0.6%	-6.5%	-2.3%	-4.0%
c1	+0.0%	+10.6%	-0.6%	-0.7%	-4.0%	+0.2%	+1.8%	-1.1%	+0.6%	-0.6%
c2	-3.5%	+0.8%	+0.9%	+2.0%	-0.3%	+0.2%	+0.8%	-0.2%	+0.1%	+1.2%
c3	+0.0%	+0.8%	+0.4%	-0.5%	-0.4%	-2.2%	-0.3%	-0.7%	+0.7%	+0.1%
c4	+0.0%	-28.0%	+0.1%	-2.6%	-1.0%	-2.2%	-1.8%	+1.9%	-1.0%	-0.2%
c5	+0.0%	+0.0%	+0.1%	-0.4%	-1.9%	+0.8%	-1.3%	+1.2%	+0.4%	-0.9%
c6	-9.4%	+14.5%	+7.5%	-0.1%	-5.1%	-0.3%	-2.6%	-2.7%	-0.5%	-2.6%
c7	+0.0%	+11.5%	+5.8%	-1.5%	-0.3%	+0.8%	-3.0%	+0.4%	-1.4%	-1.3%
d1	+1.7%	+8.7%	-3.5%	-1.2%	+2.7%	-2.7%	+0.1%	+0.9%	+1.1%	+0.5%
d2	+25.2%	-4.0%	+2.9%	+1.2%	-1.8%	+2.3%	+0.4%	-1.3%	-0.6%	-0.2%
d3	+0.0%	+4.1%	+2.8%	+3.1%	+3.0%	+4.4%	+3.1%	+2.5%	+1.6%	+2.5%
d4	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%
e1	+0.0%	-0.3%	+0.0%	+0.0%	-0.1%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%
e2	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%
$\Delta\frac{d\mathcal{B}}{dp}_{Syst}$	17.5%	44.1%	13.4%	9.5%	9.9%	11.9%	7.1%	7.5%	4.1%	5.7%
$p_{\psi(1S)}$	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
$\frac{d\mathcal{B}}{dp}(\frac{\%}{GeV})$	0.57	0.56	0.62	0.73	0.80	0.85	0.70	0.54	0.25	0.01
$\Delta\frac{d\mathcal{B}}{dp}_{Stat}$	8.6%	8.8%	7.8%	6.0%	6.0%	6.2%	5.5%	6.0%	9.7%	90.4%
a1	-1.1%	-1.2%	-0.8%	-0.3%	-0.3%	-0.1%	+0.0%	+0.0%	+0.0%	+0.0%
a2	-0.6%	-0.6%	-0.3%	-0.2%	-0.1%	-0.1%	-0.1%	+0.0%	+0.0%	+0.0%
b1	+0.8%	-0.6%	-0.6%	+0.5%	+0.6%	+1.4%	+0.5%	+1.0%	-2.1%	-16.8%
b2	+0.1%	-2.3%	-6.8%	-3.2%	-3.2%	-2.0%	-2.9%	-1.8%	-9.8%	-11.7%
b3	-2.1%	+0.5%	-1.6%	-1.1%	-0.3%	-1.5%	-0.5%	-0.8%	-1.1%	+23.6%
b4	-5.4%	-3.3%	-3.0%	-1.0%	-0.4%	-1.9%	+0.0%	+0.3%	+5.5%	+26.8%
c1	-0.4%	+0.2%	+0.2%	-1.2%	-0.8%	-2.3%	-1.3%	-0.8%	+0.5%	-9.8%
c2	-0.4%	-0.5%	-0.7%	-1.2%	-0.7%	+0.0%	-1.0%	-0.1%	+0.2%	-0.4%
c3	-0.7%	-1.1%	-0.6%	-0.9%	-0.4%	-0.3%	-0.9%	+0.2%	+0.2%	-0.4%
c4	-0.2%	-0.8%	-1.4%	-1.7%	-1.4%	-0.8%	-1.6%	-0.8%	-2.3%	-0.4%
c5	-0.7%	-0.1%	-1.4%	-0.2%	-0.1%	-0.1%	-0.5%	-0.9%	-1.0%	-3.1%
c6	-1.5%	-2.1%	-1.6%	-1.4%	-1.8%	-2.2%	-2.3%	-2.6%	-1.3%	-4.5%
c7	-0.3%	-0.3%	+0.0%	-0.1%	-0.6%	-0.6%	-1.3%	-1.0%	-0.7%	-6.1%
d1	-0.2%	-0.1%	+0.8%	+0.6%	+0.7%	+2.9%	+4.3%	+5.2%	+1.6%	+5.0%
d2	+0.3%	-0.1%	-0.6%	-0.4%	-0.4%	-2.1%	-2.8%	-3.3%	-1.1%	-3.2%
d3	+1.3%	+1.5%	+1.0%	+0.4%	+0.3%	+0.1%	+0.0%	+0.0%	+0.0%	+0.0%
d4	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%
e1	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	-0.1%	-0.7%
e2	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%
$\Delta\frac{d\mathcal{B}}{dp}_{Syst}$	5.1%	4.6%	5.8%	4.3%	4.0%	5.3%	5.8%	5.9%	7.9%	32.0%

Table 4.7: $B \to \psi(1S)[direct]X$ momentum spectrum systematic error results.

	Lower Limit of Momentum Bin (GeV)							
$p_{\psi(2S)}$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\frac{d\mathcal{B}}{dp}(\frac{\%}{GeV})$	0.01	0.03	0.07	0.15	0.20	0.26	0.30	0.24
$\Delta\frac{d\mathcal{B}}{dp}_{Stat}$	69.4%	45.0%	28.7%	19.4%	18.2%	14.7%	13.8%	16.9%
b1	+2.1%	-2.6%	-0.1%	-1.0%	-0.6%	+0.2%	+0.2%	+0.2%
b2	-18.8%	+2.9%	-5.1%	+1.5%	+3.8%	-0.4%	-2.1%	+6.3%
b3	+3.2%	+0.3%	+7.9%	+0.6%	+0.9%	+1.4%	+0.0%	+0.0%
b4	-0.1%	-6.7%	+10.4%	+6.4%	+5.7%	+5.2%	+6.1%	+4.9%
c1	+0.0%	-2.8%	-6.1%	-0.8%	-2.9%	+1.2%	-3.0%	+1.5%
c2	+0.0%	-3.0%	+0.7%	+0.3%	+0.3%	+0.5%	-0.1%	-0.4%
c3	+0.0%	-3.0%	+0.7%	+0.3%	+0.3%	+0.5%	-0.1%	-0.4%
c4	+0.0%	-0.2%	-4.3%	-0.8%	-2.1%	-0.2%	-1.6%	+1.5%
c5	+0.0%	+0.0%	+0.0%	-0.8%	-2.5%	+0.7%	-0.1%	+0.3%
c6	+0.0%	+1.7%	-10.4%	-4.4%	-7.9%	-2.8%	-3.2%	+0.5%
c7	+0.0%	+1.7%	-6.2%	-6.3%	-2.6%	+0.0%	+0.8%	-0.4%
d1	+13.0%	+2.5%	+1.8%	+5.3%	-0.2%	+2.1%	-1.3%	-0.2%
d2	-5.4%	+9.2%	-0.6%	+3.4%	+5.0%	+3.3%	+7.2%	+5.1%
d3	-5.0%	-4.4%	-2.5%	-4.2%	-4.2%	-3.5%	-3.5%	-4.1%
d4	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%
e2	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%
$\Delta\frac{d\mathcal{B}}{dp}_{Syst}$	15.1%	10.5%	17.5%	10.8%	11.5%	6.7%	8.2%	7.1%
$p_{\psi(2S)}$	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
$\frac{d\mathcal{B}}{dp}(\frac{\%}{GeV})$	0.20	0.26	0.36	0.32	0.34	0.23	0.15	0.05
$\Delta\frac{d\mathcal{B}}{dp}_{Stat}$	19.1%	15.6%	12.4%	13.4%	13.1%	16.6%	20.8%	52.3%
b1	+6.1%	-1.4%	+0.7%	+1.0%	+0.2%	+1.7%	+0.1%	+3.1%
b2	-4.0%	-11.4%	-1.4%	-8.7%	-6.9%	+1.4%	-8.5%	-5.9%
b3	+0.0%	+5.3%	+1.9%	+2.4%	-0.2%	+1.8%	+0.5%	+18.5%
b4	+4.3%	+4.3%	+4.0%	+4.1%	+3.1%	+1.3%	+2.7%	+11.8%
c1	-7.5%	-4.6%	-4.8%	-2.3%	-2.3%	+1.6%	-1.7%	-1.1%
c2	+0.3%	-0.9%	-0.1%	+1.1%	-0.3%	-1.2%	+1.0%	+2.2%
c3	+0.3%	-0.9%	-0.1%	+1.1%	-0.3%	-1.2%	+1.0%	+2.2%
c4	+0.1%	-3.8%	-0.8%	-0.3%	-0.3%	-0.4%	+0.8%	+5.9%
c5	-0.6%	+0.9%	-0.1%	+1.1%	-0.8%	-2.4%	+0.5%	+2.2%
c6	-3.9%	-1.3%	-4.8%	-3.1%	-1.7%	-0.6%	+0.9%	-5.6%
c7	+0.2%	+1.1%	-1.8%	+0.7%	+0.2%	-3.0%	+0.4%	-1.0%
d1	-0.7%	+2.3%	+0.2%	+1.6%	+0.1%	+1.2%	-2.6%	+0.6%
d2	+6.7%	+3.1%	+5.0%	+2.9%	+3.8%	+5.1%	+6.0%	+1.0%
d3	-3.3%	-3.6%	-3.4%	-3.5%	-3.8%	-3.9%	-2.9%	-4.9%
d4	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%	+0.8%
e2	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%	+2.0%
$\Delta\frac{d\mathcal{B}}{dp}_{Syst}$	11.4%	11.4%	9.1%	8.6%	6.9%	7.5%	7.7%	19.0%

Table 4.8: $B \to \psi(2S)X$ momentum spectrum systematic error results.

4.4 Bin Migration Correction

In this section we adjust our results to correct the assumption made in Section 4.1.4, namely, that "net bin migration" due to mismeasurement smearing (see Section 3.1.2) is negligible[3]. Bin migration refers to the phenomena of detecting and measuring the momentum of the ψ, but obtaining an incorrect measurement, to such an extent that the detected momentum is in a different momentum bin than the true momentum. Bin migration will happen for all momenta of ψ, so in principle it is possible that for every bin, the migration of wrongly measured events into the bin will be equal to the number migrated out of the bin. In such an unlikely case, the "net" bin migration would be zero. It is the "net" bin migration, or, the effect of bin migration, that we must correct for.

We use the Monte Carlo simulation to quantify this effect and correct for it. The degree to which this is a problem for our analysis may be seen in Figure 3.2, which shows a box plot of the reconstructed $\psi(1S)$ momentum versus the generated $\psi(1S)$ momentum for tagged $\psi(1S) \to \ell\ell$ Monte Carlo events. We see that the majority of events are reconstructed in the correct momentum bin (along the diagonal). This justifies our choice of setting the momentum bin size to 0.1 GeV. Of the events that are reconstructed in the wrong momentum bin, most migrate only one bin; however, the tails of the distribution extend out to several bins.

We will describe the algorithm used to correct for this smearing, but first need to justify it with a general discussion of the problem.

4.4.1 General Results

We denote any discrete distribution by the vector $\vec{v}$, where the elements of the vector are positive real numbers and the dimension of the vector, m, is equal to the number of bins. Define $\vec{v}^{obs}$ as the observed (post-convoluted) distribution, and $\vec{v}^{gen}$ as the generated (pre-convoluted) distribution which we wish to recover. It is not

[3] As the final results will show, the momentum spectrum after correcting for measurement smearing is almost identical to the spectrum before the correction; on this basis one could argue that the net bin migration is indeed negligible.

helpful to think of the vector spatially, since we define the norm of the vector to be the sum of its elements:

$$|\vec{v}| \equiv (1, 1, \cdots 1)\vec{v} \tag{4.3}$$

We use the variable $\mathbf{M}$ to denote the 2 dimensional square matrix which describes the convolution. The elements of $\mathbf{M}$ are real numbers between 0 and 1 and represent probabilities. Specifically, M_{rc}, the matrix element of the r-th row and c-th column, is the probability that an event generated in bin number c will be observed in bin number r. The detector matrix $\mathbf{M}$ connects the generated and observed vectors, whereas the inverse, $\mathbf{M}^{-1}$, returns the original distribution given an observed distribution:

$$\vec{v}^{obs} = \mathbf{M}\vec{v}^{gen} \tag{4.4}$$

$$\vec{v}^{gen} = \mathbf{M}^{-1}\vec{v}^{obs} \tag{4.5}$$

In this discussion, we define all matrices $\mathbf{M}$ so that they transform an un-convoluted distribution to a convoluted distribution. When we speak of a transformation that does go the other way, we will use the form $\mathbf{M}^{-1}$. We choose this convention because $\mathbf{M}$ is the "observed" quantity in that we use the Monte Carlo simulation to determine the matrix, and later calculate $\mathbf{M}^{-1}$ to to de-convolute an observed measurement.

The simplest example of a matrix is that of no convolution at all: $\mathbf{M^{perfect}} = \mathbf{I}$, in which case $\vec{v}^{gen} = \vec{v}^{obs}$.

Next, consider a detector which has 100% efficiency but an unknown measurement accuracy; define $\mathbf{M}^{\epsilon=\mathbf{1}}$ as the matrix describing such a detector. All we can say is that every generated event will be observed in one bin or another, which in turn is equivalent to

$$|\vec{v}^{obs}| = |\vec{v}^{gen}| \quad \forall \vec{v}^{gen} \tag{4.6}$$

$$\Leftrightarrow \quad |\mathbf{M}^{\epsilon=\mathbf{1}}\vec{v}^{gen}| = |\vec{v}^{gen}| \quad \forall \vec{v}^{gen} \quad \text{(from Eqn. 4.4)} \tag{4.7}$$

$$\Leftrightarrow \quad (1, 1, \cdots, 1)\mathbf{M}^{\epsilon=\mathbf{1}}\vec{v}^{gen} = (1, 1, \cdots, 1)\vec{v}^{gen} \quad \forall \vec{v}^{gen} \quad \text{(from Eqn. 4.3)} \tag{4.8}$$

$$\Leftrightarrow \quad (1, 1, \cdots, 1)\mathbf{M}^{\epsilon=\mathbf{1}} = (1, 1, \cdots, 1) \tag{4.9}$$

$$\Leftrightarrow \quad \sum_i M_{in}^{\epsilon=1} = 1 \quad \forall n \tag{4.10}$$

in other words, every column of the matrix sums to 1.

Now consider a general convolution matrix $\mathbf{M}$, with dimension m. Using simple arithmetic, we rewrite

$$\mathbf{M} \equiv \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1m} \\ M_{21} & M_{22} & \cdots & M_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mm} \end{pmatrix} \tag{4.11}$$

$$\text{as} \quad \mathbf{M} = \underbrace{\begin{pmatrix} \frac{M_{11}}{\sum_i M_{i1}} & \frac{M_{12}}{\sum_i M_{i2}} & \cdots & \frac{M_{1m}}{\sum_i M_{im}} \\ \frac{M_{21}}{\sum_i M_{i1}} & \frac{M_{22}}{\sum_i M_{i2}} & \cdots & \frac{M_{2m}}{\sum_i M_{im}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{M_{m1}}{\sum_i M_{i1}} & \frac{M_{m2}}{\sum_i M_{i2}} & \cdots & \frac{M_{mm}}{\sum_i M_{im}} \end{pmatrix}}_{\equiv \mathbf{M}^S} \underbrace{\begin{pmatrix} \sum_i M_{i1} & 0 & \cdots & 0 \\ 0 & \sum_i M_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_i M_{im} \end{pmatrix}}_{\equiv \mathbf{M}^E} \tag{4.12}$$

The "S" in matrix $\mathbf{M}^S$ stands for "smearing." Each column of $\mathbf{M}^S$ adds up to 1; therefore, this is a matrix of the $\mathbf{M}^{\epsilon=1}$ variety.

The "E" in diagonal matrix $\mathbf{M}^E$ stands for "efficiency;" each element of $\mathbf{M}^E$ is a detection efficiency. Here "detection efficiency" is defined in a very specific way, and may be obtained as follows: generate N_x Monte Carlo events in a single bin (the x-th bin, in this example). Then, the efficiency of bin x is defined as the fraction of the N_x events observed, in any bin (i.e. events which are detected but migrate to other bins count). However, events which migrate *into* the observed bin from other generated bins do not count, so that the numerator contains only events generated in bin x.

To elaborate further on the relation between the elements of $\mathbf{M}^E$ and efficiency, we note that generating N_x Monte Carlo events in the x-th bin corresponds to setting $\vec{v}^{gen} = (0, \cdots, N_x, \cdots, 0)$, where the nonzero element is the x-th element. The effect of propagating these events through the detector is given by Equation 4.4, i.e. the *observed* momentum distribution is given by $\mathbf{M}\vec{v}^{gen}$. The number of events in this distribution (summed over all possible observed momenta) is $|\mathbf{M}\vec{v}^{gen}|$. Therefore we may write the above-defined efficiency ϵ_x as

$$\epsilon_x = |\mathbf{M}\vec{v}^{gen}|/N_x \tag{4.13}$$

$$= (1, 1, \cdots, 1)\mathbf{M}\vec{v}^{gen}/N_x \tag{4.14}$$

$$= \left(N_x \sum_i M_{ix}\right)/N_x \tag{4.15}$$

$$= \sum_i M_{ix} \tag{4.16}$$

so that substituting into Equation (4.12) yields

$$\mathbf{M}^E = \begin{pmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_m \end{pmatrix} \tag{4.17}$$

To summarize, equation (4.12) is a general result that says that any matrix $\mathbf{M}$ which represents a detector's response may be factorized into a smearing matrix component and a detection efficiency component.

Substituting this result into Equation (4.4) yields

$$\vec{v}^{gen} = \mathbf{M}^{-1}\vec{v}^{obs} \tag{4.18}$$

$$= (\mathbf{M}^S\mathbf{M}^E)^{-1}\vec{v}^{obs} \tag{4.19}$$

$$= (\mathbf{M}^E)^{-1}(\mathbf{M}^S)^{-1}\vec{v}^{obs} \tag{4.20}$$

Being a diagonal matrix, $\mathbf{M}^E$ is trivial to invert. If the degree of bin migration is sufficiently small, it seems that $\mathbf{M}^S$ would also be not too difficult to invert, being quasi-diagonal.

4.4.2 Application to Data

We now apply these results to the momentum distribution measurement.

The definition of detection efficiency used in this analysis (Section 4.1.4), matches that of the previous section. This means that we have already determined the values of the diagonal elements of $\mathbf{M}^E$; they are given by the the right columns of Figures 4.3, 4.6, and 4.7. By applying bin-by-bin efficiency corrections to the data and taking into account the statistical and systematic errors, as we have already done, we have obtained $\mathbf{M}^{E^{-1}}\vec{v}^{obs}$.

We will now apply a smearing correction to the efficiency-corrected data we have already obtained. In other words, we will find $(\mathbf{M}^S)^{-1}(\mathbf{M}^E)^{-1}\vec{v}^{obs}$. Relative to Equation (4.20), however, this presents a commutation problem. The magnitude of this problem can be calculated by noting that $(\mathbf{M}^E)^{-1} \sim I$, and writing $(\mathbf{M}^E)^{-1}$ in such a way that reflects this:

$$(\mathbf{M}^E)^{-1} = \frac{1}{\epsilon}\begin{pmatrix} 1+\delta_1 & 0 & \cdots & 0 \\ 0 & 1+\delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1+\delta_m \end{pmatrix} \tag{4.21}$$

$$\Rightarrow \quad \frac{[(\mathbf{M}^S)^{-1}(\mathbf{M}^E)^{-1} - (\mathbf{M}^E)^{-1}(\mathbf{M}^S)^{-1}]_{rc}}{[(\mathbf{M}^E)^{-1}(\mathbf{M}^S)^{-1}]_{rc}} = \begin{cases} 0 & if \quad r = c \\ \frac{\delta_c - \delta_r}{1+\delta_r} & if \quad r \neq c \end{cases} \tag{4.22}$$

The efficiencies in Figures 4.3, 4.6, and 4.7 are flat as a function of p_ψ to about 15%. Therefore by multiplying by $(\mathbf{M}^S)^{-1}$ on the wrong side, we will obtain a result in which only the off-diagonal elements will be incorrect, by at worst 15% of themselves. The off-diagonal elements are small to begin with (see Equation (4.24)), thus making this a negligible correction.

$\mathbf{M}^S$ describes the measurement error of events that are known to have been detected; this is exactly what is depicted in Figure 3.2. To obtain $\mathbf{M}^S$ from Figure 3.2 it is only necessary to normalize each column individually such that each column sums to 1. Combining electron and muon channels, $\mathbf{M}^S$ and $(\mathbf{M}^S)^{-1}$ are calculated and listed below:

$\mathbf{M}^S \times 1000 =$

$$
\begin{pmatrix}
846 & 19 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
135 & 846 & 39 & 4 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 92 & 870 & 46 & 5 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
12 & 25 & 68 & 862 & 56 & 7 & 4 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
8 & 9 & 10 & 70 & 850 & 58 & 8 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 6 & 6 & 12 & 68 & 853 & 61 & 8 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 3 & 10 & 65 & 852 & 62 & 7 & 4 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 1 & 1 & 5 & 8 & 61 & 852 & 73 & 10 & 4 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 3 & 4 & 8 & 60 & 845 & 73 & 11 & 5 & 3 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 3 & 7 & 57 & 840 & 81 & 11 & 5 & 2 & 2 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 3 & 7 & 59 & 835 & 81 & 12 & 6 & 3 & 2 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 & 6 & 55 & 835 & 84 & 12 & 6 & 3 & 2 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 7 & 56 & 834 & 85 & 12 & 7 & 4 & 2 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 6 & 53 & 824 & 84 & 15 & 7 & 4 & 2 & 4 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 5 & 60 & 824 & 94 & 15 & 7 & 4 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 5 & 59 & 815 & 98 & 15 & 8 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 6 & 55 & 808 & 111 & 18 & 10 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 5 & 57 & 808 & 130 & 17 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 4 & 47 & 799 & 257 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 4 & 36 & 704
\end{pmatrix}
\tag{4.23}
$$

$$(\mathbf{M}^S)^{-1} \times 100 =$$

$$\begin{pmatrix}
119 & -3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-19 & 119 & -5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & -12 & 116 & -6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -2 & -9 & 117 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & -1 & -10 & 119 & -8 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & -1 & -1 & -9 & 119 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & -9 & 119 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -8 & 119 & -10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & -8 & 120 & -10 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 121 & -12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 121 & -12 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 121 & -12 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 122 & -12 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 123 & -12 & -1 & -1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 123 & -14 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 125 & -15 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 126 & -17 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 126 & -21 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -7 & 128 & -47 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & 144
\end{pmatrix} \tag{4.24}$$

Applying the latter matrix to the efficiency-corrected yield vector, and assuming that the bin errors are uncorrelated, we can calculate the efficiency-corrected, smearing-corrected momentum distribution. As a cross-check, we also fit to find this distribution. 20 histograms, are created, each one showing the measured distribution from a single event generated in each of the bins. In other words, the histograms are the rows of Equation (4.23). We then do a χ^2 fit to a linear superposition of the 20 histograms. The two methods yield nearly identical results. We note that the result of the smearing correction is subtle; even at high momentum, where the correction is the most drastic, the change in central values is within 1σ of statistical error.

The statistical errors on each bin increase by about 25% (of themselves) for all momentum bins; the increase occurs because neighboring bins of the corrected dis-

tribution are correlated. The magnitude of the increase is primarily driven by the diagonal elements of Equation (4.24). Though the errors seem to have become worse, we note that if we were to add the bins up to obtain an inclusive branching fraction, proper consideration of the 20×20 error correlation matrix would lead to the same statistical error on the sum of the bins. In other words, the increase in the statistical error applies to the shape of the momentum distribution only, not the sum.

The procedure for obtaining the smearing-corrected $\psi(2S)$ spectrum is identical to that of $\psi(1S)$. There is a small complication in that we reconstruct the $\psi(2S)$ in both the dilepton and $\psi(1S)\pi^+\pi^-$ channels. However, from $\psi(2S)$ MC studies (plots similar to those in Figure 3.2),we find that the invariant mass signal width in both $\psi(2S) \to e^+e^-$ and $\psi(2S) \to \psi(1S)[\to e^+e^-]\pi^+\pi^-$ is 38 MeV. Similarly, we find resolutions of 24 and 25 MeV, respectively, for the muon modes $\psi(2S) \to \mu^+\mu^-$ and $\psi(2S) \to \psi(1S)[\to \mu^+\mu^-]\pi^+\pi^-$. The nearly identical widths justify our using the $\psi(2S) \to e^+e^-$ to deconvolute the smearing in $\psi(2S) \to \psi(1S)[\to e^+e^-]\pi^+\pi^-$ (and likewise for the corresponding muon decays.)

4.5 Final Results, Momentum Distribution

The final results of efficiency-corrected, measurement smearing-corrected momentum distributions of $\psi(1S)$ and $\psi(2S)$ produced directly from B decays are displayed in Figure 4.14 and listed in the final columns of Table 4.9 [for $\psi(1S)$] and Table 4.10 [for $\psi(2S)$].

4.6 Discussion of Results

4.6.1 Comparing to CLEO Monte Carlo

The distributions of $\psi(1S)$ and $\psi(2S)$ as produced in the signal Monte Carlo are compared to that of the efficiency-corrected data in the left column of Figure 4.15. There are significant differences between MC and data, especially in the feed-down modes.

	$d\mathcal{B}/dp$ (%/GeV)	
$p_\psi(1S)$ (GeV)	Uncorrected	Corrected
$0.0 - 0.1$	0.015 ± 0.005	0.017 ± 0.007
$0.1 - 0.2$	0.026 ± 0.016	0.021 ± 0.020
$0.2 - 0.3$	0.096 ± 0.05	0.095 ± 0.029
$0.3 - 0.4$	0.190 ± 0.034	0.192 ± 0.039
$0.4 - 0.5$	0.218 ± 0.040	0.215 ± 0.048
$0.5 - 0.6$	0.247 ± 0.047	0.237 ± 0.056
$0.6 - 0.7$	0.328 ± 0.048	0.324 ± 0.057
$0.7 - 0.8$	0.413 ± 0.049	0.398 ± 0.058
$0.8 - 0.9$	0.524 ± 0.049	0.533 ± 0.058
$0.9 - 1.0$	0.432 ± 0.050	0.402 ± 0.061
$1.0 - 1.1$	0.572 ± 0.049	0.574 ± 0.060
$1.1 - 1.2$	0.557 ± 0.049	0.537 ± 0.059
$1.2 - 1.3$	0.621 ± 0.048	0.605 ± 0.059
$1.3 - 1.4$	0.734 ± 0.044	0.736 ± 0.055
$1.4 - 1.5$	0.804 ± 0.048	0.778 ± 0.059
$1.5 - 1.6$	0.851 ± 0.053	0.859 ± 0.064
$1.6 - 1.7$	0.700 ± 0.038	0.700 ± 0.048
$1.7 - 1.8$	0.537 ± 0.032	0.554 ± 0.040
$1.8 - 1.9$	0.250 ± 0.024	0.268 ± 0.031
$1.9 - 2.0$	0.014 ± 0.013	0.002 ± 0.019

Table 4.9: Final Results, $B \to \psi(1S)(direct)X$ momentum spectrum. "Uncorrected" is the spectrum without correcting for measurement smearing, and corresponds to Figure 4.12. "Corrected" is the spectrum after correcting for measurement smearing, and corresponds to Figure 4.14. The final systematic errors for each bin are shown in Table 4.7.

This discrepancy between the momentum distributions of Monte Carlo and data has the potential to bias the polarization measurement (if, for example, the polarization were a strongly dependent function of p_ψ.) We therefore take steps to change the Monte Carlo so that it better matches the data. A "rejection skim" program is executed for the signal Monte Carlo in which the generated lab-frame momentum of the ψ determines the event's probability of survival. The values of the probabilities are calculated from the above results. After running the signal Monte Carlo events through this skim, the momentum distributions of $\psi(1S)$ and $\psi(2S)$ more closely match those of the data. This may be seen in the right ("after skim") column of

	$d\mathcal{B}/dp$ (%/GeV)	
$p_\psi(2S)$ (GeV)	Uncorrected	Corrected
$0.0-0.1$	0.008 ± 0.008	0.010 ± 0.008
$0.1-0.2$	0.028 ± 0.028	0.028 ± 0.016
$0.2-0.3$	0.067 ± 0.067	0.061 ± 0.023
$0.3-0.4$	0.154 ± 0.154	0.161 ± 0.037
$0.4-0.5$	0.196 ± 0.196	0.190 ± 0.044
$0.5-0.6$	0.262 ± 0.262	0.262 ± 0.047
$0.6-0.7$	0.298 ± 0.298	0.307 ± 0.050
$0.7-0.8$	0.244 ± 0.244	0.239 ± 0.050
$0.8-0.9$	0.196 ± 0.196	0.181 ± 0.047
$0.9-1.0$	0.262 ± 0.262	0.252 ± 0.051
$1.0-1.1$	0.359 ± 0.359	0.372 ± 0.056
$1.1-1.2$	0.315 ± 0.315	0.306 ± 0.053
$1.2-1.3$	0.336 ± 0.336	0.348 ± 0.055
$1.3-1.4$	0.230 ± 0.230	0.235 ± 0.049
$1.4-1.5$	0.151 ± 0.151	0.155 ± 0.040
$1.5-1.6$	0.050 ± 0.050	0.049 ± 0.034

Table 4.10: Final Results, $B \to \psi(2S)(direct)X$ momentum spectrum. "Uncorrected" is the spectrum without correcting for measurement smearing, and corresponds to Figure 4.12. "Corrected" is the spectrum after correcting for measurement smearing, and corresponds to Figure 4.14. The final systematic errors for each bin are shown in Table 4.8.

Figure 4.15.

We use the corrected Monte Carlo in the polarization measurement.

4.6.2 Comparing to Other Measurements and Theory

The most recent CLEO analysis of inclusive $B \to \psi$ was based on 2.02 fb^{-1} of $\Upsilon(4S)$ data [35]. Among the results were branching fraction measurements: $\mathcal{B}(B \to \psi(1S)X) = 1.12 \pm 0.04 \pm 0.06$ and $\mathcal{B}(B \to \psi(1S)(direct)X) = 0.80 \pm 0.08$. These results are the primary measurements used by PDG 2001. With 4.5 times as much data, the statistical errors in this analysis have been reduced by a factor of 2, and the systematic errors have also been improved. The results reported here are consistent with these earlier findings.

Recent results reported from BaBar [64]: $\mathcal{B}(B \to \psi(1S)X) = 1.044 \pm 0.013 \pm 0.028$, $\mathcal{B}(B \to \psi(1S)(direct)X) = 0.789 \pm 0.010 \pm 0.034$, and $\mathcal{B}(B \to \psi(2S)X) = 0.275 \pm 0.020 \pm 0.029$ are consistent with our findings to within 2σ. Despite having twice the integrated luminosity, their systematic errors dominate and the overall precision of our results is equivalent to that of BaBar.

Several improvements were made relative to the last CLEO publication in the momentum spectrum measurement. Our analysis subtracts *measured* momentum distributions of $\psi(1S)$ from $\psi(2S)$ and $\psi(1S)$ from χ_{c1}, whereas the previous analysis was forced to resort to MC because of low statistics. This is a significant difference, as the left column of Figure 4.15 shows. We also subtract the measured continuum charmonium contribution. Our final results in Figure 4.14 include a correction for detector measurement smearing and include systematic error calculations for every momentum bin.

Figure 4.14 also shows the contributions to the momentum distributions from exclusive two-body decays $B \to \psi Y$, where $Y \in \{\pi, K, K^*, K_1(1270)\}$. A significant part of the observed $\psi(1S)$ and $\psi(2S)$ distributions lies below the two-body region of p_ψ, showing that both $B \to \psi(1S)X$ and $B \to \psi(2S)X$ have substantial multibody components. This was noted soon after the first CLEO momentum distribution measurement, and was interpreted as being consistent with significant color-octet ($c\overline{c}$) production [29]. The reason is that color-octet $c\overline{c}$ states must emit a gluon to become a ψ; the emitted gluon is likely to create additional light hadrons in the final state.

A peak in $p_\psi(2S)$ at about 650 MeV is observed that was not apparent in the previous CLEO measurement (Figure 4.12). Figures 4.6 and 4.7 show that this effect can nominally be seen in each of the 4 independent data sets that we employ to measure the $p_{\psi(2S)}$ spectrum. Taking the overlayed two-body modes into account, the new bump may be simply be interpreted as being the peak of the multibody $B \to \psi(2S)X$ component; similarities between the distributions of multibody $\psi(1S)$ and $\psi(2S)$ are apparent.

Section 1.6.3 briefly discussed a recent conjecture that the discovery of a bump at $p_{\psi(1S)} \sim 500$ GeV might be a signature of triple charm decay $b \to c\overline{c}c$ [39]. We observe no such enhancement in this region. The hint of a bump which was seen in CLEO

II data was most likely a statistical fluctuation. The CLEO II and CLEO II+II.V momentum spectrum results (where both results are uncorrected for measurement smearing and show statistical errors only) are shown for $B \to \psi(direct)X$ in Figure 4.12.

We hope that the final momentum distributions shown in Figure 4.14, Table 4.9, and Table 4.10 will prove useful in further constraining the color octet NRQCD matrix elements.

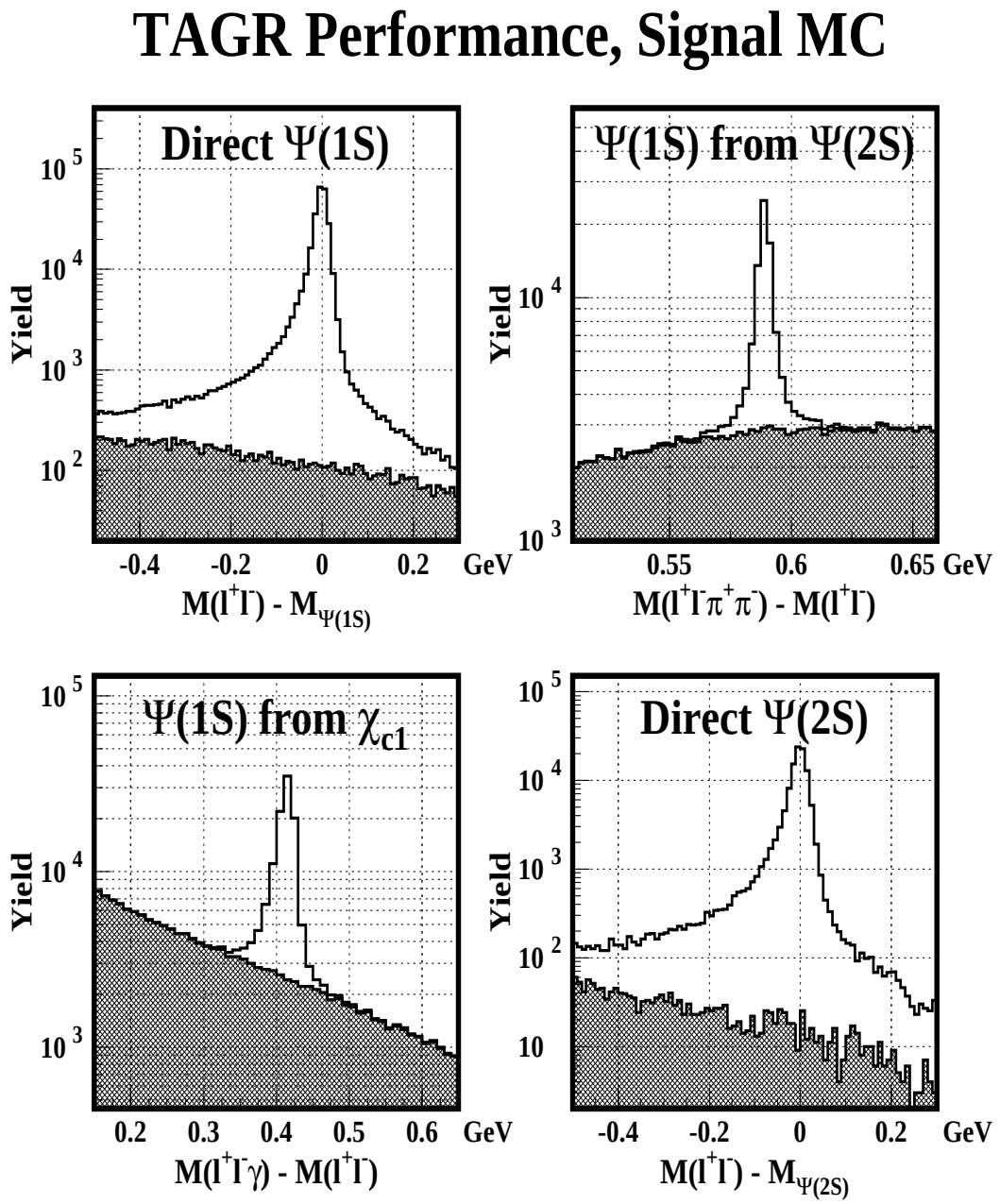

Figure 4.1: CLEO Tagger performance on signal Monte Carlo of the four search modes. Shaded area is the combinatoric background component as determined by TAGR. The vertical scale is logarithmic. In the $\psi \to \ell\ell$ modes, the beginning of the low-side radiative tail can be seen. The selected signal events are used as the signal lineshapes.

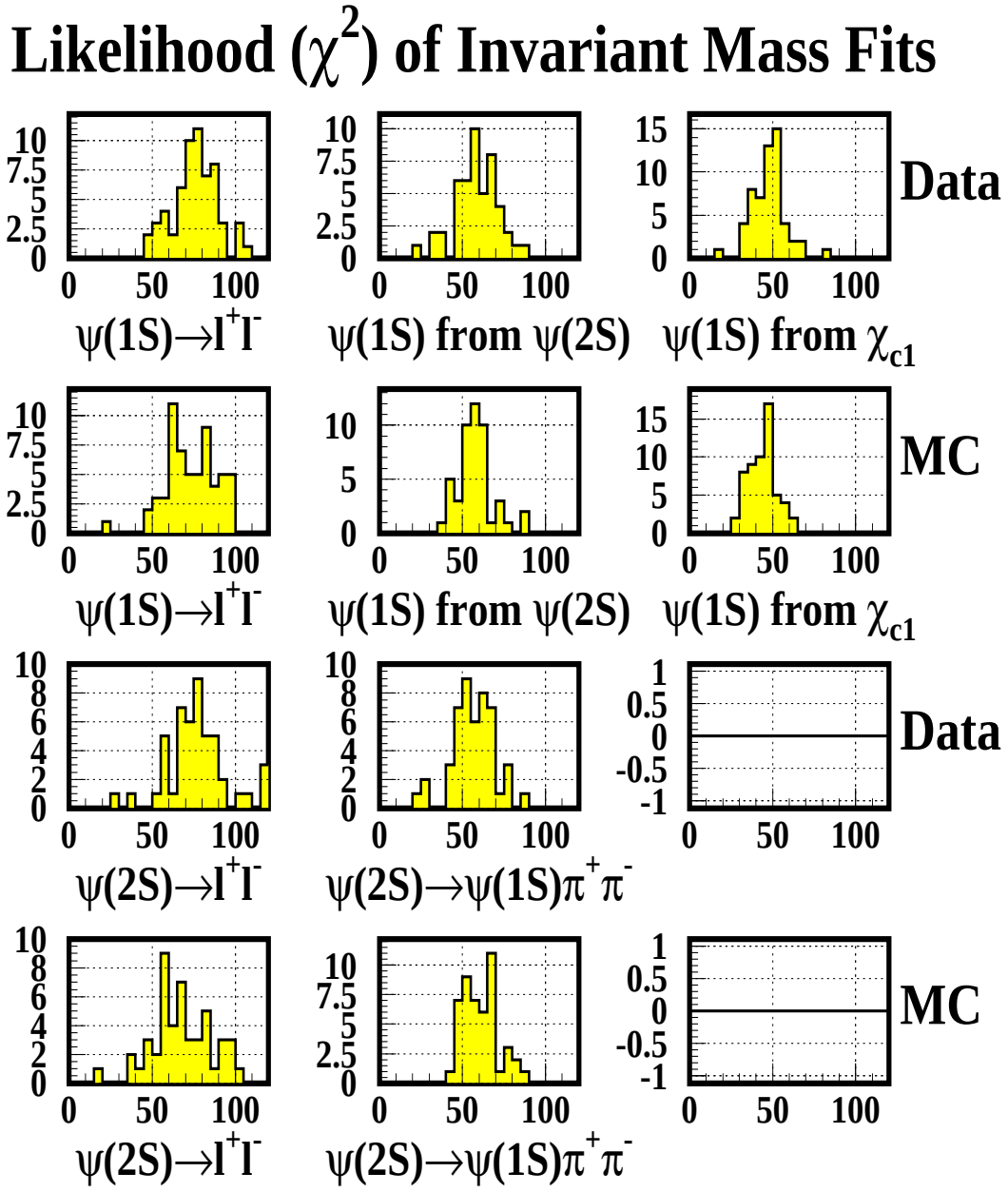

Figure 4.2: Distributions of the Likelihood function (χ^2) of the 552 fits for yield in the momentum distribution measurement.

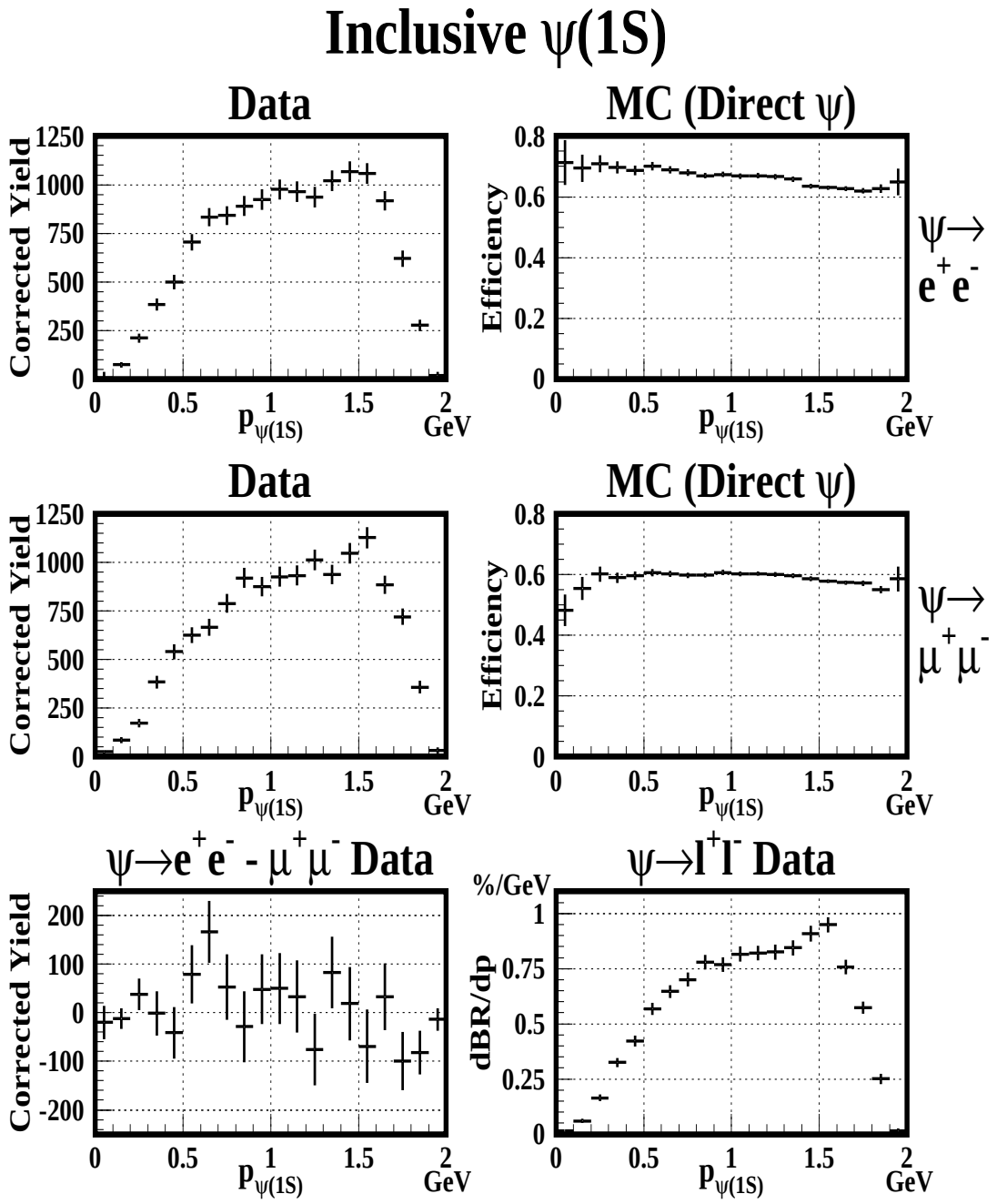

Figure 4.3: "$\psi(1S)$ inclusive" momentum distributions. Top left: Efficiency-corrected data, inclusive $\psi(1S) \to ee$. Top right: Detection efficiency, direct $\psi(1S) \to ee$. Middle left: Efficiency-corrected data, inclusive $\psi(1S) \to \mu\mu$. Middle right: Detection efficiency, direct $\psi(1S) \to \mu\mu$. Bottom left: Efficiency-corrected data, difference between $\psi(1S) \to ee$ and $\psi(1S) \to \mu\mu$ modes. Bottom right: Efficiency-corrected data, $\psi(1S) \to \ell\ell$, with the vertical axis rescaled in terms of $\mathcal{B}(B \to \psi(1S)X)$. As discussed in Sec. 4.1.4 and shown in Figure 4.8, the efficiencies for direct $\psi(1S) \to ee$ and inclusive $\psi(1S) \to ee$ are the same for a given bin in $p_{\psi(1S)}$.

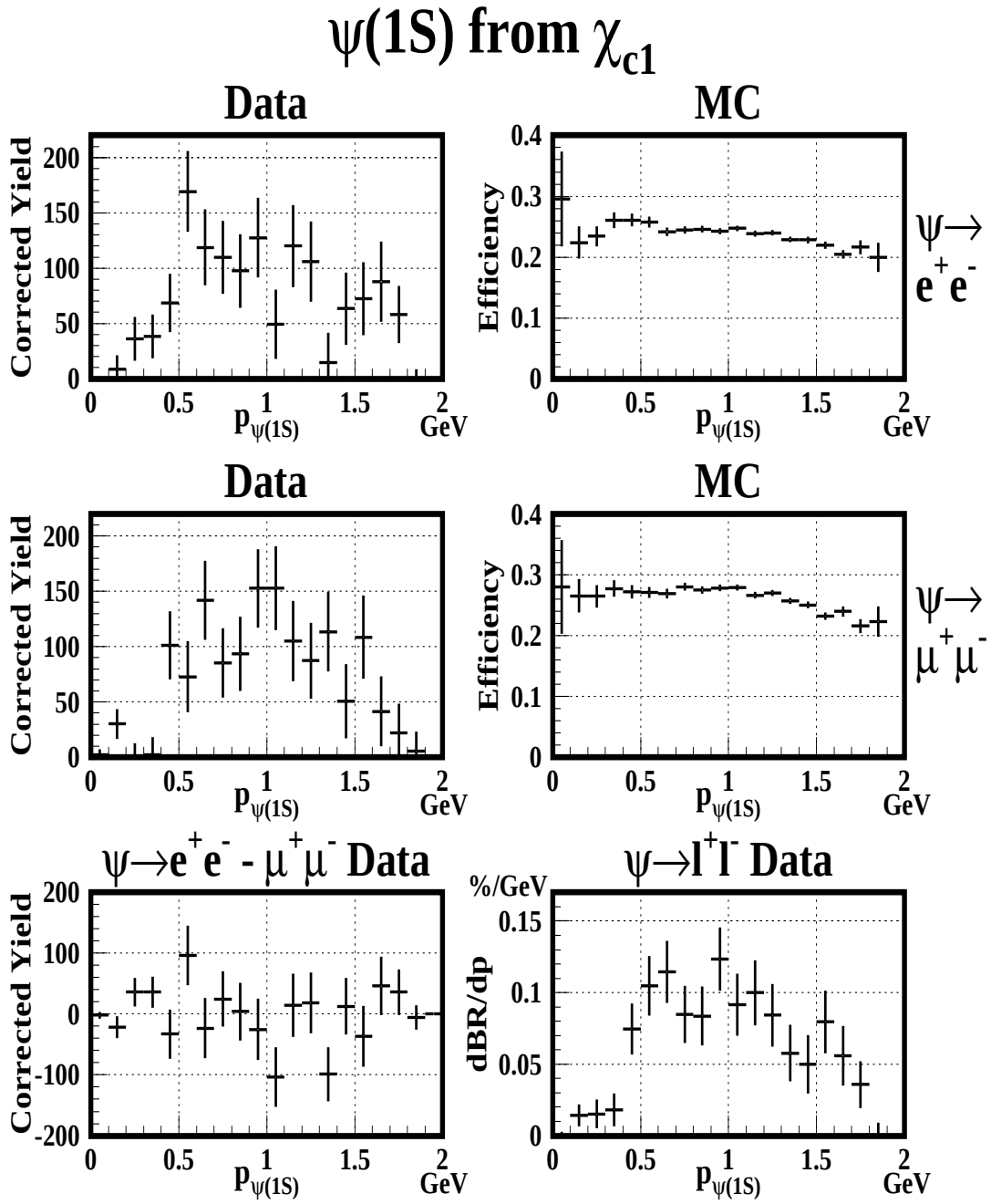

Figure 4.4: "$\psi(1S)$ from χ_{c1}" momentum distributions. Top row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to ee$. Middle row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to \mu\mu$. Bottom left: Efficiency-corrected data, difference between $\psi(1S) \to ee$ and $\psi(1S) \to \mu\mu$ modes. Bottom right: Efficiency-corrected data, $\psi(1S) \to \ell\ell$, with the vertical axis rescaled in terms of $\mathcal{B}(B \to \psi(1S)X)$.

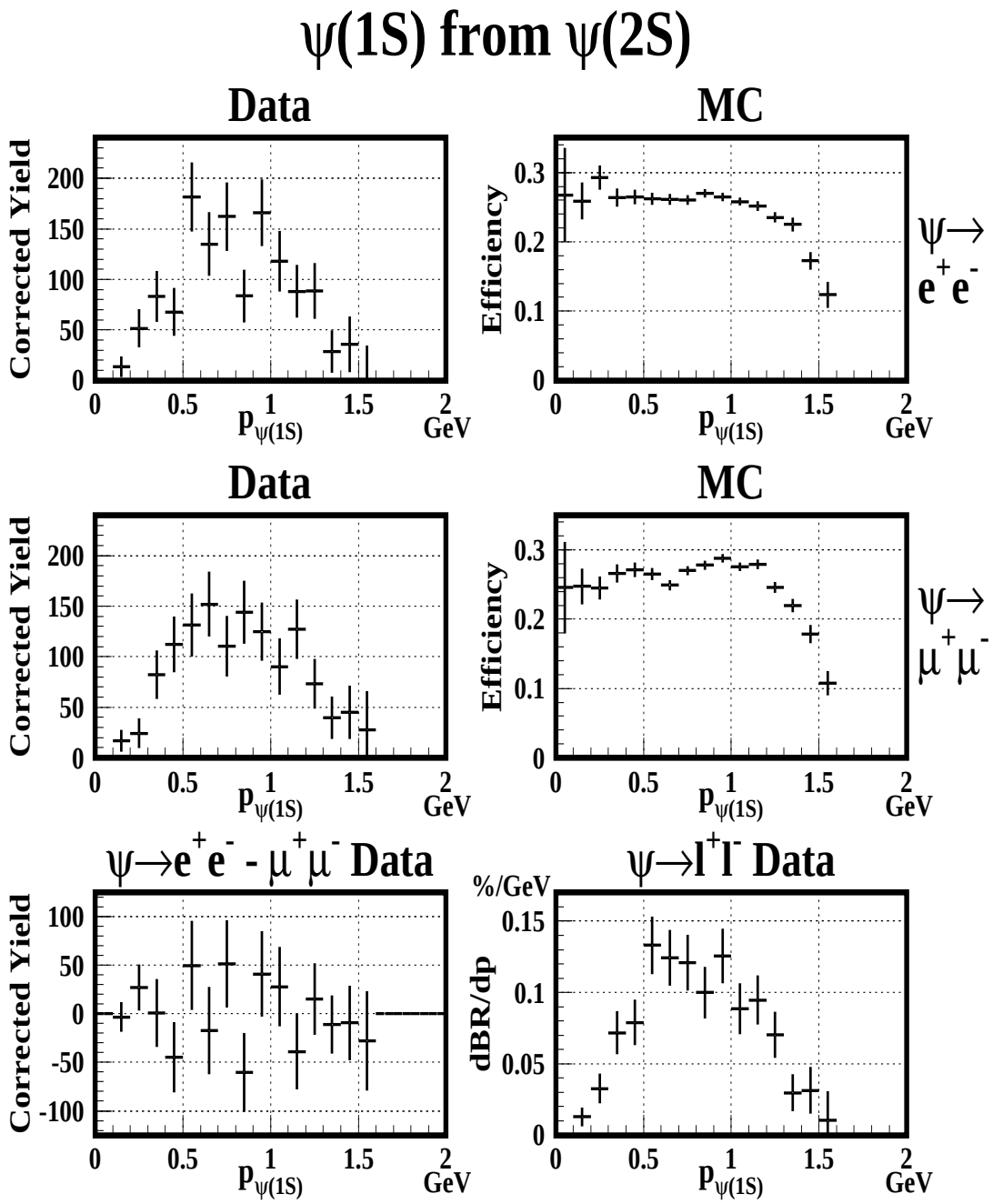

Figure 4.5: "$\psi(1S)$ from $\psi(2S)$" momentum distributions. Top row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to ee$. Middle row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to \mu\mu$. Bottom left: Efficiency-corrected data, difference between $\psi(1S) \to ee$ and $\psi(1S) \to \mu\mu$ modes. Bottom right: Efficiency-corrected data, $\psi(1S) \to \ell\ell$, with the vertical axis rescaled in terms of $\mathcal{B}(B \to \psi(1S)X)$.

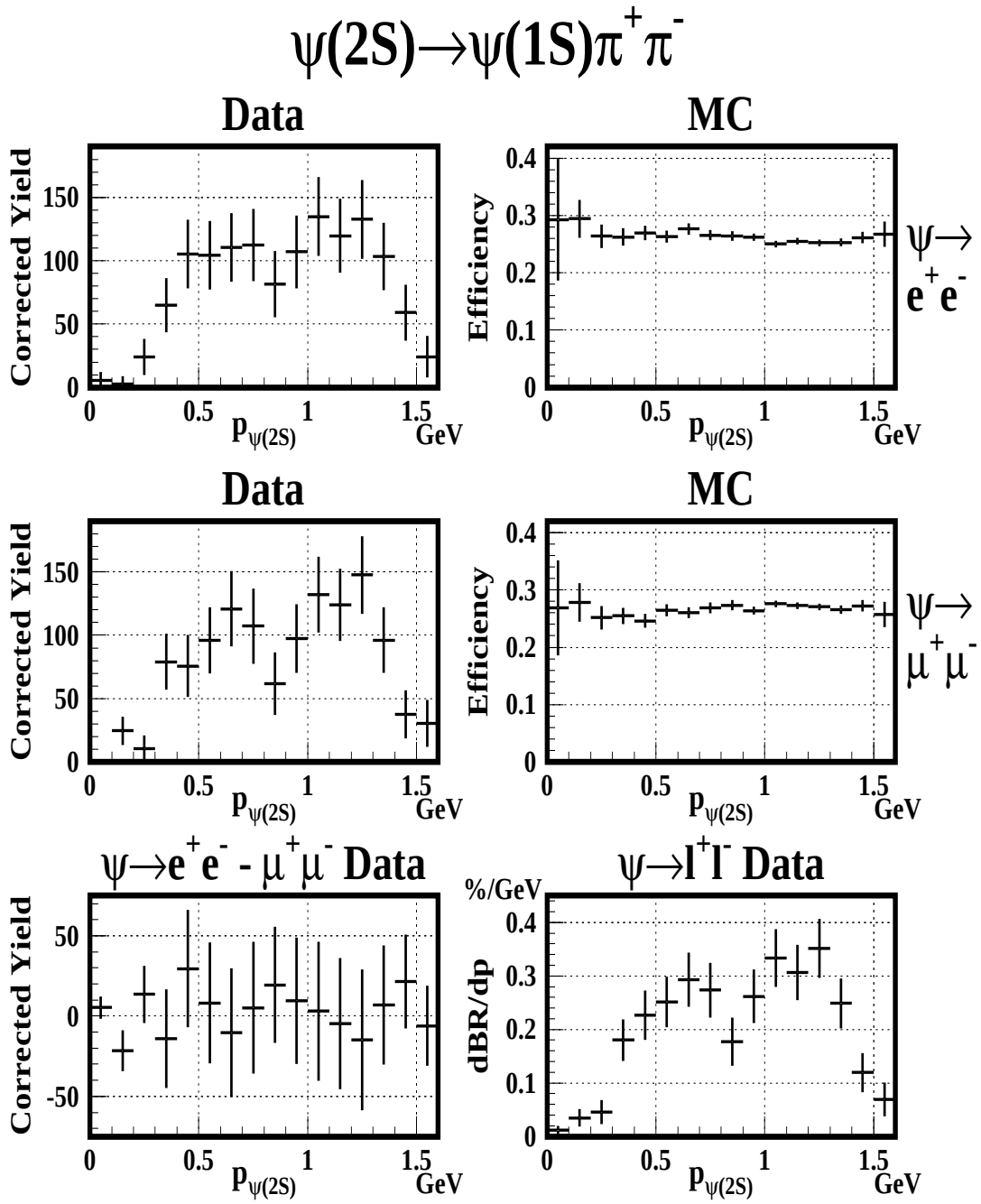

Figure 4.6: "$\psi(2S)$ via $\psi\pi\pi$" momentum distributions. Top row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to ee$. Middle row: Efficiency-corrected data and detection efficiency, $\psi(1S) \to \mu\mu$. Bottom left: Efficiency-corrected data, difference between $\psi(1S) \to ee$ and $\psi(1S) \to \mu\mu$ modes. Bottom right: Efficiency-corrected data, $\psi(1S) \to \ell\ell$, with the vertical axis rescaled in terms of $\mathcal{B}(B \to \psi(2S)X)$.

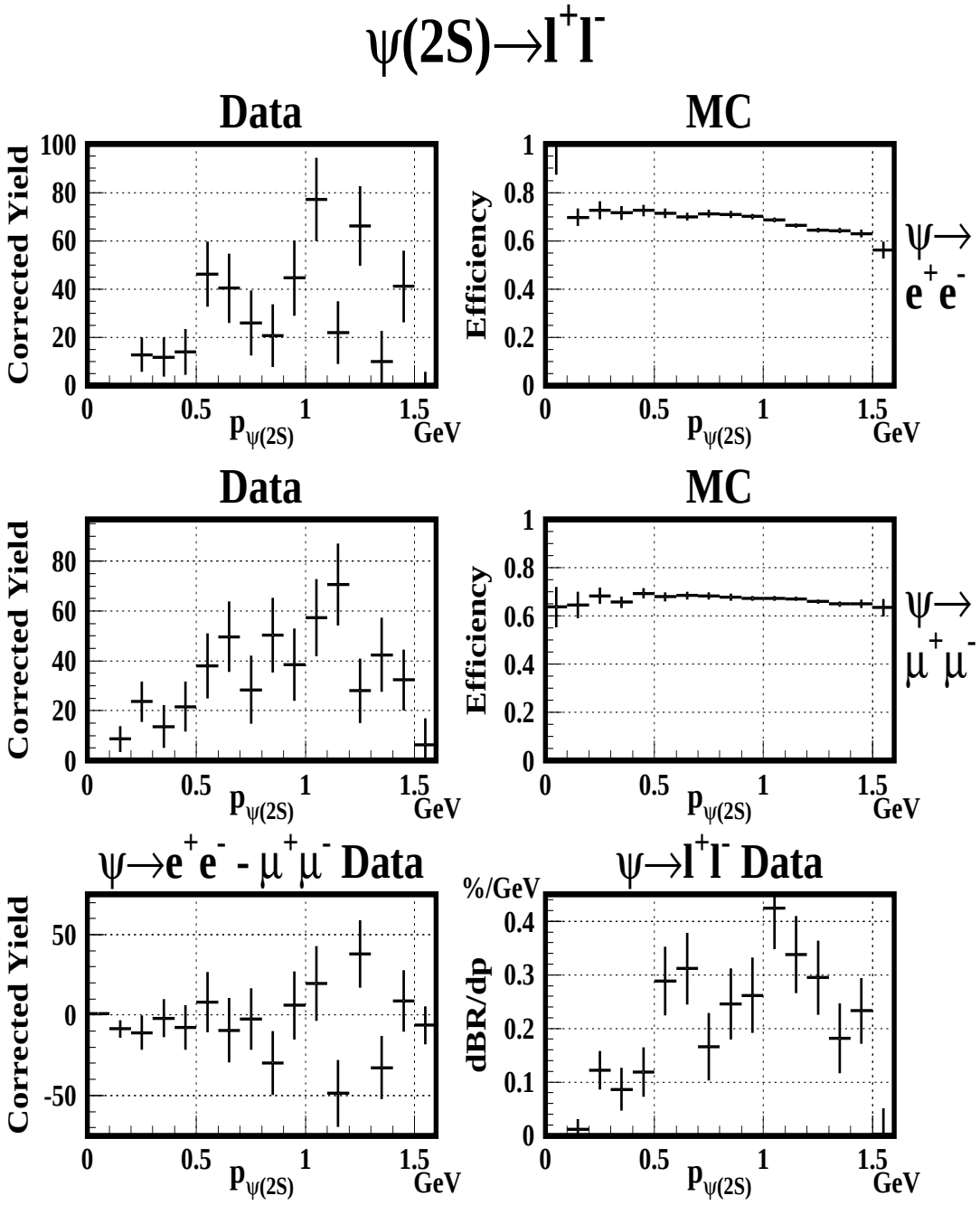

Figure 4.7: "$\psi(2S)$ via dileptons" momentum distributions. Top row: Efficiency-corrected data and detection efficiency, $\psi(2S) \to ee$. Middle row: Efficiency-corrected data and detection efficiency, $\psi(2S) \to \mu\mu$. Bottom left: Efficiency-corrected data, difference between $\psi(2S) \to ee$ and $\psi(2S) \to \mu\mu$ modes. Bottom right: Efficiency-corrected data, $\psi(2S) \to \ell\ell$, with the vertical axis rescaled in terms of $\mathcal{B}(B \to \psi(2S)X)$.

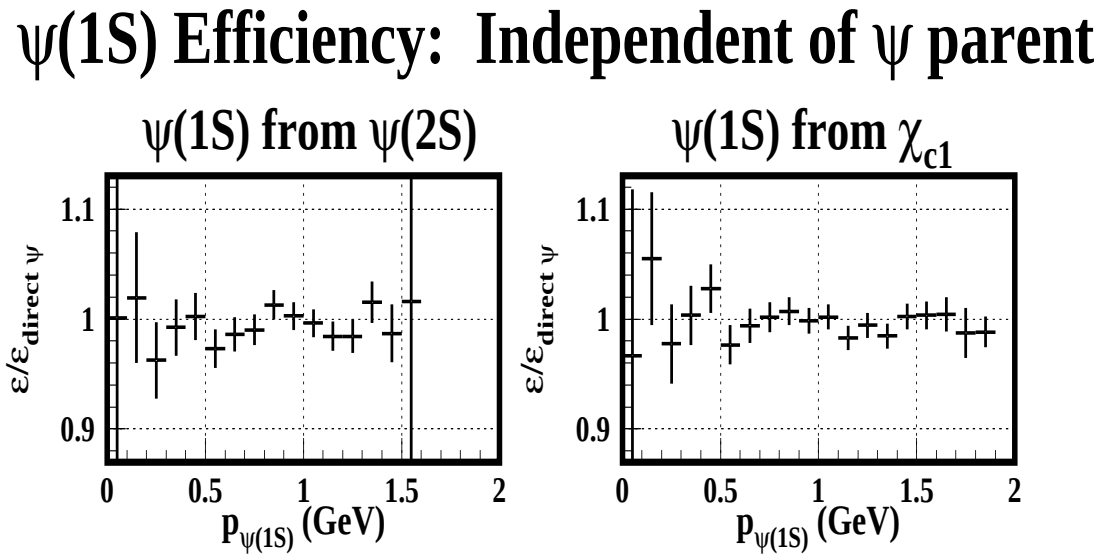

Figure 4.8: Efficiency of finding $\psi(1S)$ as a function of the $\psi(1S)$ parent, from signal Monte Carlo. Vertical axis, left: $\frac{\epsilon^{parent=\psi(2S)}_{\psi(1S)}}{\epsilon^{parent=B}_{\psi(1S)}}$. Vertical axis, right: $\frac{\epsilon^{parent=\chi_{c1}}_{\psi(1S)}}{\epsilon^{parent=B}_{\psi(1S)}}$. H=-1/3 signal Monte Carlo events. In these plots, efficiency is defined as the probability of detecting the $\psi(1S)$. These plots demonstrate that the probability of detecting the $\psi(1S)$ is independent of the direct parent of the $\psi(1S)$; therefore, Monte Carlo efficiency estimations for directly produced $\psi(1S)$ may be applied to inclusively produced $\psi(1S)$ in the data. See Sec. 4.1.4 for further discussion.

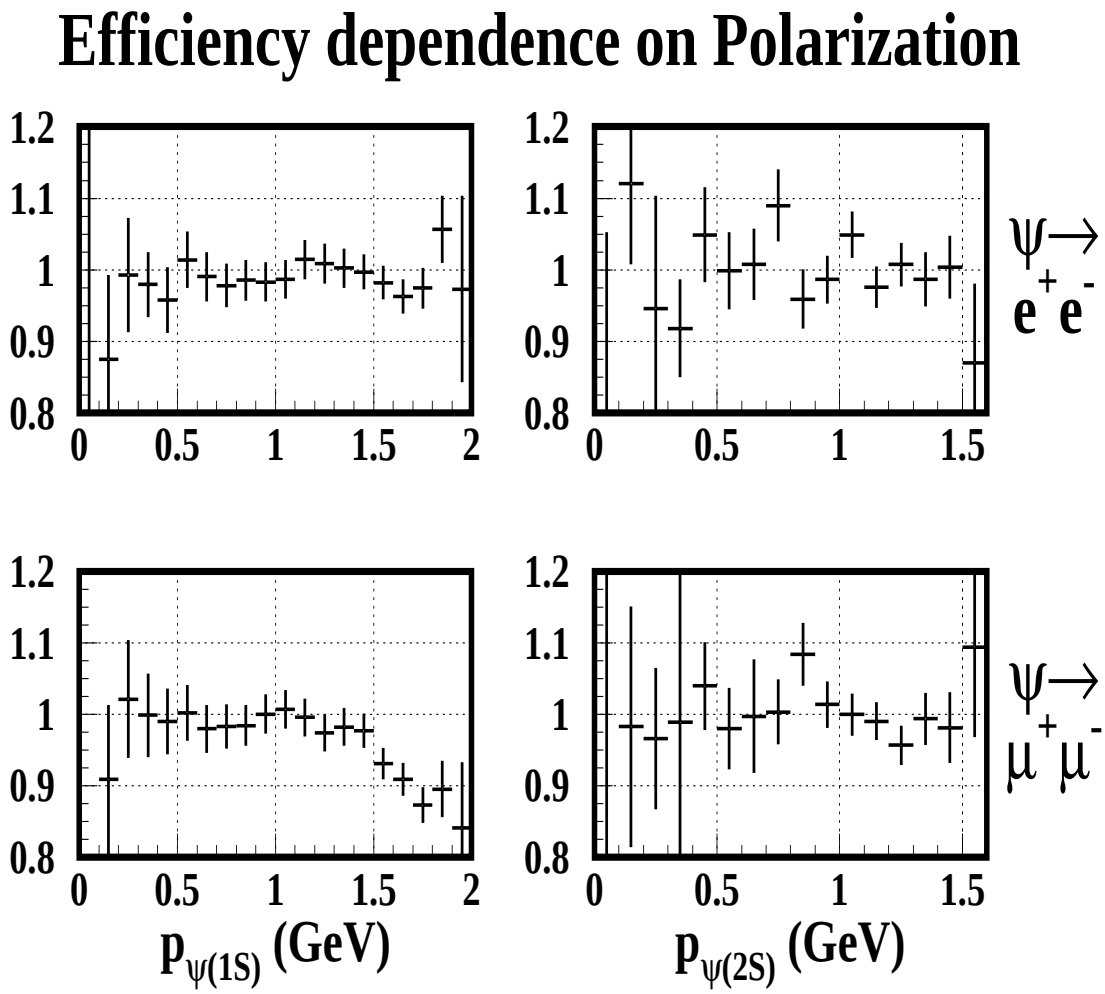

Figure 4.9: The effect of $\psi(1S)$ polarization on detection efficiency, from signal Monte Carlo. Vertical axis: $\frac{\epsilon^{transverse}_{\psi(1S)direct}}{\epsilon^{longitudinal}_{\psi(1S)direct}}$. Horizontal axis: $p_{\psi(1S)}$. Top: $\psi(1S) \to ee$. Middle: $\psi(1S) \to \mu\mu$. The dip in efficiency at high $p_{\psi(1S)}$ for $\psi(1S) \to \mu\mu$ occurs because high-momentum $\psi(1S)$ which decay preferentially along the direction of flight will produce one fast muon and one slow muon.

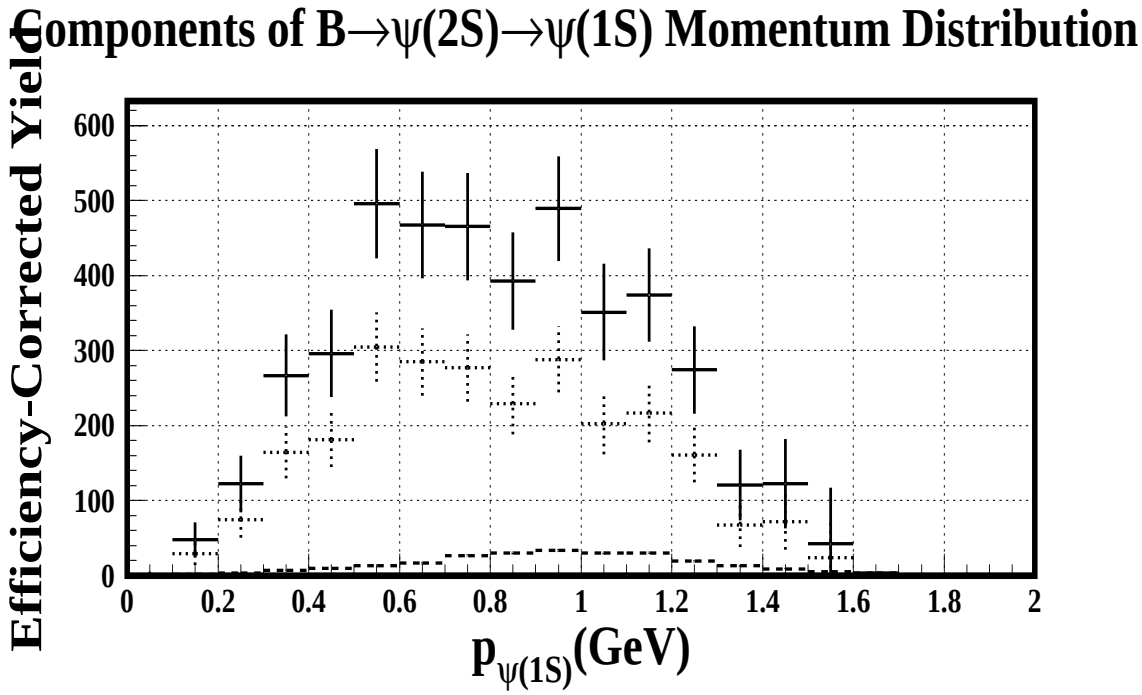

Figure 4.10: $B \to \psi(2S) \to \psi(1S)$ Momentum Distribution Components. Solid: Total, efficiency-corrected distribution, from the sum of $\psi(2S) \to \psi(1S)\pi^+\pi^-$, $\psi(2S) \to \psi(1S)\pi^0\pi^0$, and $\psi(2S) \to \psi(1S)X_{misc}$ decays. Dotted: Measured contribution from $\psi(2S) \to \psi(1S)\pi^+\pi^-$. $\psi(1S) \to e^+(\gamma)e^-(\gamma)$ or $\mu^+\mu^-$. Dashed: Expected contribution of $\psi(2S) \to \psi(1S)X_{misc}$, from MC.

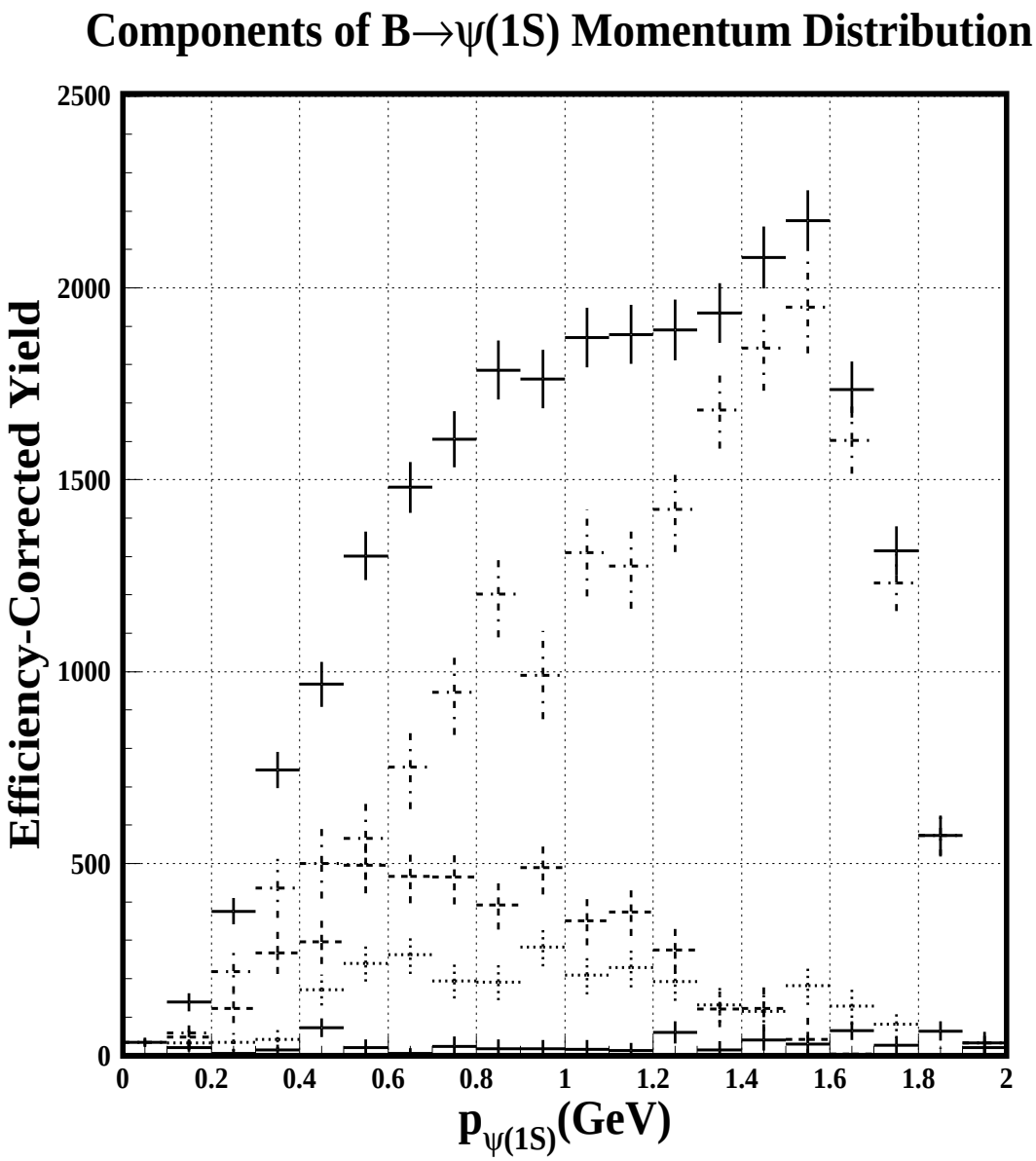

Figure 4.11: $B \to \psi(1S)$ Momentum Distribution Components. Top Solid: Inclusive $\psi(1S)$ (measured). Dot-Dashed: $\psi(1S)$ directly produced from B (inferred). Dashed: $\psi(1S)$ from $\psi(2S)$ (measured). Dotted: $\psi(1S)$ from χ_{c1} (measured). Bottom Solid: Continuum $\psi(1S)$ (measured). $\psi(1S) \to e^+(\gamma)e^-(\gamma)$ or $\mu^+\mu^-$.

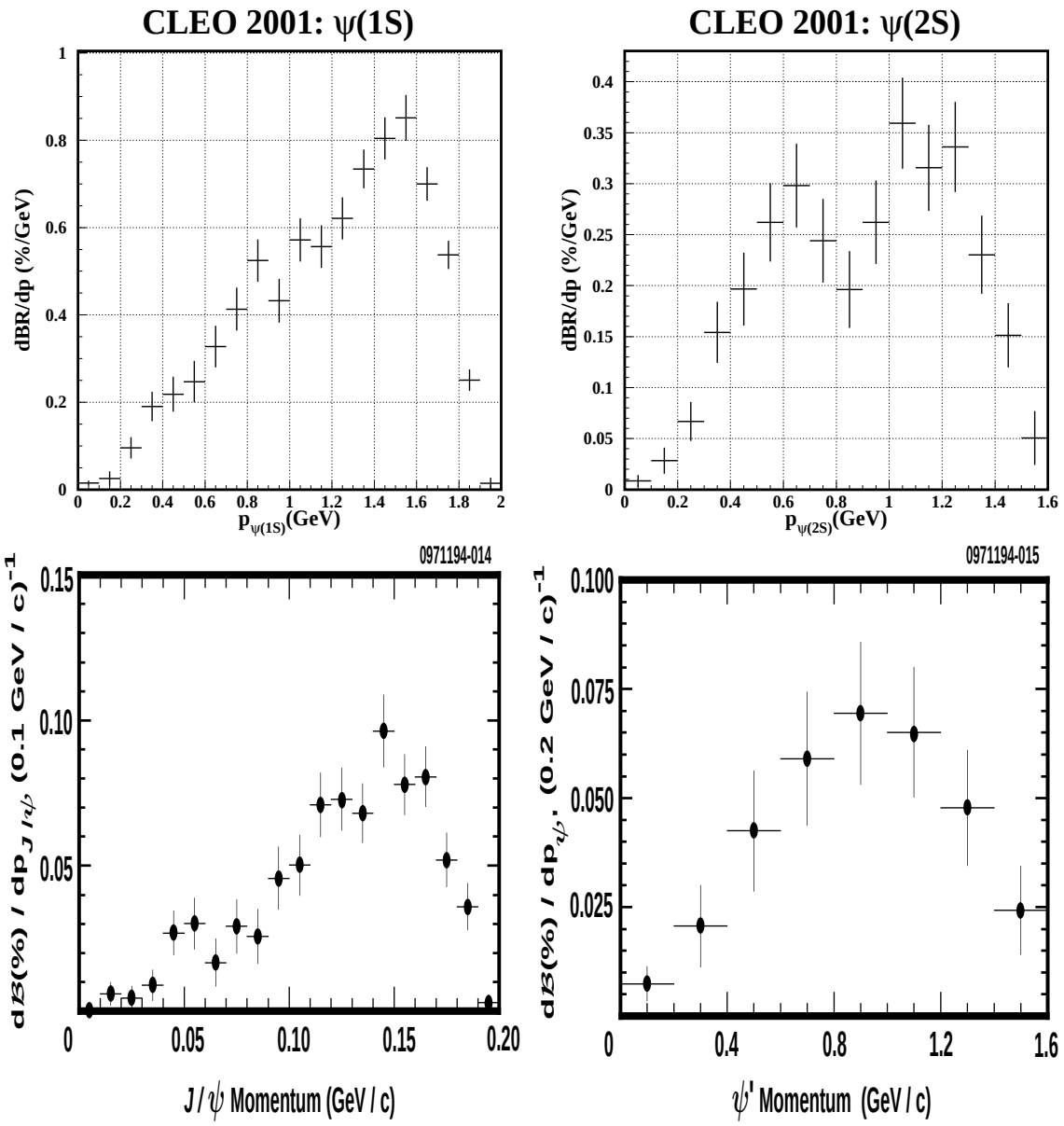

Figure 4.12: Momentum Distribution Results, efficiency-corrected data. Vertical error bars are statistical only. Top: CLEO 2001 results. Bottom: CLEO 1995 results. Left: Momentum spectrum, direct $B \to \psi(1S)X$. Right: Momentum spectrum, direct $B \to \psi(2S)X$.

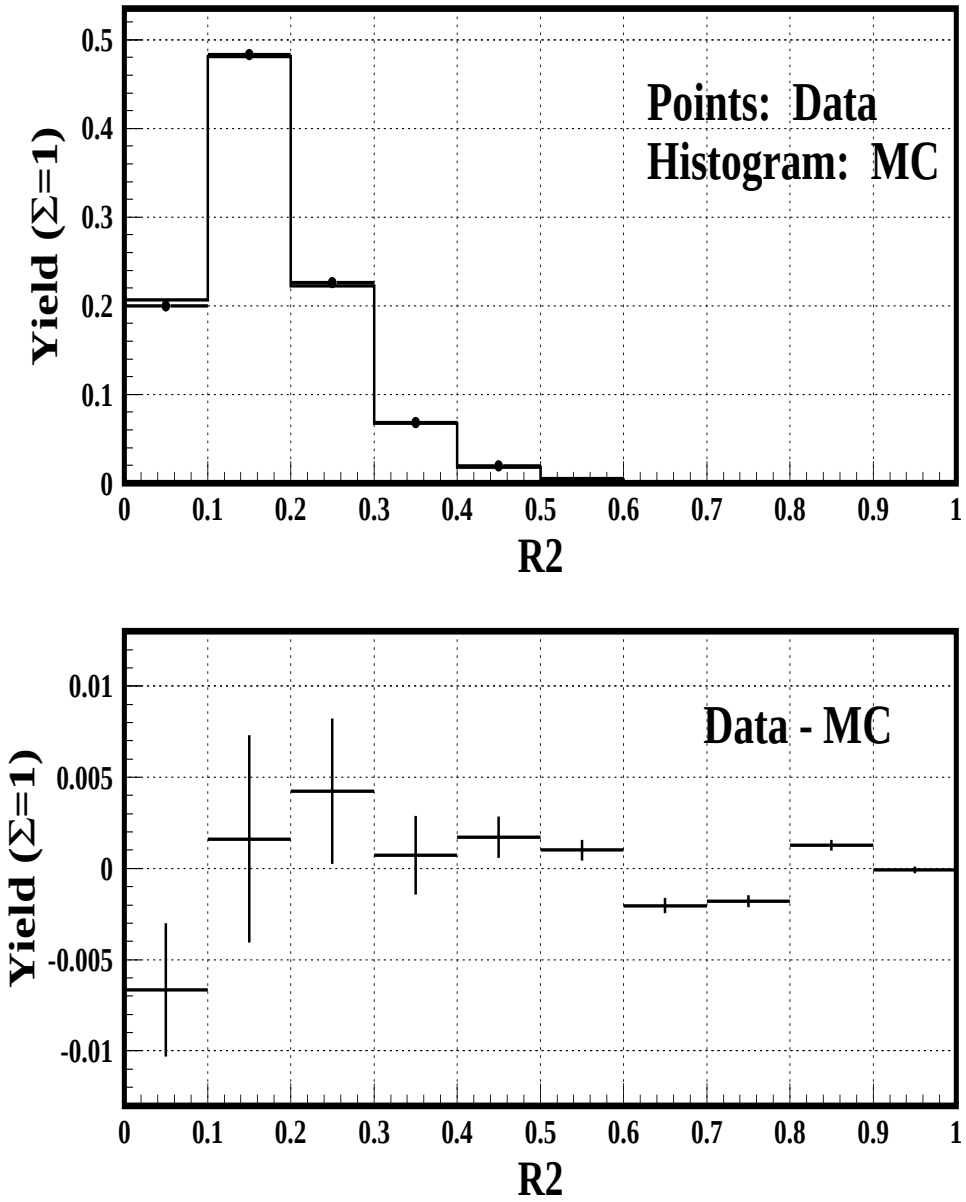

Figure 4.13: R2GL: Data vs. Monte Carlo. The top plot compares the two, the bottom plot shows their difference. Both are normalized so that the total area of the histograms is 1.

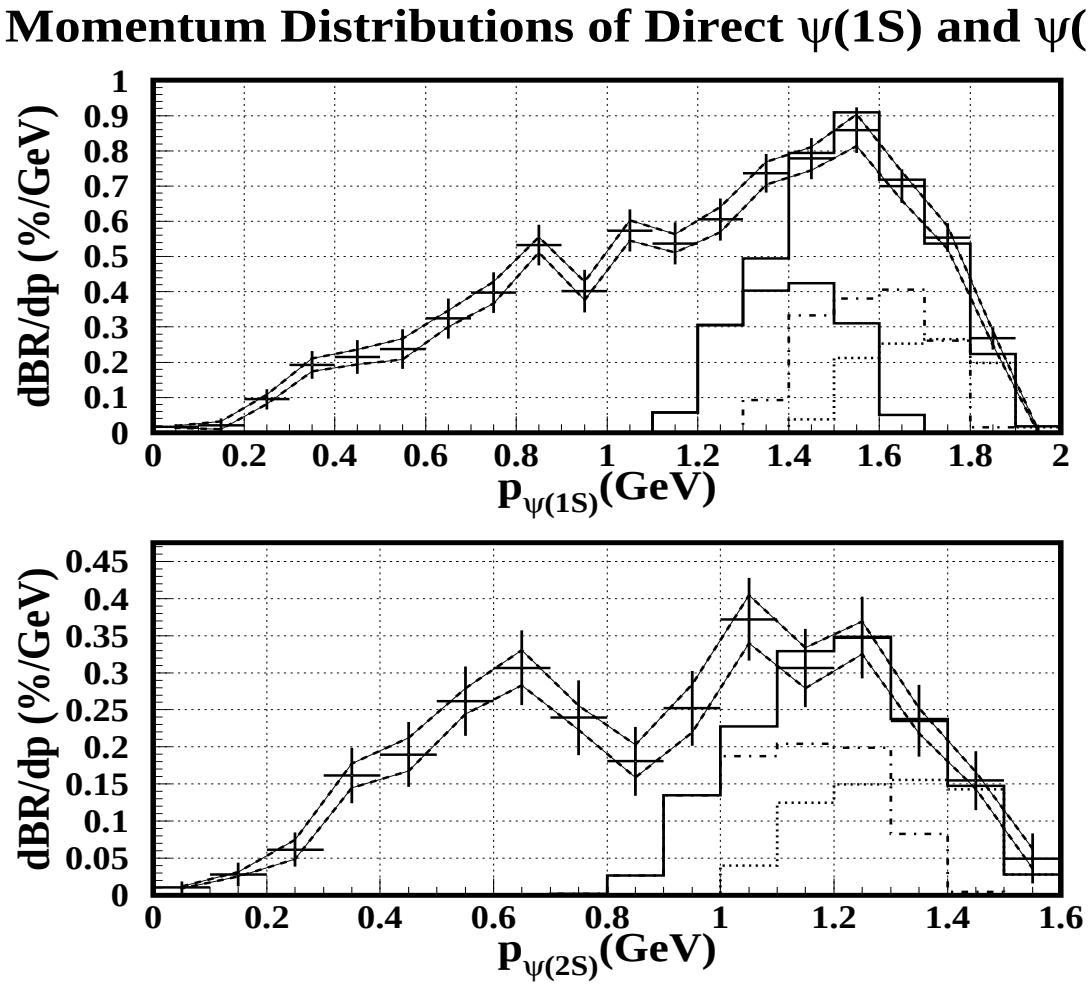

Figure 4.14: Final results, momentum distributions of $\psi(1S)$ and $\psi(2S)$ produced directly from B decay. Data is corrected for continuum $\psi(1S)$, detector efficiency, and detector smearing. Vertical bars are statistical errors; the region between the two lines is the systematic error region. For the $\psi(1S)$ distribution, the top solid histogram is the contribution, obtained from Monte Carlo simulation, from the two-body decays $B \to \psi(1S)Y$, where $Y \in \{\pi, K, K^*, K_1(1270)\}$. The lower solid histogram is from $B \to \psi(1S)K_1$, the dot-dash histogram is from $B \to \psi(1S)K^*$, and the dotted histogram is from $B \to \psi(1S)K$ (the contribution from $B \to \psi(1S)\pi$ is small.) For the $\psi(2S)$ distribution, the top solid histogram is the contribution, obtained from Monte Carlo simulation, from the two-body decays $B \to \psi(2S)Z$, where $Z \in \{K, K^*\}$. The dot-dash histogram is from $B \to \psi(2S)K^*$, and the dotted histogram is from $B \to \psi(2S)K$. The branching fractions used to set the normalization of each exclusive-mode histogram were obtained from the 2001 PDG, except for $B \to \psi(1S)K_1$ [75] and $B \to \psi(2S)Z$ [76]. As of this writing, there are no known two-body decays with non-upper-limit branching fractions, other than those listed above. There is an additional scale uncertainty (not shown on the plots) of 1.2% for $\psi(1S)$ and 5.1% for $\psi(2S)$ due to uncertainty in the branching fractions that we reconstruct the ψ in (see Section 4.1.6).

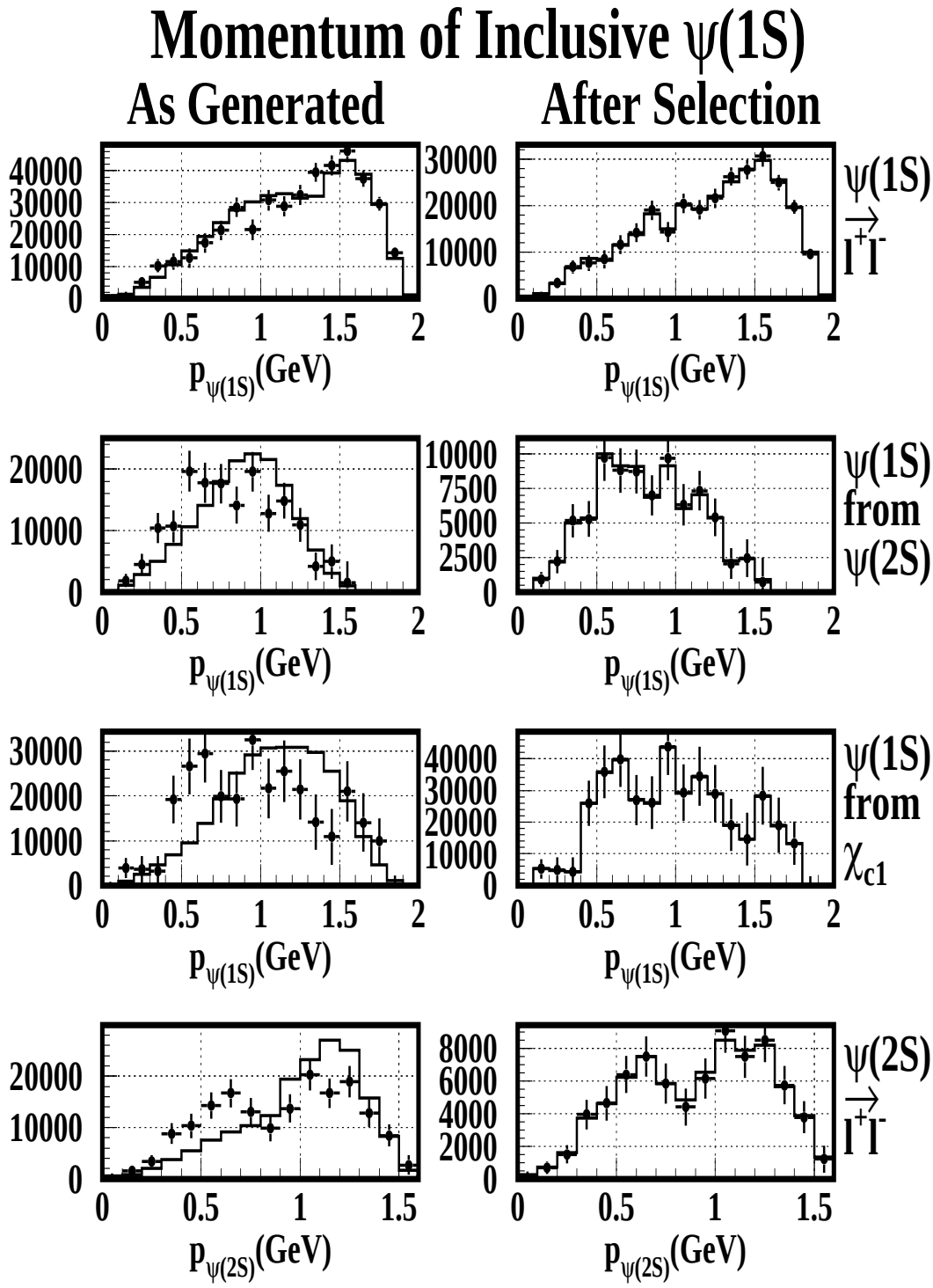

Figure 4.15: MC versus data: points show data (efficiency- and smearing-corrected), histograms show Monte Carlo (as generated). Relative normalization is equal-area. Left: Momentum distributions of $\psi(1S)$ as generated by the CLEO Monte Carlo. Right: The same, but after a rejection technique is used to force the MC distributions to conform to the data; these are the MC event sets which are used in the polarization measurement.

Chapter 5

Polarization Measurement

5.1 Analysis

5.1.1 Analysis Outline

The polarization variable α is simply a parameter in one dimension which describes the shape of the distribution in $\cos\theta$. So, just as in the momentum distribution measurement, which was constructed by combining ψ yield measurements in partitions of p_ψ, in this measurement we obtain angular distributions by grouping together yield measurements in partitions of $\cos\theta$.

As in the momentum distribution measurement, we must again take into account differences in efficiency across the angular distribution, as well as the fact that about 30% of the $\psi(1S)$ from B decays have intermediate parents, $\psi(2S)$ or χ_{c1}.

Ideally, one could obtain α by fitting the efficiency-corrected angular distribution from the data, to the function $1 + \alpha\cos^2\theta$. However, as discussed in Section 3.1.1, when we measure $\cos\theta$, we use p_ψ as measured in the lab frame when we should ideally use p_ψ as measured in the B frame. Therefore we actually measure $\cos\theta^{lab}$; the difference between this smeared measurement and the actual $\cos\theta$ is shown in the lower right plot of Figure 3.1.

To solve this problem, we extract the polarization from the *measured* distributions – that is, distributions that suffer from both the B-frame smearing discussed above,

as well as measurement smearing. Instead of attempting to unsmear data so that we can fit it to an exact function $(1 + \alpha \cos^2 \theta)$, we generate Monte Carlo events with known angular distributions, propagate them through the Monte Carlo simulation to include the effects of B-frame and measurement smearing, then compare the resulting distributions to the data.

As discussed in Section 3.6, we generate signal Monte Carlo in two separate subsets: one with the ψ sample completely transversely polarized, and the other, with all ψ completely longitudinally polarized. Said another way, the Monte Carlo generates events so that in the limit of infinite statistics, the angular distributions of the two subsets are exactly $1 + \cos^2 \theta$ and $1 - \cos^2 \theta$, respectively. We note that θ in this context is the "true" B-frame θ, not the smeared measurement θ^{lab}. These events are then propagated through the detector simulation, thus incorporating B-frame and measurement smearing. Angular distributions are then constructed by fitting for yields in bins of $\cos \theta^{lab}$, in exactly the same manner as the data. This gives us smeared angular distributions, in $\cos \theta^{lab}$, of the "pure-polarization" Monte Carlo. By fitting the (smeared) data angular distribution to a weighted sum of the two Monte Carlo "pure-polarization" smeared angular distributions, we determine the number of transversely and longitudinally polarized ψ in the data. Finally, we use Equation 1.52 to extract α.

This strategy works to the extent that we can trust the Monte Carlo to get the smearing right. For symmetric colliders, in the reaction $e^+e^- \to \Upsilon(4S) \to B\bar{B}$, the energy of the B meson is given by the beam energy. The fact that the beam energy is slightly larger than the rest mass of the B meson is the source of B-frame smearing. Modeling this smearing is therefore a matter of simple kinematics. However, it is important to note that the Monte Carlo simulation also incorporates the subtle variations in beam energy [77] that have occurred over the life of CLEO.

Relative to B-frame smearing, smearing due to mismeasurement is a comparatively subtle effect. Evidence that the Monte Carlo also models this source of smearing extremely well is found in our comparisons of data versus Monte Carlo signal lineshapes in the $\psi(1S) \to \ell\ell$ mode, as discussed earlier in Section 4.2.3 (systematic knob b4.)

5.1.2 Signal Lineshapes

Signal lineshapes were obtained from signal Monte Carlo. The procedures used to obtain the signal lineshapes are identical to those of the momentum distribution measurement (see Section 4.1.2). However, the Monte Carlo in this analysis is "post-rejection-skim" (see Figure 4.15), so that the Monte Carlo momentum distributions of $\psi(1S)$ and $\psi(2S)$ match those of the data.

5.1.3 Invariant Mass Fits

In obtaining yields from invariant mass fits, the only difference between the polarization and momentum distribution measurements is in the definition of the partitions. We use the same fitting procedure in both measurements.

For the polarization analysis, there are a total of 720 ($5\times3\times4\times3\times4$) on-resonance partitions to consider, each involving a mass fit, and defined as follows:

- $\cos\theta$ [1] (5):
  - $-1.0 \leq \cos\theta < -0.6$
  - $-0.6 \leq \cos\theta < -0.2$
  - $-0.2 \leq \cos\theta < +0.2$
  - $+0.2 \leq \cos\theta < +0.6$
  - $+0.6 \leq \cos\theta < +1.0$
- Data source (3):
  - Signal MC, ψ Helicity $= \pm1$
  - Signal MC, ψ Helicity $= 0$
  - 4s2-4sT on-resonance data
- Search mode (4):

[1] Starting from here, we will use $\cos\theta$ to denote either $\cos\theta$ or $\cos\theta^{lab}$; after reading the discussion in Section 5.1.1, it should be clear from context which is being referred to. Since $\cos\theta^{lab}$ is what we actually measure, this is the usual meaning.

- "$\psi(1S)$ inclusive"
- "$\psi(1S)$ from χ_{c1}"
- "$\psi(1S)$ from $\psi(2S)$"
- "$\psi(2S)$ via dileptons"

- Dilepton mode lepton species (3):
  - electrons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$
  - muons: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow \mu^+\mu^-$
  - both: $\psi(1S)$ [or $\psi(2S)$] $\rightarrow e^+(\gamma)e^-(\gamma)$ or $\mu^+\mu^-$

- p_ψ (measured momentum) (4):
  - "low": 0.0 GeV$\leq p_{\psi(1S)} <$ 0.8 GeV, 0.0 GeV$\leq p_{\psi(2S)} <$ 0.7 GeV
  - "mid": 0.8 GeV$\leq p_{\psi(1S)} <$ 1.4 GeV, 0.7 GeV$\leq p_{\psi(2S)} <$ 1.1 GeV
  - "high": 1.4 GeV$\leq p_{\psi(1S)} <$ 2.0 GeV, 1.1 GeV$\leq p_{\psi(2S)} <$ 1.6 GeV
  - "all" : 0.0 GeV$\leq p_{\psi(1S)} <$ 2.0 GeV, 0.0 GeV$\leq p_{\psi(2S)} <$ 1.6 GeV

We also generate angular distributions of continuum data. Based on Figure 3.5, we assume that the production rate of continuum $\psi(2S)$ and χ_{c1} is negligible. Furthermore, because of limited statistics, we do not partition by momentum. The 15 (5×3) partitions for off-resonance data are defined as follows:

- $\cos\theta$ (5): as above
- Data source (1): 4s2-4sT off-resonance data
- Search mode (1): "$\psi(1S)$ inclusive"
- Dilepton mode lepton species (3): as above
- $p_{\psi(1S)}$ (measured momentum) (1): 0.0 GeV$\leq p_{\psi(1S)} <$ 2.0 GeV

Our choice of binning in $\cos\theta$ is justified by several observations. Since the angular distributions of longitudinally and transversely polarized ψ both involve only even powers of $\cos\theta$, we could have binned in $|\cos\theta|$ instead. However, this would have removed a useful "symmetry check" of our distributions, with no improvement in statistical power. So, we fit in $\cos\theta$ over the interval [-1, 1]. We also wish to define the bins in such a way that does not diminish the differentiation between the endpoints, $\cos\theta = \pm 1$, and the midpoint, $\cos\theta = 0$. This restricts us to an odd number of bins. The minimum number of bins is therefore 3. Now, the fitter returns N and α, i.e. the total number of ψ and their polarization. With 2 parameters in the fit, a 3-bin fit would give little information on fit quality. We therefore choose the next largest number of bins, 5. The binsize that results (0.4 in $\cos\theta$) is justified in terms of the B-frame smearing of $\cos\theta$ (Figure 3.2); keeping the binsize this coarse keeps the amount of bin migration to a minimum.

As in the momentum distribution measurement, we measure the $\psi(1S)$ or $\psi(2S)$ yield in each partition with a binned maximum likelihood fit. The signal is modeled by lineshapes from tagged signal Monte Carlo. The background is modeled with a cubic polynomial and no constraints are imposed on any of the fit variables. MIGRAD errors were used.

In fewer than than 1% of the fits, partitions with very low statistics returned negative yields. For these fits, we attempted to force the yield to be positive; however, this resulted in undependable error measurements. Sometimes the error obtained was smaller by orders of magnitude than fits with a small positive yield. In this measurement, the error measurement is vital to the procedure because an error that is falsely reported as being near zero can unduly pull the angular fits. Therefore we do not constrain the yield to be positive; when a fit returns a negative yield, we use the error obtained by the fit, and manually reset the central value of the fit to zero.

The χ^2 of each fit was saved to a histogram and the distributions are shown in Figure 5.1 and Table 5.2. The signal lineshapes are not statistically independent from the signal Monte Carlo, so the χ^2 of those fits tend to be lower than that in the data. The χ^2 distributions in the data plots show that the fits obtained are reasonable.

Figures 5.2 ($H = 0$) and 5.3 ($H = \pm 1$) are efficiency-weighted angular distribu-

tions obtained from fitting purely polarized signal Monte Carlo. The normalization of these plots is such that the sum of the columns is equal to the detection efficiency. Each point is the $\psi(1S)$ signal yield for that bin, divided by the number of events generated in *all 5* angular bins. The plots illustrate a combination of effects: the difference between the two $\psi(1S)$ polarization states, detector efficiency and resolution, and smearing in the measurement of θ (Figure 3.1). These figures show that the difference between $H = 0$ and $H = \pm 1$ events is unmistakable in our detector despite the inherently smeared θ measurement. Similar plots are also made for the other ψ search modes but are not shown here.

Uncorrected yields of the data are shown for each search mode in Figures 5.4, 5.5, 5.6, and 5.7.

5.1.4 Continuum Correction [$\psi(1S)$ only]

We correct for continuum production of $\psi(1S)$ only in the partitions that include all p_ψ, due to low statistics in the continuum $\psi(1S)$ mode. Any plots involving the "all-momentum" partitions, such as the bottom row of Figure 5.4, will use solid lines to denote the continuum-subtracted distribution and dashed lines to denote the distribution before continuum subtraction. Figure 5.4 shows that the contribution of the continuum is at the sub-percent level. The fact that we do not execute the continuum subtraction in the lower-statistics partitions leads to a negligible error in the measurement of α in these partitions.

5.1.5 Illustration: $\alpha_{\psi(1S)}$ for Inclusive $B \to \psi(1S)X$

Although we are interested in the polarization of ψ produced directly from B decays, we take a brief detour to derive the polarization of *inclusively* produced $\psi(1S)$, over all $\psi(1S)$ momenta and leptonic decay modes.

All of the steps are depicted and explained in Figure 5.8. We can extract a couple of cross-checks from the results. The first fit (Figure 5.8, lower left) yields $N_0 = 12939 \pm 499$ and $N_\pm = 12457 \pm 516$. Using Equations (1.48) and (1.52) and ignoring errors, we obtain $N = 25396$ and $\alpha = -0.350$. Since the second fit (Figure

5.8, lower right) simply incorporates these change-of-basis equations into the fit, it is no surprise that the results from the second fit match exactly: $N = 25396 \pm 285$ and $\alpha = -0.350 \pm 0.034$.

The data angular distribution (upper left, Figure 5.8) is of continuum-corrected data over all ψ momenta and both lepton channels. So, the result for N should match the efficiency-corrected, continuum-corrected yield obtained previously in the inclusive branching fraction measurement. From Table 4.2, $N_{BF} = 25682 \pm 289$. The two results agree to 0.7σ.

5.1.6 Correcting for Feed-Down, $\psi(1S)$

In order to obtain the polarization of $\psi(1S)$ which are produced *directly* from B decays, we must first subtract the angular distributions of the $\psi(1S)$ from $\psi(2S)$ and χ_{c1} from the distribution of the inclusively produced $\psi(1S)$. In the polarization analysis, however, this correction is more complicated than that of the momentum distribution measurement. In the momentum distribution measurement, we obtain efficiency-corrected yields for all search modes, i.e. inclusive $\psi(1S)$ as well as feed-down $\psi(1S)$. In the polarization measurement, we do not want to correct for efficiency; we need to know the observed (smeared) distribution of directly produced $\psi(1S)$. Therefore, when correcting for the feed-down $\psi(1S)$, we need to adjust the feed-down distributions *partially* for efficiency; namely, we correct for the efficiency loss incurred due to adding extra particles to the $\psi(1S)$ to form the intermediate parents $\psi(2S)$ or χ_{c1}. The strategy to correct for the other complication, namely, that we must take into account fraction of the feed-down that we actually measure, remains the same as in the momentum measurement.

All of the above concerns are addressed in the following:

$$Y_{direct} = Y_{inclusive} - \epsilon^{\psi(1S)}_{\psi(2S)} \beta^{\psi(1S)}_{\psi(2S)} Y_{\psi(2S)} - \epsilon^{\psi(1S)}_{\chi_{c1}} \beta^{\psi(1S)}_{\chi_{c1}} Y_{\chi_{c1}} \tag{5.1}$$

Here, Y is the (uncorrected) observed yield in each bin of the data. $Y_{inclusive}$ is the yield of $\psi(1S)$ from any source; $Y_{\psi(2S)}$, the yield of $\psi(1S)$ which, when combined with $\pi^+\pi^-$, form a $\psi(2S)$; and $Y_{\chi_{c1}}$, which is the yield of $\psi(1S)$ which, when combined

with a γ, form a χ_{c1}.

The "marginal" efficiency $\epsilon_{\psi(2S)}^{\psi(1S)}$ is the efficiency of finding a $\psi(1S)$, divided by the efficiency of finding that same $\psi(1S)$ and combining it with a π^+ and π^- to form a $\psi(2S)$. Likewise, $\epsilon_{\chi_{c1}}^{\psi(1S)}$ is the ratio of the $\psi(1S)$ efficiency to the $[\psi(1S)+\gamma]$ efficiency.

In calculating ϵ we must account for the fact that the ψ detection efficiency may vary as a function of $\cos\theta$. By combining $H = \pm 1$ and $H = 0$ Monte Carlo in a 2:1 ratio, we create populations of $\psi(1S)$ with zero net polarization, i.e., generated flat in $\cos\theta$. By propagating these events through the Monte Carlo simulation, fitting yields in bins of $\cos\theta$ to form angular distributions, and normalizing by the total number of events generated, we derive angular efficiency plots. These efficiency plots are shown for the three search modes in Figures 5.9 [direct $\psi(1S)$], 5.10 [$\psi(1S)$ from $\psi(2S)$], and 5.11 [$\psi(1S)$ from χ_{c1}]. The efficiency loss due to detecting the extra $\pi^+\pi^-$ or γ is clear from comparing the vertical scales of the plots; subtleties of the detector response are apparent as well. Bin-by-bin division of the efficiencies depicted in Figure 5.10 by the corresponding efficiencies of Figure 5.9 yields $\epsilon_{\psi(2S)}^{\psi(1S)}$; $\epsilon_{\chi_{c1}}^{\psi(1S)}$ is obtained similarly.

Finally, we must correct for the fact that we only reconstruct one decay mode of $\psi(2S) \to \psi(1S)X$ (as discussed in Section 4.1.5); the fraction which we do measure is the factor $\beta_{\psi(2S)}^{\psi(1S)}$ in Equation (5.1). We assume that the angular distribution for $\psi(1S)$ generated from $\psi(2S) \to \psi(1S)X$ is identical to that of $\psi(2S) \to \psi(1S)\pi^+\pi^-$. This assumption is made in the above equation by the $\beta_{\psi(2S)}^{\psi(1S)}$ term, which is equal to $\mathcal{B}(\psi(2S) \to \psi(1S)X)/\mathcal{B}(\psi(2S) \to \psi(1S)\pi^+\pi^-)$. Using PDG 2001 values, $\beta_{\psi(2S)}^{\psi(1S)} = 1.752 \pm 0.119$. The error in $\beta_{\psi(2S)}^{\psi(1S)}$ will be handled later in the systematic error study. As for $\beta_{\chi_{c1}}^{\psi(1S)}$, PDG lists only one decay mode for $\chi_{c1} \to \psi(1S)X$, and that is $\chi_{c1} \to \psi(1S)\gamma$, which we measure. We therefore take $\beta_{\chi_{c1}}^{\psi(1S)} = 1$ and assign no systematic uncertainty to this.

Having obtained all of the required ingredients in Eqn. 5.1, we may now carry out the feed-down subtraction. Figure 5.12 shows "partially corrected" distributions, where the inclusive $\psi(1S)$ is corrected only for the $\psi(2S)$ feed-down mode, i.e. $Y_{inclusive} - \epsilon_{\psi(2S)}^{\psi(1S)}\beta_{\psi(2S)}^{\psi(1S)}Y_{\psi(2S)}$. Figure 5.13 shows the results after correcting for both feed-down modes (direct $\psi(1S)$.)

The statistical errors are added in quadrature in the feed-down subtraction. This

is equivalent to making the assumption that the three corrected distributions are completely uncorrelated to each other. This assumption is conservative because the statistical error in Y_{direct} would decrease if we took the correlations into account.

5.1.7 Fitting for α

We now obtain α by fitting the feed-down corrected distributions, using the same procedure described in Section 5.1.5. Results for α (statistical error only, at this point) are shown in Table 5.7; and will be discussed after the systematic error study is described in the next section. Table 5.1 shows the confidence levels of the fits for α. We conclude that the angular fits obtained are reasonable.

$\psi(1S)$	$0.0-0.8$ GeV	$0.8-1.4$ GeV	$1.4-2.0$ GeV	$0.0-2.0$ GeV
e	83.7%	98.3%	9.4%	67.0%
μ	97.1%	27.7%	94.0%	81.2%
e or μ	92.9%	42.8%	33.4%	99.0%
$\psi(2S)$	$0.0-0.7$ GeV	$0.7-1.1$ GeV	$1.1-1.6$ GeV	$0.0-1.6$ GeV
e	84.4%	71.0%	56.7%	77.5%
μ	63.8%	46.1%	12.1%	23.9%
e or μ	92.5%	69.9%	8.6%	34.2%

Table 5.1: Confidence Levels for Fits of $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$.

5.2 Systematic Error Study

5.2.1 Overview

The knob-turning study for the polarization measurement is similar to that of the momentum distribution measurement: we redo the procedures that lead to α many times, each time activating a different "knob" designed to test the importance of a single systematic error source. The last two columns of Table 4.3 list the knobs which applied to the polarization analysis.

Detailed results, showing the deviation in α attained from applying each knob, are listed for $\alpha_{\psi(1S)}$ in Table 5.5 and for $\alpha_{\psi(2S)}$ in Table 5.6.

5.2.2 Uncertainty in Unmeasured Modes

Knobs a1, a3, a4: Polarization and Branching Fraction of Unmeasured Modes ($\psi(1S)$ only) As previously discussed in Section 4.2.2, there are several decay chains of the form $B \to \psi(2S)X \to \psi(1S)X$ which we do not attempt to measure. Define $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^+\pi^-}$ to be the average polarization of $\psi(1S)$ which occurs from the intermediate decay $\psi(2S) \to J/\psi\pi^+\pi^-$. Similarly, let $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^0\pi^0}$ be the average polarization of $\psi(1S)$ which occurs from the intermediate decay $\psi(2S) \to J/\psi\pi^0\pi^0$. Finally, let $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)X_{misc}}$ be the average polarization of $\psi(1S)$ which occurs from any one of the following intermediate decays: $\psi(2S) \to \chi_{c2}\gamma \to \psi(1S)\gamma\gamma$, $\psi(2S) \to \psi(1S)\pi^0$, $\psi(2S) \to \chi_{c0}\gamma \to \psi(1S)\gamma\gamma$, or $\psi(2S) \to \psi(1S)\pi^+\pi^-$.

Our standard procedure is to measure $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^+\pi^-}$ and then assume that

$$\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^0\pi^0} = \alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)X_{misc}} = \alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^+\pi^-} \tag{5.2}$$

The first assumption is that $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^0\pi^0} = \alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^+\pi^-}$, i.e. the polarization of $\psi(1S)$ created via $\psi(2S) \to \psi(1S)\pi^0\pi^0$ is identical to the polarization of $\psi(1S)$ from $\psi(2S) \to \psi(1S)\pi^+\pi^-$. Both $\psi(2S)$ and $\psi(1S)$ are isospin zero; therefore the two-pion state must also be created in an $(I, I_3) = (0, 0)$ state. The particular manifestation of this $(0, 0)$ isospin state – be it $\pi^+\pi^-$ or $\pi^0\pi^0$ – will have no effect on the $\psi(1S)$.

Even after making this assumption, we also need to know how often $\psi(2S) \to \psi(1S)\pi^0\pi^0$ occurs. The relevant result from PDG 2001 is $\frac{\mathcal{B}(\psi(2S)\to\psi(1S)X)}{\mathcal{B}(\psi(2S)\to\psi(1S)\pi^+\pi^-)} = 1.752 \pm 0.119$. The standard procedure is to use the central value of the PDG result. Knob a1 uses 1.8716 (central value $+1\sigma$) instead.

Now we deal with the second assumption, $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)X_{misc}} = \alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)\pi^+\pi^-}$. In this case, there is absolutely no fundamental reason that this equality should hold. We estimate the degree to which this assumption affects the final result by recalculating α twice; once assuming $\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)X_{misc}} = +1$ and once again assuming

$\alpha_{\psi(1S)}^{\psi(2S)\to\psi(1S)X_{misc}} = -1$. These are knobs a3 and a4.

5.2.3 Invariant Mass Fit Procedure

Knobs b1-b4: Fitting Procedure As in the momentum distribution measurement, we re-fit the invariant mass plots using alternative parametrizations of background (knobs b1, b2) and signal lineshape (knobs b3, b4). The mean χ^2 of these fits are listed in Tables 5.2 and 5.3.

Knob:	Standard			b1 (Float Quadratic)			b2 (Fixed Cubic)		
	d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$	
search mode		Data	MC		Data	MC		Data	MC
$\psi(1S)$ Inclusive	75	75.8	74.1	76	78.2	75.3	78	99.8	78.9
$\psi(1S)$ from $\psi(2S)$	55	62.4	53.1	56	63.2	54.0	58	66.1	59.3
$\psi(1S)$ from χ_{c1}	45	51.5	45.9	46	55.0	54.1	48	56.0	50.2
$\psi(2S)$ via dileptons	75	80.3	71.7	76	80.3	72.2	78	106.2	73.8

Table 5.2: Mean χ^2 for invariant mass fits using standard and alternative background parametrizations.

Knob:	b3 (Smooth Hist.)			b4 (Gaussian)		
	d.o.f.	$\overline{\chi^2}$		d.o.f.	$\overline{\chi^2}$	
search mode		Data	MC		Data	MC
$\psi(1S)$ Inclusive	76	76.5	83.6	75	76.1	161.7
$\psi(1S)$ from $\psi(2S)$	56	61.0	59.0	55	61.8	111.2
$\psi(1S)$ from χ_{c1}	46	50.7	46.5	45	50.7	97.1
$\psi(2S)$ via dileptons	76	79.6	71.6	75	80.5	115.7

Table 5.3: Mean χ^2 for invariant mass fits using alternative signal parametrizations.

5.2.4 Monte Carlo Accuracy, Luminosity-Related Uncertainty

Knobs c, d, e These are the same as in the momentum distribution measurement.

5.2.5 Procedure for Feed-Down Correction and α Extraction

Knobs f1-f4: Handling Feed-Down; Determining α We attempt four alternative procedures in handling feed-down and determining α.

For knobs f1 and f2, instead of obtaining the final variables α and N directly from the fitter, we obtain N_0, $N_\pm$, and the 2×2 correlation matrix from the fitter, then derive α and N. Knobs f1 and f2 differ in how they extract α and N from these variables.

For Knob f1, we obtain α analytically from the $(N_\pm, N_0)$ basis. The central value of α is given by Equation 1.52. The correlation matrix returned by the fitter has the form

$$C \equiv \begin{pmatrix} C_{\pm\pm} & C_{\pm 0} \\ C_{0\pm} & C_{00} \end{pmatrix} \tag{5.3}$$

where $C_{\pm 0} = C_{0\pm}$, $C_{\pm\pm} = \Delta N_\pm^2$, and $C_{00} = \Delta N_{00}^2$. The analytical solution for the error in α is then

$$\Delta\alpha = \frac{4}{(N_\pm + 2N_0)^2}\sqrt{N_0^2 \Delta N_\pm^2 + N_\pm \Delta N_0^2 - 2N_0 N_\pm C_{0\pm}} \tag{5.4}$$

Knob f2 makes use of a Monte Carlo method to obtain α. We use random number routines in `CERNLIB` to generate 100,000 pairs of numbers $(n_\pm, n_0)$, such that the distributions of each of these variables is Gaussian, with widths $(\Delta N_\pm, \Delta N_0)$ and central values $(N_\pm, N_0)$. The generated numbers are also correlated, consistent with the off-diagonal correlation matrix element $C_{0\pm}$. For each pair $(n_\pm, n_0)$, we calculate the corresponding value of α and save it to a histogram. The resulting distribution of α is then fit to a bifurcated Gaussian, where the Gaussian widths and central value are allowed to float. The ability to fit to a bifurcated Gaussian and thus obtain asymmetric errors in α is an advantage of this method.

Knobs f3 and f4, used only for the $\alpha_{\psi(1S)}$ measurement, differ from the standard procedure at an early stage. In the standard procedure, we correct the angular distributions for feed-down, then derive α from fitting the angular distribution of direct $\psi(1S)$. Knobs f3 and f4 reverse the order: for each distribution of the directly observed search modes [inclusive $\psi(1S)$, $\psi(1S)$ from $\psi(2S)$, and $\psi(1S)$ from χ_{c1}], we

perform the angular fit, thus obtaining

$$(N_{\pm}^{incl}, N_0^{incl}), (N_{\pm}^{\psi(2S)}, N_0^{\psi(2S)}), (N_{\pm}^{\chi_{c1}}, N_0^{\chi_{c1}}) \tag{5.5}$$

and the correlation matrices

$$\begin{pmatrix} C_{\pm\pm}^{incl} & C_{\pm 0}^{incl} \\ C_{0\pm}^{incl} & C_{00}^{incl} \end{pmatrix} \begin{pmatrix} C_{\pm\pm}^{\psi(2S)} & C_{\pm 0}^{\psi(2S)} \\ C_{0\pm}^{\psi(2S)} & C_{00}^{\psi(2S)} \end{pmatrix} \begin{pmatrix} C_{\pm\pm}^{\chi_{c1}} & C_{\pm 0}^{\chi_{c1}} \\ C_{0\pm}^{\chi_{c1}} & C_{00}^{\chi_{c1}} \end{pmatrix}. \tag{5.6}$$

When fitting, we impose a constraint that all yields N be non-negative.

At this point, we again have two methods of converting these results to α. Knob f3 is the analytic calculation (similar to Knob f1), where the central value of α is

$$\alpha = \frac{N_{\pm} - 2N_0}{N_{\pm} + 2N_0} = \frac{\left(N_{\pm}^{incl} - N_{\pm}^{\psi(2S)} - N_{\pm}^{\chi_{c1}}\right) - 2\left(N_0^{incl} - N_0^{\psi(2S)} - N_0^{\chi_{c1}}\right)}{\left(N_{\pm}^{incl} - N_{\pm}^{\psi(2S)} - N_{\pm}^{\chi_{c1}}\right) + 2\left(N_0^{incl} - N_0^{\psi(2S)} - N_0^{\chi_{c1}}\right)} \tag{5.7}$$

Finding an analytic solution to the error on α requires that we have the 6×6 correlation matrix. However, the information we are given by the fitter is three 2×2 correlation matrices, so we construct a 6×6 correlation matrix as follows:

$$C \equiv \begin{pmatrix} C_{\pm\pm}^{incl} & C_{\pm 0}^{incl} & 0 & 0 & 0 & 0 \\ C_{0\pm}^{incl} & C_{00}^{incl} & 0 & 0 & 0 & 0 \\ 0 & 0 & C_{\pm\pm}^{\psi(2S)} & C_{\pm 0}^{\psi(2S)} & 0 & 0 \\ 0 & 0 & C_{0\pm}^{\psi(2S)} & C_{00}^{\psi(2S)} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{\pm\pm}^{\chi_{c1}} & C_{\pm 0}^{\chi_{c1}} \\ 0 & 0 & 0 & 0 & C_{0\pm}^{\chi_{c1}} & C_{00}^{\chi_{c1}} \end{pmatrix} \tag{5.8}$$

By setting all of the off-block-diagonal elements to zero, we are again treating the results from the three fits as being uncorrelated.

Knob f4 uses the same Monte Carlo technique as Knob f2. A computer program throws 100,000 experimental trials of three pairs of numbers, distributed with the central values, errors, and (pairwise) correlation coefficients that are obtained from the three fits. Every time a set of random numbers is generated, α is calculated by Equation 5.7; the distribution in α is plotted and fit with a bifurcated Gaussian to obtain α and its error. Because we throw pairs of numbers that satisfy the 2×2 correlation matrices, each pair of random numbers generated has no correlation with the other pairs, i.e. once again we treat the results from the three fits as being uncorrelated.

5.2.6 Results

Activating each of the 20 knobs for $\psi(1S)$, and 13 knobs for $\psi(2S)$, we obtain varying results for α. We combine the deviations from the standard procedure for each knob, Δ_{knob}, to obtain a final systematic error, Δ_{Syst}. As in the momentum spectrum measurement, pairs of knobs which probe the same systematic uncertainty are averaged before adding in quadrature. The systematic error is obtained from the individual knob deviations as follows:

$$\begin{aligned}
\Delta^2_{Syst} &= \Delta^2_{a1} + [(|\Delta_{a3}| + |\Delta_{a4}|)/2]^2 \\
&+[(|\Delta_{b1}| + |\Delta_{b2}|)/2]^2 + [(|\Delta_{b3}| + |\Delta_{b4}|)/2]^2 \\
&+\Delta^2_{c1} + \Delta^2_{c2} + \Delta^2_{c3} + \Delta^2_{c4} + \Delta^2_{c5} + \Delta^2_{c6} + \Delta^2_{c7} \\
&+\Delta^2_{d3} + \Delta^2_{e1} \\
&+[(|\Delta_{f1}| + |\Delta_{f2}| + |\Delta_{f3}| + |\Delta_{f4}|)/4]^2 \qquad (5.9)
\end{aligned}$$

Table 5.7 shows the final polarization results, with statistical and systematic errors.

Table 5.4 is a systematic error summary for our two central results, the all-momentum, both-lepton partitions of $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$. The table shows "subtotal" systematic errors from groupings of similar knobs; each knob is weighted according to Equation 5.9. In both the $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$ measurements, the largest systematic error comes from uncertainty in the Monte Carlo modelling [the (c) knobs], particularly in lepton ID. Almost as significant is the variation due to invariant mass fitting procedures [the (b) knobs], within which the dominant error comes from using a fixed background shape. The total systematic error is less than 4% in both $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$, which is small compared to either of the statistical errors.

The deviation from the central value obtained from each knob, as well as the systematic error, for all 12 partitions in momentum and lepton decay mode, are shown for $\alpha_{\psi(1S)}$ in Table 5.5 and for $\alpha_{\psi(2S)}$ in Table 5.6.

5.2.7 Low-Statistics $\psi(2S)$ Partitions

In the $\psi(2S)$ polarization measurement – particularly for the partitions with low statistics – we sometimes observe 1) nonphysical central values of α and 2) unstable results when applying knobs b2 or f2.

Unphysical Results for α: The central values determined for α for high-momentum $\psi(2S) \to ee$ and $\psi(2S) \to \mu\mu$ lie outside the physically valid region of $[-1, +1]$ (although with very large errors). To see how this is possible, recall that α is the value for which $1 + \alpha \cos^2 \theta$ is the shape that best fits the $\cos\theta$ distribution. If the $\cos\theta$ distribution is extremely concave or extremely convex (for example, due to statistical fluctuation), the fitter returns a value of α which, although mathematically possible, is physically not allowed.

Examining the angular distributions (third row, left and center, Figure 5.7), we see that both of them include a bin with negligible yield. Even completely polarized samples do not include bins with zero yield (see Figures 5.2 and 5.3). Therefore the mere existence of a zero-yield bin means that the distribution cannot have been created by an *additive* linear superposition of pure-helicity events.

In these atypical cases, artificially forcing the fit to return $|\alpha| > 1$ complicates the calculation and interpretation of the resulting error. Therefore we report the unconstrained fit result.

Unstable results from knob b2: Applying knob b2 on the $\psi(2S)$ data results in 3 partitions reporting polarization results that differ by more than 1.0 relative to the standard procedure, as seen in Table 5.6. Knob b2 applies a fixed background shape to the invariant mass fits. Figure 5.14 shows the middle-momentum $\psi(2S) \to ee$ partition. The top row shows the invariant mass fit of data events in the fifth angular bin, i.e. $+0.6 < \cos\theta < +1.0$. The standard, floating-cubic background fit is on the left, and the fixed background (knob b2) is on the right. With knob b2 activated, the fit returns a negative value, which is modified to be zero. The effect of this change is evident in the middle row, which plots the angular distributions resulting from the 5 mass fits. The dramatic drop in yield in the crucial fifth angular bin pulls

the fit to be more convex than the completely longitudinal MC, thus pulling α to an unphysical value, -1.33. The statistics are small enough that a relatively minor change in background shape changes the yield drastically enough to make a large change in α.

Unstable results from knob f2: This method fits the direct-ψ angular distribution for N_0 and $N_\pm$. The fitter results are then converted into the (N, α) basis by throwing random numbers for N_0 and $N_\pm$, converting them to α, and making a histogram of the many random trials. The bottom row of Figure 5.14 shows example histograms. The left histogram is of the "grand" partition of $\psi(2S)$. When the errors on N_0 and $N_\pm$ are small, as they are in this case, then the errors on α are well-behaved. In the high-momentum muon partition, however, the errors on N_0 and N_1 are large to the point that the α that is calculated for most of the thrown values is out of the allowed region. Because of this the errors on α are highly non-Gaussian. Fitting a bifurcated Gaussian is woefully inadequate in this case (lower right, Figure 5.14) and the knob deviation reflects this.

Conclusion: The cases of nonphysical central value and knob instability simply reflect the fact that there are insufficient statistics to make a robust measurement of $\alpha_{\psi(2S)}$ when dividing the dataset into momentum bins; therefore the results for these bins, particularly the high-momentum bins, should be viewed with caution. However, our primary result is for the combined-momentum partitions, for which none of the points raised in this subsection apply.

knob(s)	source of uncertainty	Systematic Error $\alpha_{\psi(1S)}$	$\alpha_{\psi(2S)}$
a1	$\mathcal{B}(\psi(2S) \to \psi(1S)\pi^+\pi^-)$	0.007	0.000
a3-a4	$\alpha(X_{misc})$	0.010	0.000
	subtotal (a)	0.012	0.000
b1-b2	background parametrization	0.020	0.020
b3-b4	signal lineshape	0.011	0.003
	subtotal (b)	0.023	0.021
c1	tracking	0.000	0.015
c2-c3	crystal calorimeter	0.012	0.008
c4-c5	muon ID	0.010	0.021
c6-c7	electron ID	0.018	0.005
	subtotal (c)	0.024	0.027
d3	$M(\pi^+\pi^-)$ cut efficiency	0.005	0.000
e1	continuum scale factor	0.000	0.000
f1-f4	feed-down, α procedure	0.009	0.016
	TOTAL	0.037	0.038

Table 5.4: Systematic error summary for the polarization measurements which include all ψ momenta and both leptonic decay modes.

$p_{\psi(1S)}$	low $(0.0-0.8$ GeV)			mid $(0.8-1.4$ GeV)		
$\psi(1S) \rightarrow$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$
α	0.91	-0.13	0.32	-0.24	-0.45	-0.37
a1	0.13	0.00	0.05	0.01	0.01	0.01
a3	-0.17	-0.01	-0.07	-0.03	-0.02	-0.02
a4	-0.07	0.03	-0.01	-0.01	0.00	0.00
b1	0.09	0.00	-0.02	0.01	-0.07	0.00
b2	-0.03	-0.02	-0.14	-0.21	0.07	-0.01
b3	-0.02	0.03	-0.03	-0.01	-0.01	0.01
b4	0.02	-0.02	0.02	-0.03	-0.01	-0.01
c1	0.02	0.02	0.00	0.00	-0.04	0.01
c2	0.01	0.00	-0.01	-0.01	-0.04	0.00
c3	0.01	0.01	-0.01	0.00	-0.01	0.01
c4	0.00	-0.07	-0.01	0.00	0.00	0.03
c5	0.00	-0.03	-0.01	0.00	0.00	0.02
c6	-0.13	0.00	0.08	0.01	0.00	0.01
c7	0.01	0.00	-0.01	-0.01	0.00	0.01
d3	0.10	0.00	0.03	0.01	0.00	0.01
e1	0.00	0.00	0.00	0.00	0.00	0.00
f1	0.09	0.00	0.00	0.00	0.00	0.00
f2	-0.16	-0.14	-0.14	-0.03	-0.03	-0.02
f3	-0.04	0.13	-0.02	0.02	0.01	0.01
f4	-0.17	0.06	-0.14	0.01	0.00	-0.01
$\Delta\alpha_{Syst}$	0.27	0.12	0.15	0.12	0.09	0.04
$p_{\psi(1S)}$	high $(1.4-2.0$ GeV)			all $(0.0-2.0$ GeV)		
$\psi(1S) \rightarrow$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$
α	-0.65	-0.40	-0.52	-0.23	-0.37	-0.30
a1	0.00	0.00	0.00	0.01	0.00	0.01
a3	-0.01	-0.01	-0.01	-0.02	-0.01	-0.02
a4	0.00	-0.01	0.00	-0.01	0.00	0.00
b1	0.01	-0.01	0.01	0.02	-0.02	0.00
b2	-0.09	0.09	0.01	-0.13	0.04	-0.04
b3	0.01	0.00	-0.01	-0.04	-0.01	0.00
b4	0.00	0.00	-0.01	-0.03	0.00	-0.02
c1	-0.01	0.01	0.00	0.00	0.00	0.00
c2	0.02	-0.01	0.00	0.00	-0.02	-0.01
c3	0.01	0.00	0.00	0.00	-0.01	-0.01
c4	0.00	0.00	0.00	0.00	0.01	0.01
c5	0.00	-0.01	0.00	0.00	0.00	0.00
c6	0.01	0.00	0.01	0.04	0.00	0.02
c7	0.02	0.00	0.01	0.01	0.00	0.00
d3	0.00	0.00	0.00	0.01	0.00	0.01
e1	0.00	0.00	0.00	0.00	0.00	0.00
f1	0.00	0.00	0.00	0.00	0.00	0.00
f2	-0.02	-0.02	-0.01	-0.02	-0.02	-0.01
f3	-0.01	-0.10	-0.04	-0.02	0.00	-0.01
f4	-0.02	-0.11	-0.05	-0.03	-0.01	-0.02
$\Delta\alpha_{Syst}$	0.06	0.08	0.03	0.10	0.04	0.04

Table 5.5: $\alpha_{\psi(1S)}$ systematic error results.

$p_{\psi(2S)}$	low $(0.0-0.8$ GeV$)$			mid $(0.8-1.4$ GeV$)$		
$\psi(2S) \rightarrow$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$
α	-0.57	-0.54	-0.57	-0.56	-0.91	-0.83
b1	0.01	-0.01	0.00	0.00	0.00	0.00
b2	0.07	0.15	0.11	-0.77	0.01	-0.30
b3	0.01	-0.03	0.01	0.03	-0.03	0.01
b4	0.06	-0.01	0.00	-0.02	0.04	0.01
c1	0.05	-0.05	-0.01	-0.04	0.01	-0.01
c2	-0.05	0.02	-0.01	-0.08	0.01	-0.02
c3	-0.05	0.02	-0.01	-0.08	0.01	-0.02
c4	0.00	-0.15	-0.09	0.00	0.05	0.03
c5	0.00	0.13	0.07	0.00	0.03	0.02
c6	0.03	0.00	0.00	0.05	0.00	0.00
c7	0.04	0.00	0.01	-0.02	0.00	-0.01
f1	0.00	0.00	0.00	0.00	0.00	0.00
f2	-0.30	-0.25	-0.15	-0.24	-0.14	-0.09
$\Delta\alpha_{Syst}$	0.14	0.23	0.13	0.41	0.08	0.16

$p_{\psi(2S)}$	high $(1.4-2.0$ GeV$)$			all $(0.0-2.0$ GeV$)$		
$\psi(2S) \rightarrow$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$	e^+e^-	$\mu^+\mu^-$	$\ell^+\ell^-$
α	-1.24	5.28	0.19	-0.81	0.03	-0.45
b1	0.00	-0.12	0.01	0.00	-0.01	0.00
b2	0.04	9.09	1.39	-0.27	0.33	0.04
b3	0.03	1.00	0.08	0.01	-0.05	0.00
b4	0.00	-0.36	-0.03	0.01	0.01	0.00
c1	-0.03	0.41	0.00	-0.03	0.00	-0.01
c2	0.02	-0.16	0.04	-0.03	0.01	-0.01
c3	0.02	-0.16	0.04	-0.03	0.01	-0.01
c4	0.00	-0.10	0.01	0.00	-0.04	-0.02
c5	0.00	-0.72	-0.07	0.00	0.04	0.01
c6	-0.06	0.00	-0.04	-0.03	0.00	0.00
c7	-0.02	0.00	0.00	-0.01	0.00	0.00
f1	0.00	0.01	0.00	0.00	0.00	0.00
f2	-0.08	-4.07	-0.38	-0.08	-0.19	-0.06
$\Delta\alpha_{Syst}$	0.08	4.84	0.72	0.15	0.19	0.04

Table 5.6: $\alpha_{\psi(2S)}$ systematic error results.

$\alpha_{\psi(1S)}$	$0.0-0.8$ GeV	$0.8-1.4$ GeV	$1.4-2.0$ GeV	$0.0-2.0$ GeV
e	$+0.91^{+0.82}_{-0.55} \pm 0.27$	$-0.24^{+0.15}_{-0.14} \pm 0.12$	$-0.65^{+0.10}_{-0.09} \pm 0.06$	$-0.23^{+0.11}_{-0.10} \pm 0.10$
μ	$-0.13^{+0.38}_{-0.30} \pm 0.12$	$-0.45^{+0.14}_{-0.13} \pm 0.09$	$-0.40^{+0.12}_{-0.11} \pm 0.08$	$-0.37^{+0.10}_{-0.09} \pm 0.04$
ℓ	$+0.32^{+0.37}_{-0.30} \pm 0.15$	$-0.37^{+0.10}_{-0.09} \pm 0.04$	$-0.52^{+0.08}_{-0.07} \pm 0.03$	$-0.30^{+0.07}_{-0.07} \pm 0.04$
$\alpha_{\psi(2S)}$	$0.0-0.7$ GeV	$0.7-1.1$ GeV	$1.1-1.6$ GeV	$0.0-1.6$ GeV
e	$-0.57^{+0.65}_{-0.43} \pm 0.14$	$-0.56^{+0.52}_{-0.35} \pm 0.41$	$-1.24^{+0.29}_{-0.24} \pm 0.08$	$-0.81^{+0.24}_{-0.20} \pm 0.15$
μ	$-0.54^{+0.55}_{-0.40} \pm 0.23$	$-0.91^{+0.38}_{-0.30} \pm 0.08$	$+5.28^{+22.35}_{-3.26} \pm 4.84$	$+0.03^{+0.46}_{-0.35} \pm 0.19$
ℓ	$-0.57^{+0.37}_{-0.29} \pm 0.13$	$-0.83^{+0.27}_{-0.22} \pm 0.16$	$+0.19^{+0.75}_{-0.50} \pm 0.72$	$-0.45^{+0.22}_{-0.19} \pm 0.04$

Table 5.7: Final results, polarization measurement. The first error is statistical error, the second error is systematic. The summed partitions (all partitions involving ℓ and/or $0.0-2.0$ GeV) are obtained from independent fits, not averages of the unsummed partitions.

5.3 Results and Discussion

Our primary results for $B \to \psi(direct)X$ polarization, which include all ψ momentum and both leptonic decay modes, are shown below, where the first error is statistical and the second is systematic. These are the first measurements of the polarization of $\psi(1S)$ or $\psi(2S)$ produced directly from B decays. Results for $\alpha_{\psi(1S)}$ and $\alpha_{\psi(2S)}$ as a function of p_ψ are shown in Table 5.7.

$$\alpha_{\psi(1S)} = -0.30 \pm 0.07 \pm 0.04 \tag{5.10}$$

$$\alpha_{\psi(2S)} = -0.45^{+0.22}_{-0.19} \pm 0.04 \tag{5.11}$$

5.3.1 Cross-check: Lepton Universality

The polarization parameter should be independent of the flavor of the daughter leptons. The data are consistent with this, as for each momentum bin, the results for electron and muon bins are all within 2σ of each other. The low momentum bin for $\psi(1S)$ may be cause for some concern, with central values near 1 for electrons and near 0 for muons, although with large errors. The upper left and upper middle plots of direct-$\psi(1S)$ distributions, Figure 5.13, show that indeed, the electron distribution looks like $1+\cos^2\theta$, whereas the muon distribution is more flat. Figure 5.12 shows the "partially feed-down corrected" distributions which result when we correct for $\psi(2S)$ feed-down but not χ_{c1} feed-down. The low-momentum electron and muon bins agree better than in the fully corrected distributions. This is confirmed by Figure 5.6, which clearly shows a significant difference between the low-momentum electron and muon bins for the χ_{c1} feed-down sample. Recall that each of the points in these plots is the result of a mass fit; the mass fits that make up these two plots are displayed in Figures 5.15 and 5.16. Inspection of these plots shows that nothing obviously pathological is happening with the fits, supporting the hypothesis that this difference is a statistical fluctuation.

5.3.2 Comparing to Other Measurements

The last time CLEO reported findings on ψ polarization was conference proceedings in 1992 [78]. The result was $\frac{\Gamma_L}{\Gamma}_{\psi(1S)} = 0.59 \pm 0.15$, corresponding to $\alpha = -0.48 \pm 0.24$. Given the poor statistics available at the time it is highly unlikely that this result corrected for feed-down $\psi(1S)$.

A January 2000 thesis from KEK-BELLE using 0.42 fb^{-1} of on-resonance data [79] obtained $\alpha_{\psi(1S)} = -0.92 \pm 0.16 \pm 0.09$, for the entire momentum range of $\psi(1S)$. This analysis does not correct for feed-down $\psi(1S)$.

In March 2001, BaBar [64] presented a preliminary result of $\alpha_{\psi(1S)} = -0.424 \pm 0.023$, using about 20 fb^{-1}, about double the data used for our result. Comparing the statistical error of this result to that of our inclusive ψ measurement of $\alpha_{\psi(1S)} = -0.350 \pm 0.034$ (Section 5.1.5 and Figure 5.8) , it is fairly certain that the BaBar result includes the contributions of $\psi(1S)$ feed-down modes.

5.3.3 Conclusions

Our result for $\alpha_{\psi(1S)}$ is almost 4σ away from zero. This measurement therefore rules out the color evaporation model, which predicts that charmonia are produced with zero net polarization, regardless of the production mechanism.

Within 2σ in combined statistical and systematic error, we find that $\alpha_{\psi(1S)} \in [-0.48, -0.12]$. According to the latest NRQCD predictions for α (Section 1.6.5, Table 1.5), our result tends to favor a higher b quark mass. However, these calculations are leading-order in α_S; therefore the links between our results and the NRQCD matrix elements are currently not firmly established. In the future, with continued improvements in the theoretical calculations, we hope that these results will provide a strong constraint on the NRQCD color-octet matrix elements.

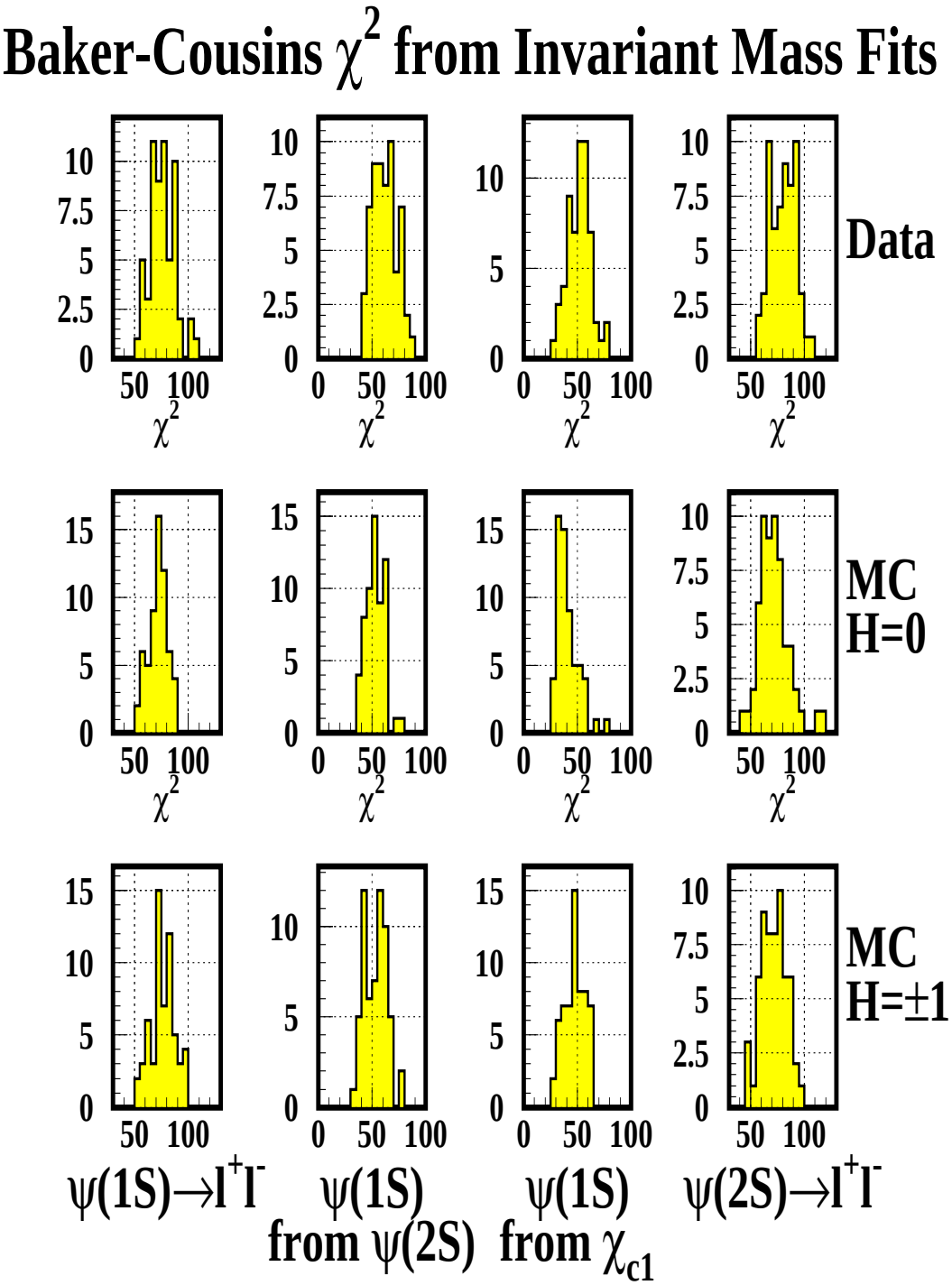

Figure 5.1: Distributions of the Likelihood function (χ^2) of the 720 fits for yield in the polarization analysis.

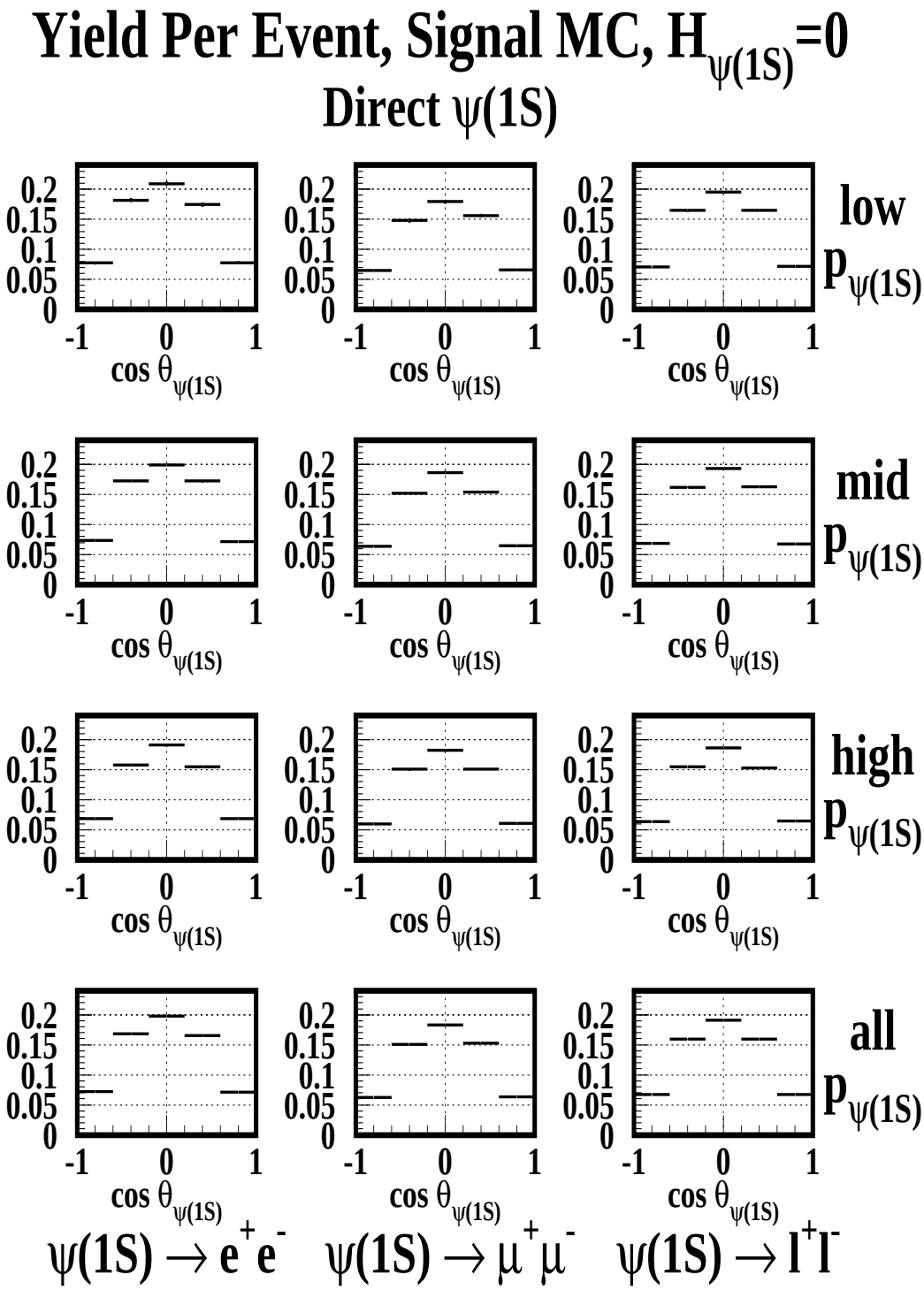

Figure 5.2: Detected Angular Yield per Monte Carlo Event, $\psi(1S)_{H=0}$ produced directly from B decays. Momentum Range Definitions: "Low" = 0.0-0.8 GeV, "Mid" = 0.8-1.4 GeV, "High" = 1.4-2.0 GeV, "All" = 0.0-2.0 GeV.

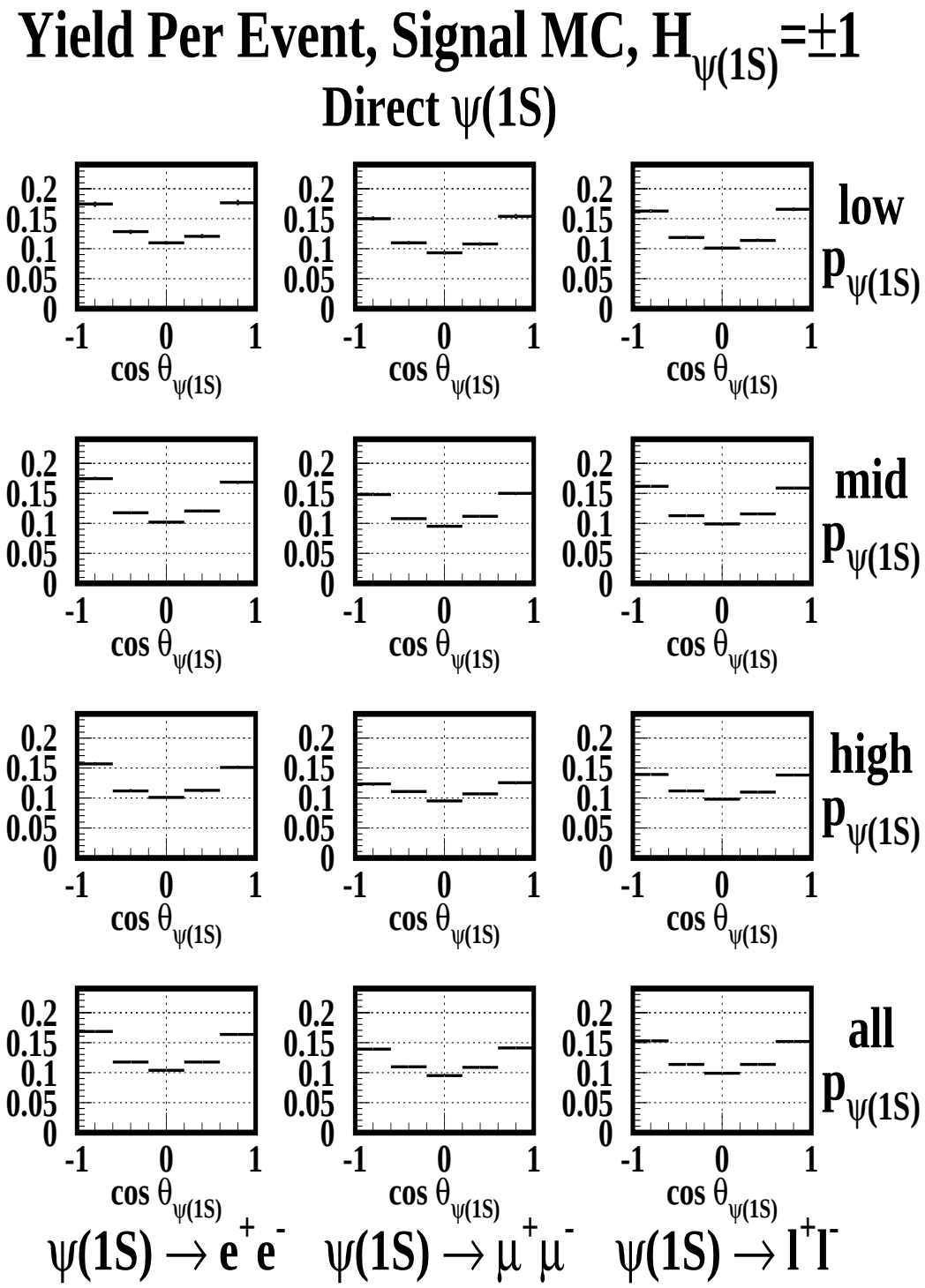

Figure 5.3: Detected Angular Yield per Monte Carlo Event, $\psi(1S)_{H=\pm1}$ produced directly from B decays. Momentum Range Definitions: "Low" = 0.0-0.8 GeV, "Mid" = 0.8-1.4 GeV, "High" = 1.4-2.0 GeV, "All" = 0.0-2.0 GeV.

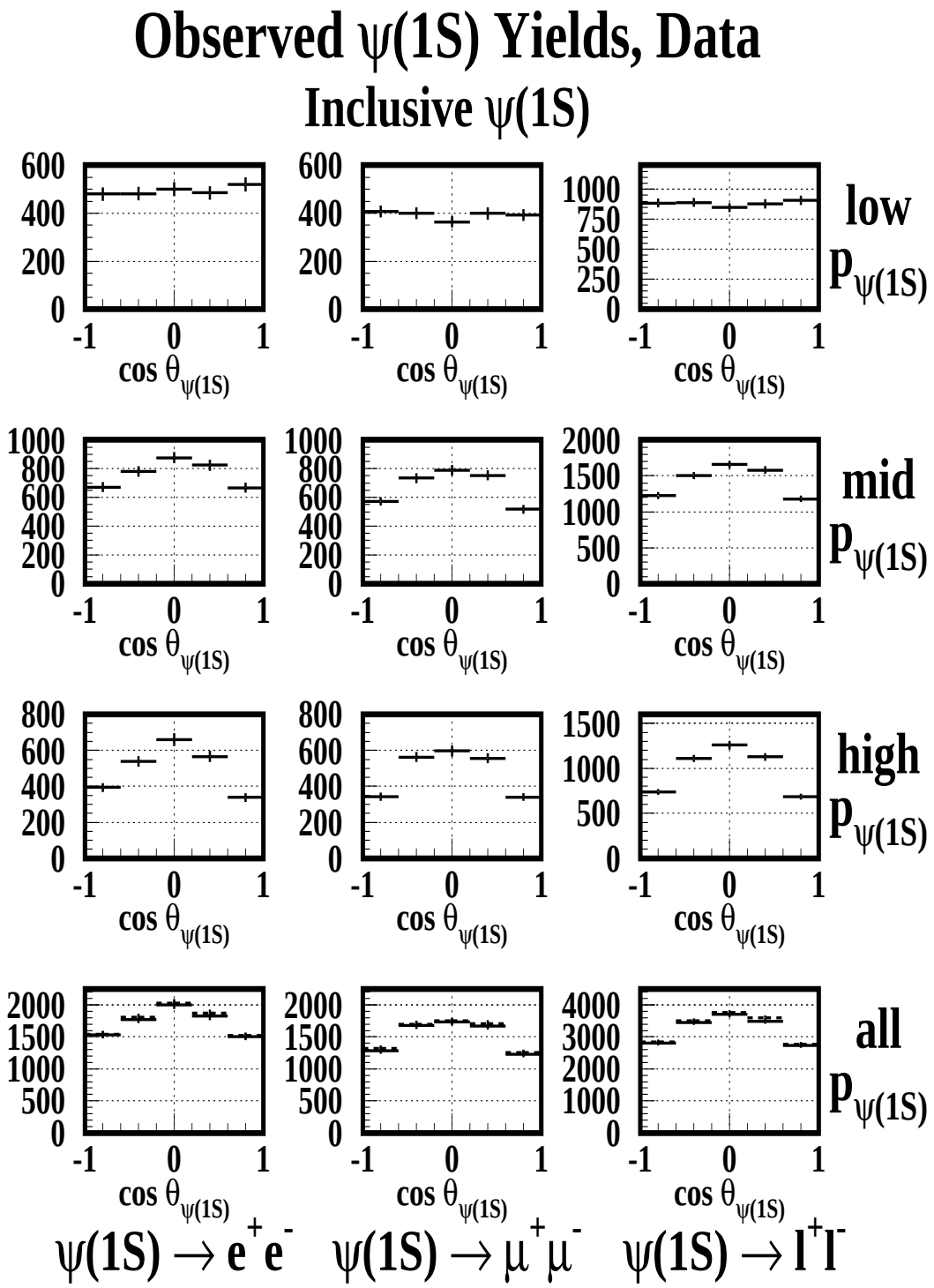

Figure 5.4: Data: inclusively produced $\psi(1S)$. Bottom row, dashed: before subtracting the distribution of continuum $\psi(1S)$.

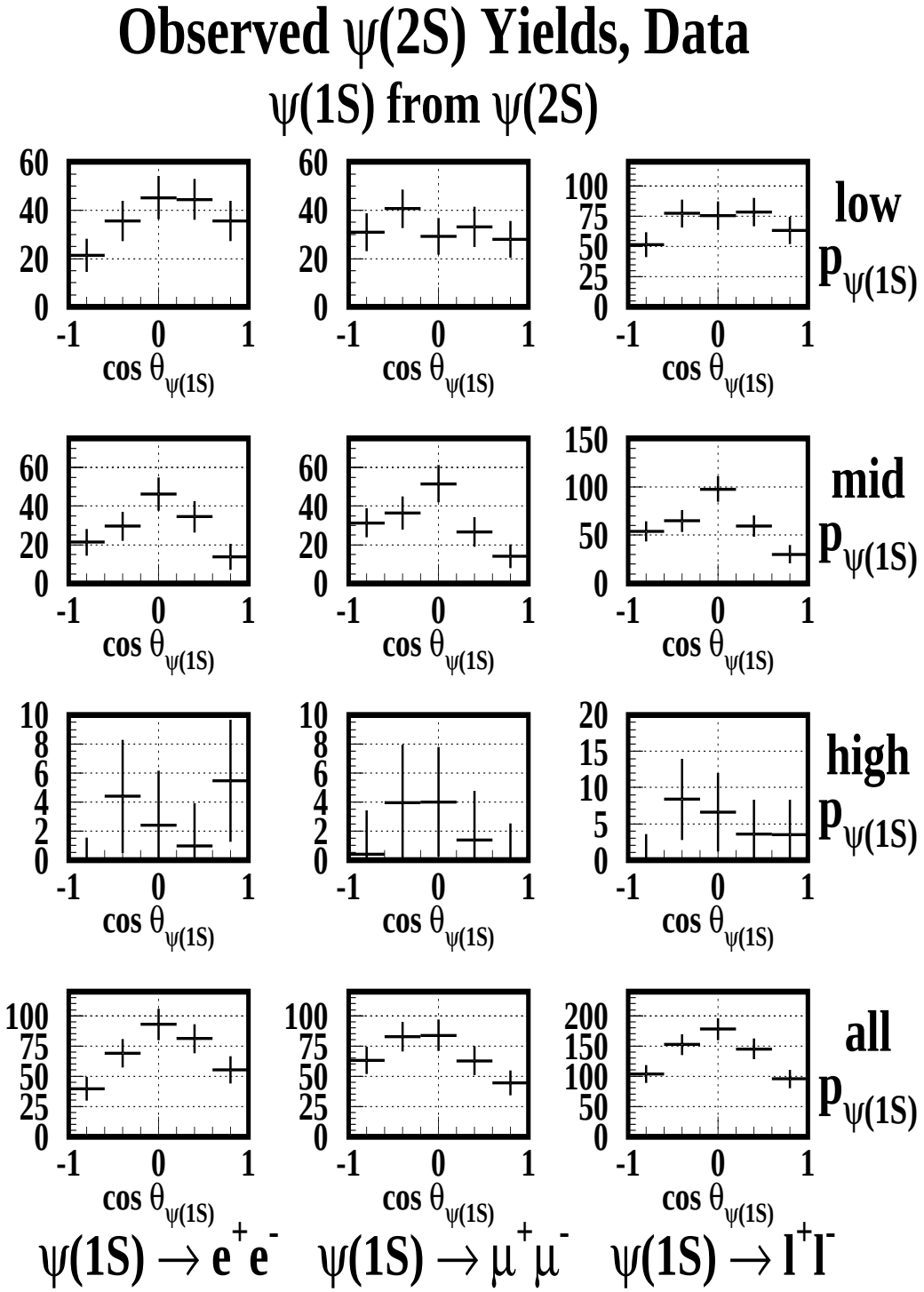

Figure 5.5: Data: $\psi(1S)$ produced via $\psi(2S) \to \psi(1S)\pi^+\pi^-$.

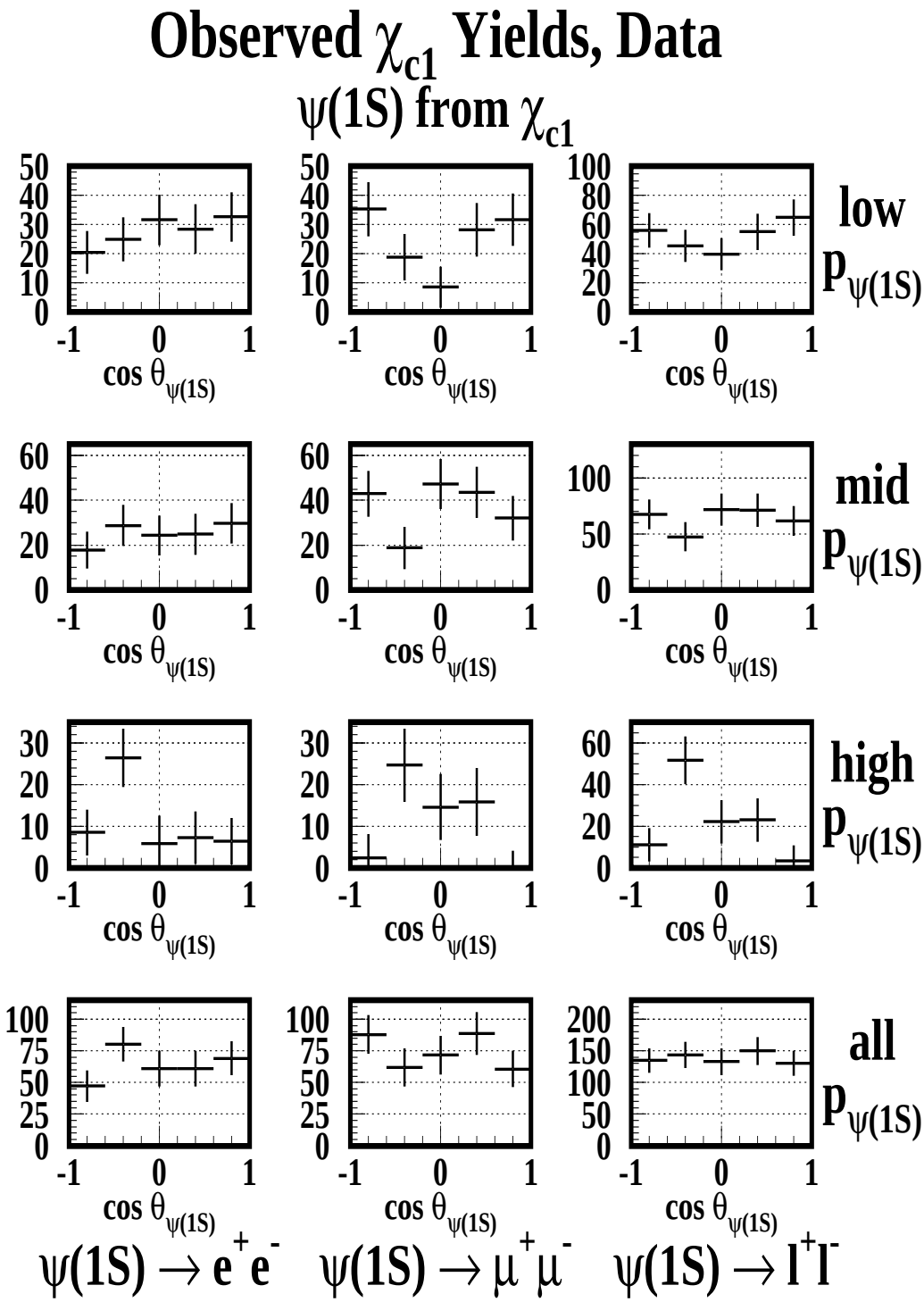

Figure 5.6: Data: $\psi(1S)$ produced via $\chi_{c1} \to \psi(1S)\gamma$.

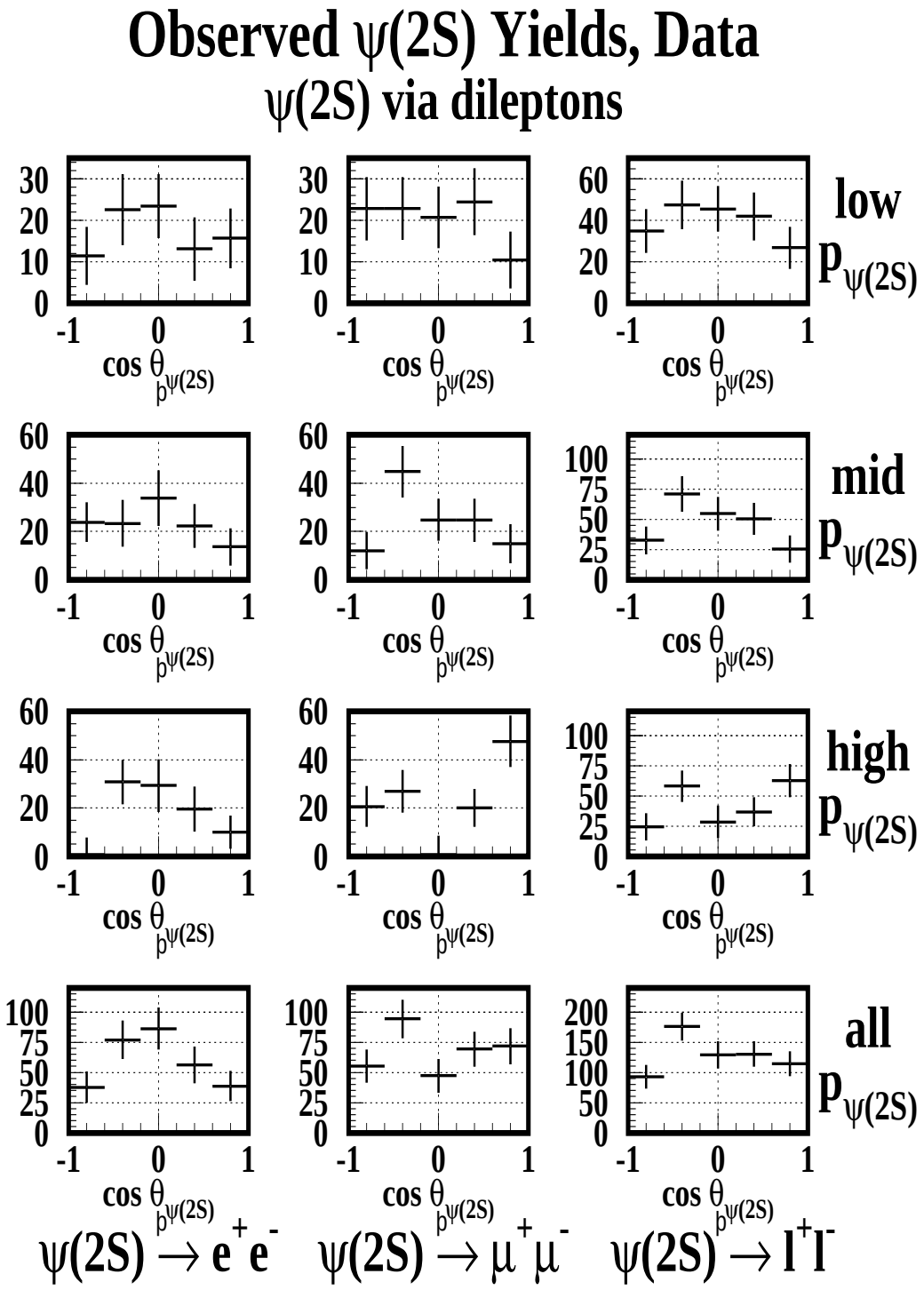

Figure 5.7: Data: $\psi(2S)$ decaying via $\psi(2S) \to \ell^+\ell^-$.

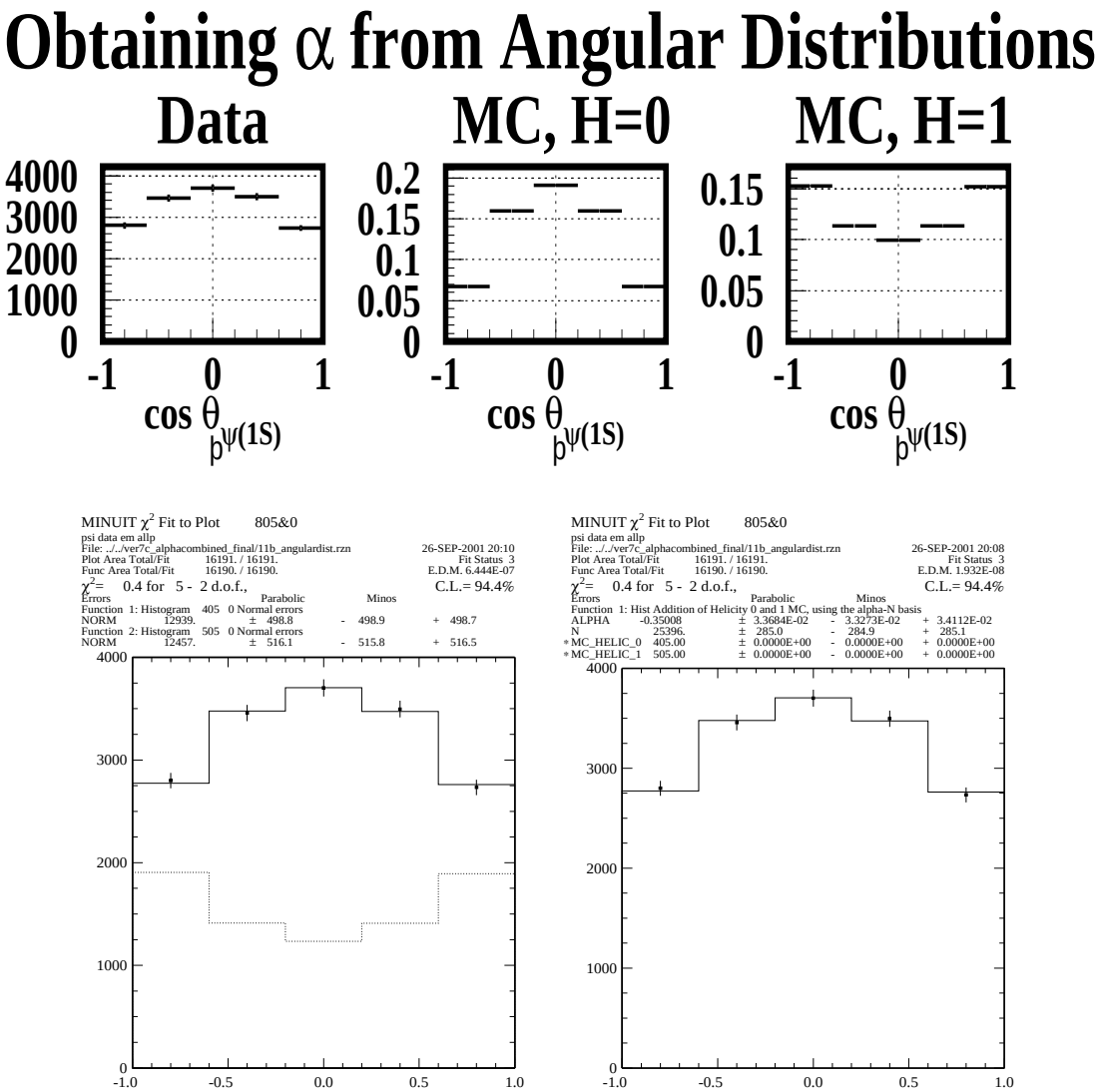

Figure 5.8: An illustrative example of how α is obtained from the angular distributions – see Section 5.1.5 for more discussion. All plots are from the all-momentum, combined e-μ partition. **Top left**: Data, continuum-corrected inclusive $\psi(1S)$, from Figure 5.4. **Top center**: MC, direct helicity 0 $\psi(1S)$, from Figure 5.2. **Top right**: MC, direct helicity ± 1 $\psi(1S)$, from Figure 5.3. The data plot is a measured, uncorrected distribution. The MC plots show the average contribution of one pure-helicity $\psi(1S)$ to the measured, uncorrected distribution. For the same reasons as in the momentum measurement, we may use direct ψ Monte Carlo to adjust inclusive ψ data, see Figure 5.8. **Lower left**: results of fitting the data to a linear combination of the two MC plots, thus obtaining $N_\pm$ and N_0, the number of transversely and longitudinally polarized $\psi(1S)$ in the data. The dashed line is the contribution of the $H = \pm 1$ component. **Lower right**: Results of fitting directly for α and N, the polarization and total number of $\psi(1S)$. The two fits are mathematically equivalent.

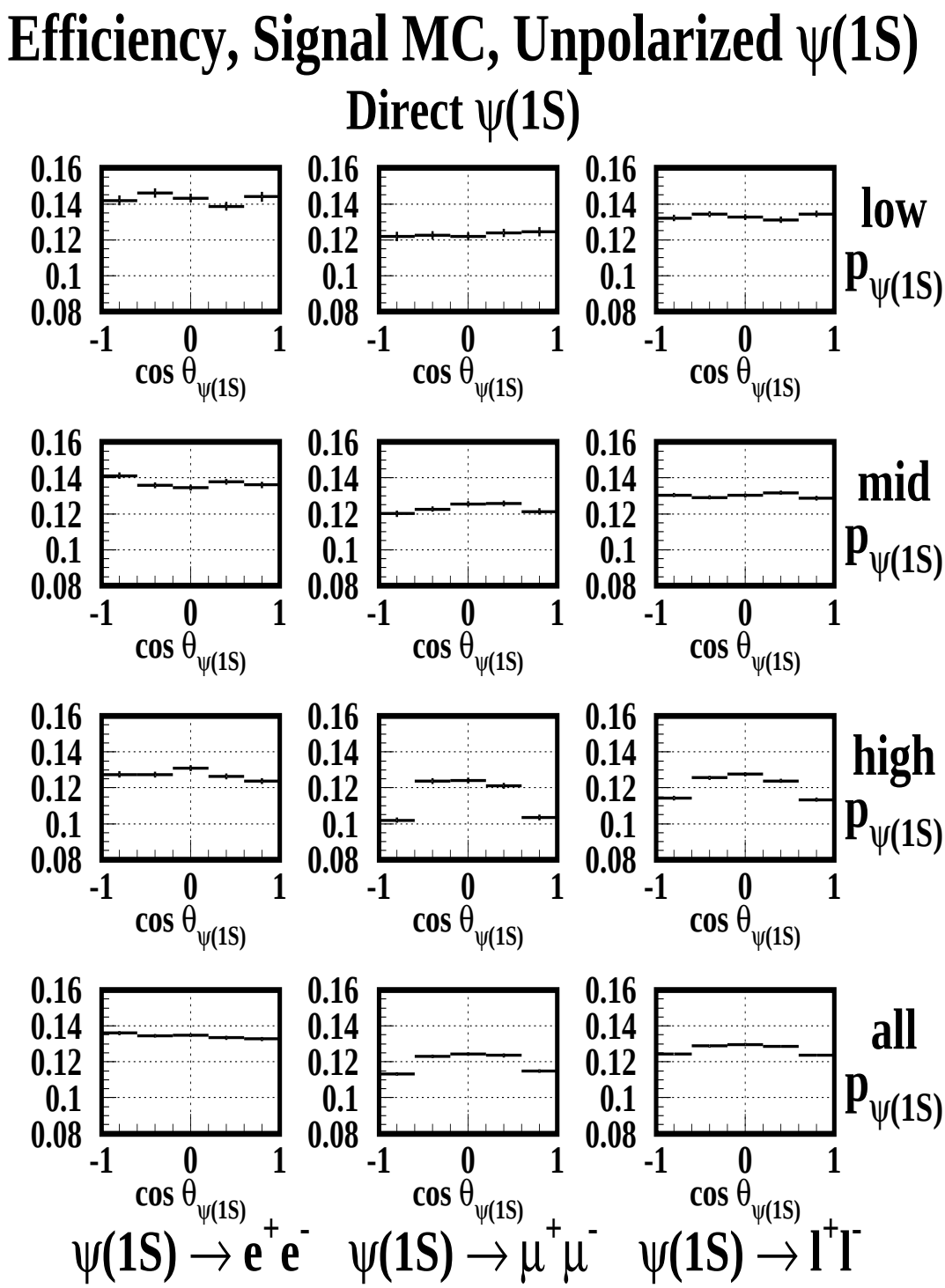

Figure 5.9: Efficiency as a function of $\cos\theta$ and p_ψ for unpolarized $\psi(1S)$ produced directly from B decays. Momentum Range Definitions: "Low" = 0.0-0.8 GeV, "Mid" = 0.8-1.4 GeV, "High" = 1.4-2.0 GeV, "All" = 0.0-2.0 GeV.

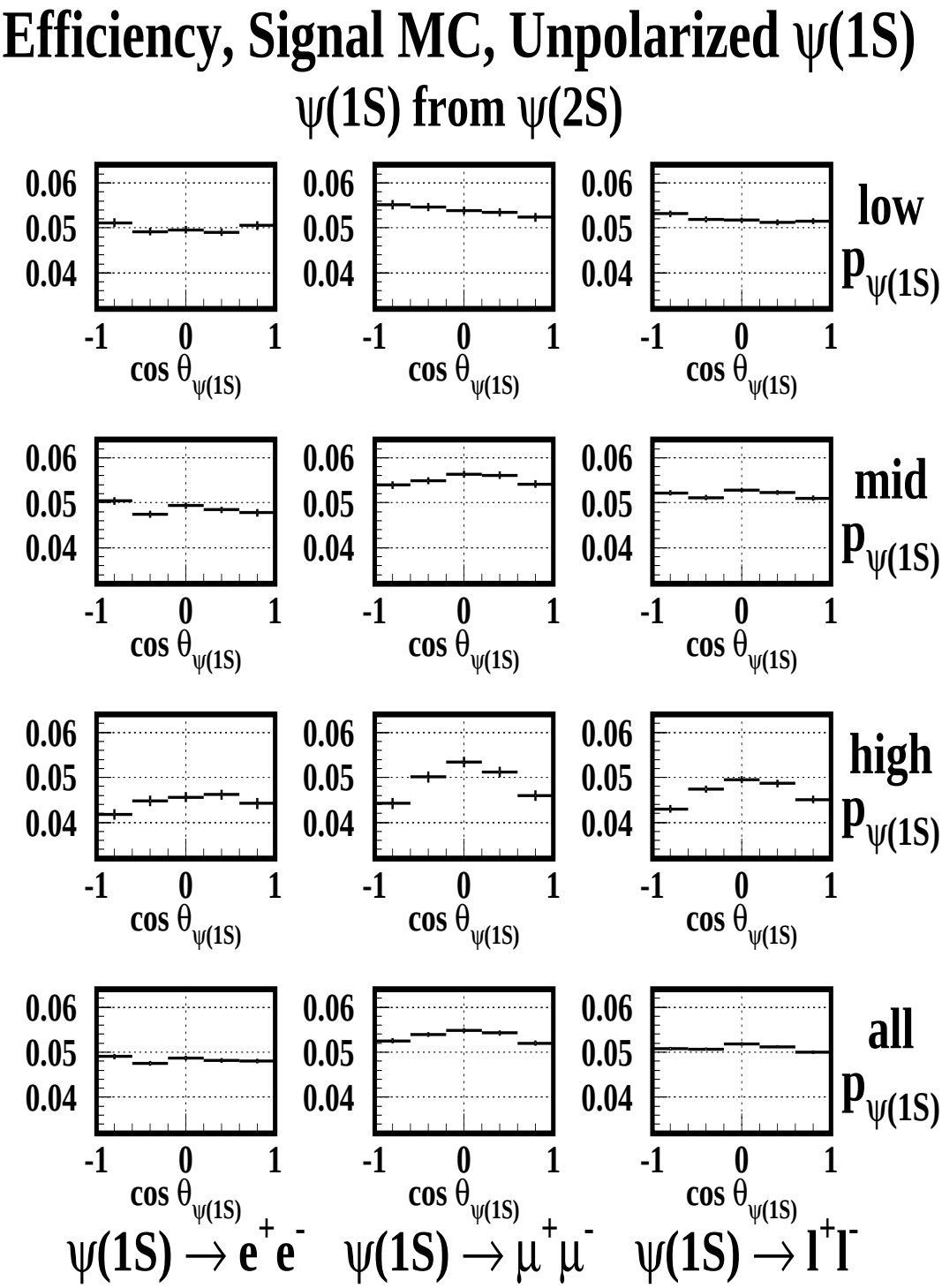

Figure 5.10: Efficiency as a function of $\cos\theta$ and p_ψ for unpolarized $\psi(1S)$ produced via $\psi(2S) \rightarrow \psi(1S)\pi^+\pi^-$. Momentum Range Definitions: "Low" = 0.0-0.8 GeV, "Mid" = 0.8-1.4 GeV, "High" = 1.4-2.0 GeV, "All" = 0.0-2.0 GeV.

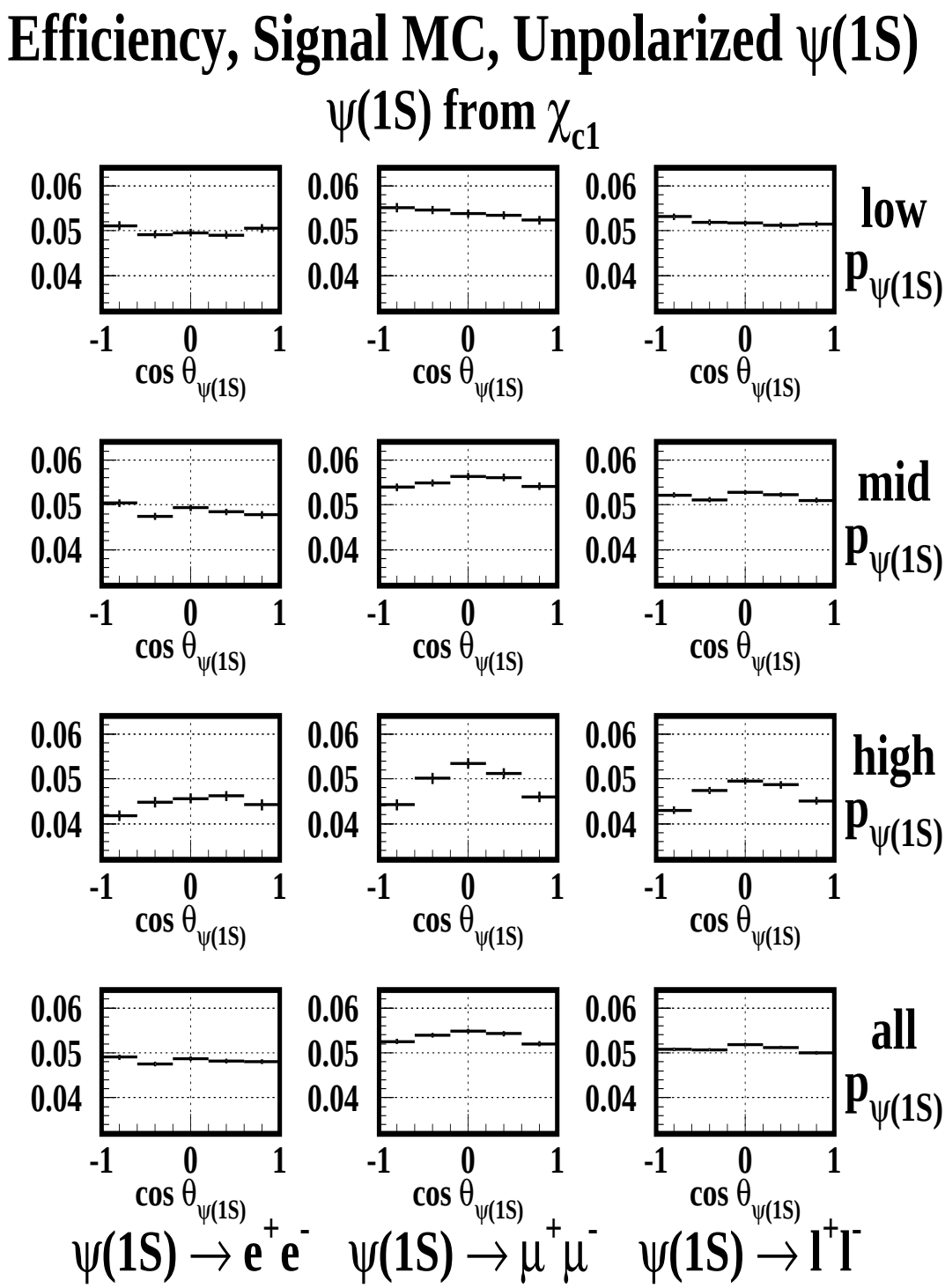

Figure 5.11: Efficiency as a function of $\cos\theta$ and p_{psi} for unpolarized $\psi(1S)$ produced via $\chi_{c1} \to \psi(1S)\gamma$. Momentum Range Definitions: "Low" = 0.0-0.8 GeV, "Mid" = 0.8-1.4 GeV, "High" = 1.4-2.0 GeV, "All" = 0.0-2.0 GeV.

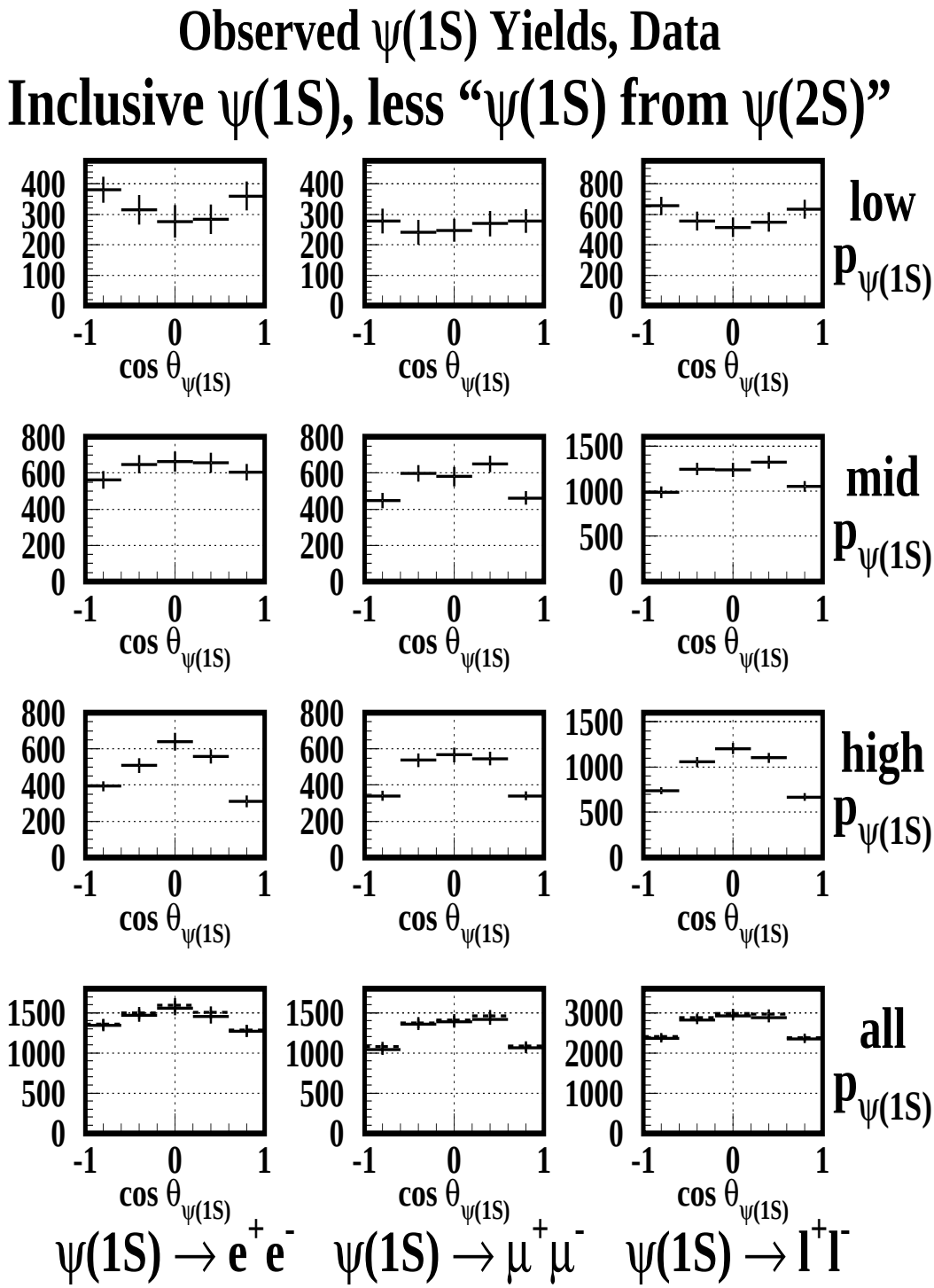

Figure 5.12: Data: partially corrected angular distributions. The distributions of $[\psi(1S)$ from $\psi(2S)]$ are subtracted from the distributions of inclusively produced $\psi(1S)$. Bottom row, dashed: before subtracting the distribution of continuum $\psi(1S)$.

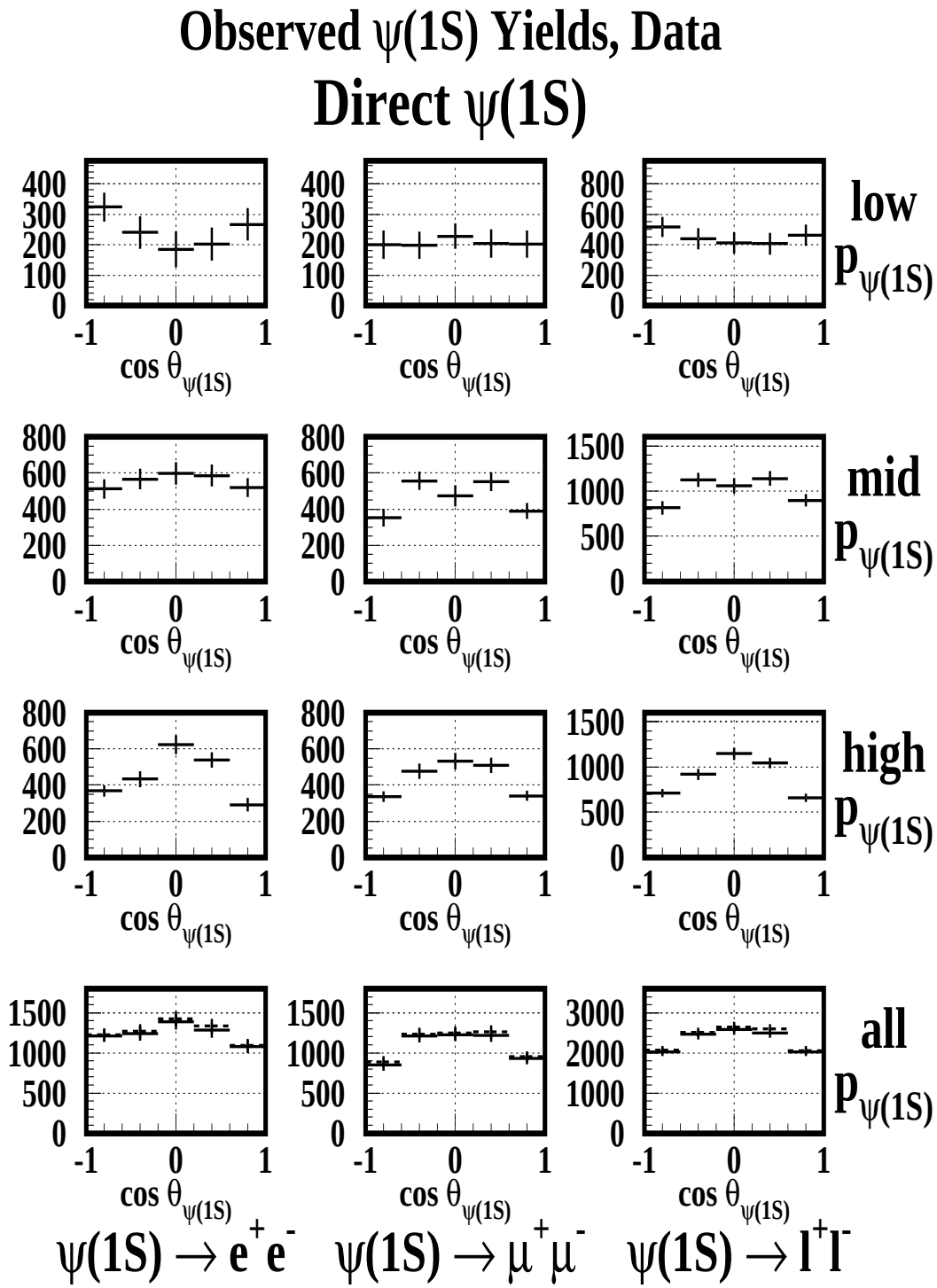

Figure 5.13: Data: Directly Produced $\psi(1S)$. The distributions of [$\psi(1S)$ from $\psi(2S)$] and [$\psi(1S)$ from χ_{c1}] are subtracted from the distributions of inclusively produced $\psi(1S)$. Bottom row, dashed: before subtracting the distribution of continuum $\psi(1S)$. The vertical scales are set identically to Figure 5.12.

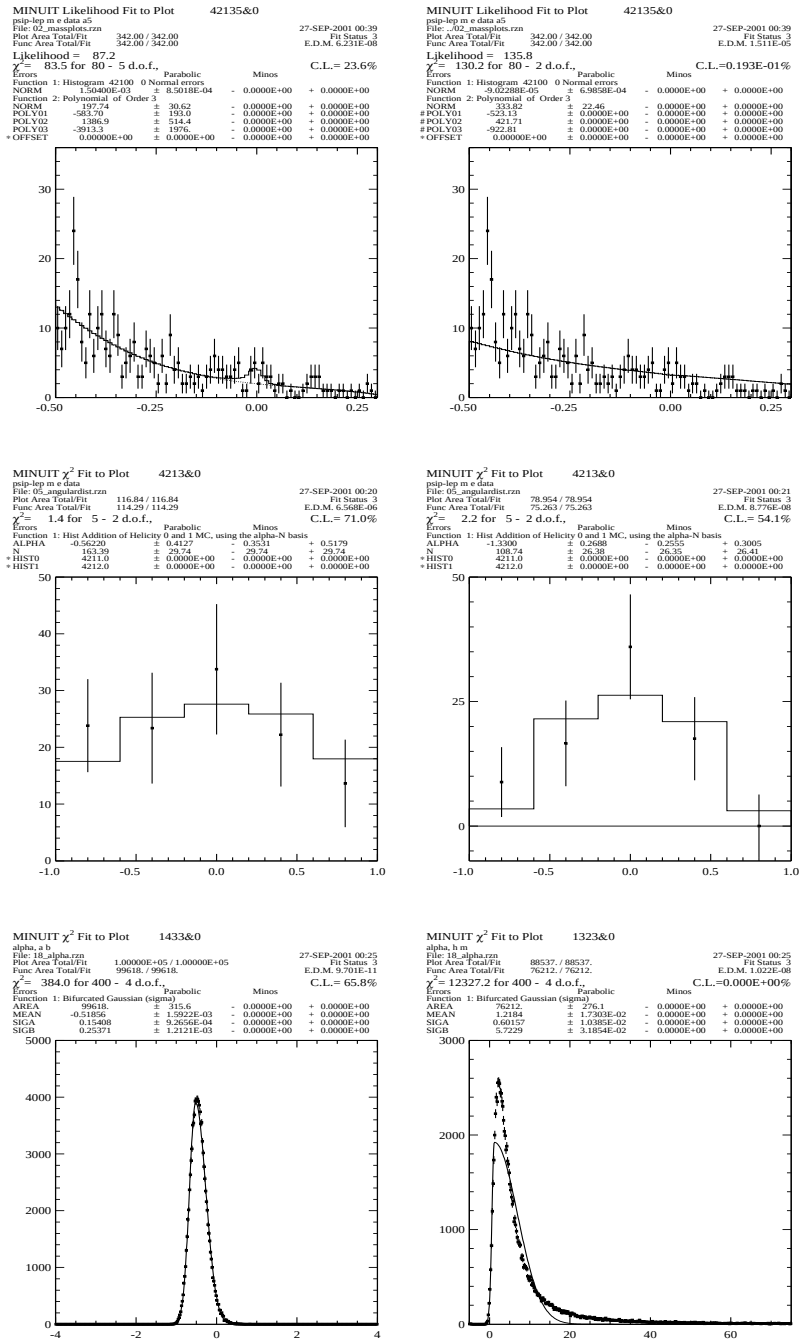

Figure 5.14: Plots illustrating large knob deviations in low-statistics $\psi(2S)$ partitions. See Sec. 5.2.7 for more details. **Top Row**: Mass fits for $\psi(2S) \to e^+e^-$, $0.6 < \cos\theta < 1.0$. Left: standard fitting procedure; Right: fixed background shape [knob (b2)]. Horizontal axis is GeV. **Middle row**: Angular distributions for $\psi(2S) \to e^+e^-$; the fifth bin is given by the yield on the fits in the top row. Left: standard fitting procedure; Right: fixed background shape [knob (b2)]. Horizontal axis is $\cos\theta$. **Bottom Row**: Fits for α using the Toy MC technique. Left: α for the unpartitioned full dataset. Right: α for the high-momentum muon partition. The central value for the highly complicated distribution that results is poorly estimated by the bifurcated Gaussian. Horizontal axis is α.

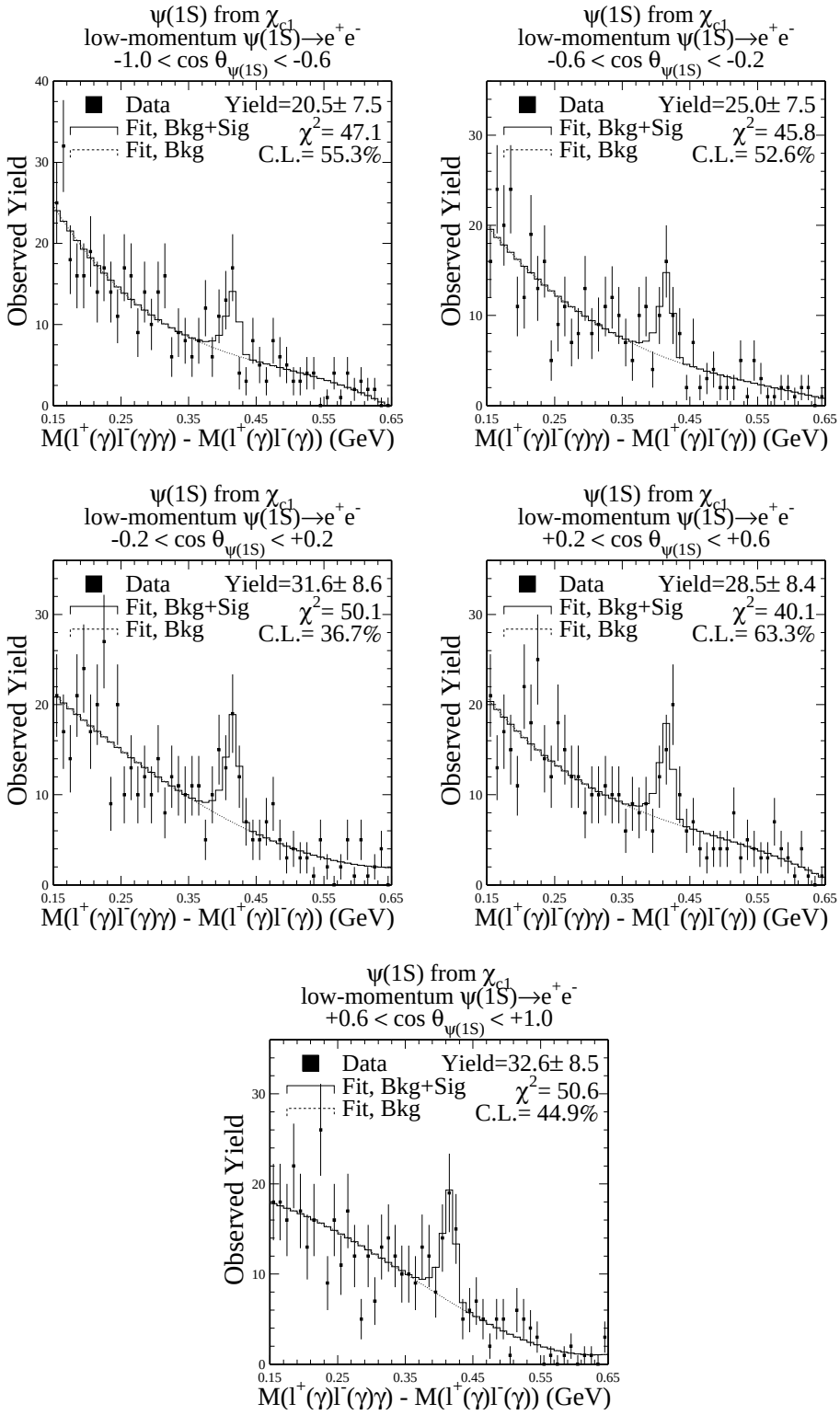

Figure 5.15: Data: Mass fits for $\psi(1S)$ from χ_{c1}, low momentum bin, $\psi(1S) \to ee$. The yields of these fits are the entries of the bins in the upper left plot of Figure 5.6.

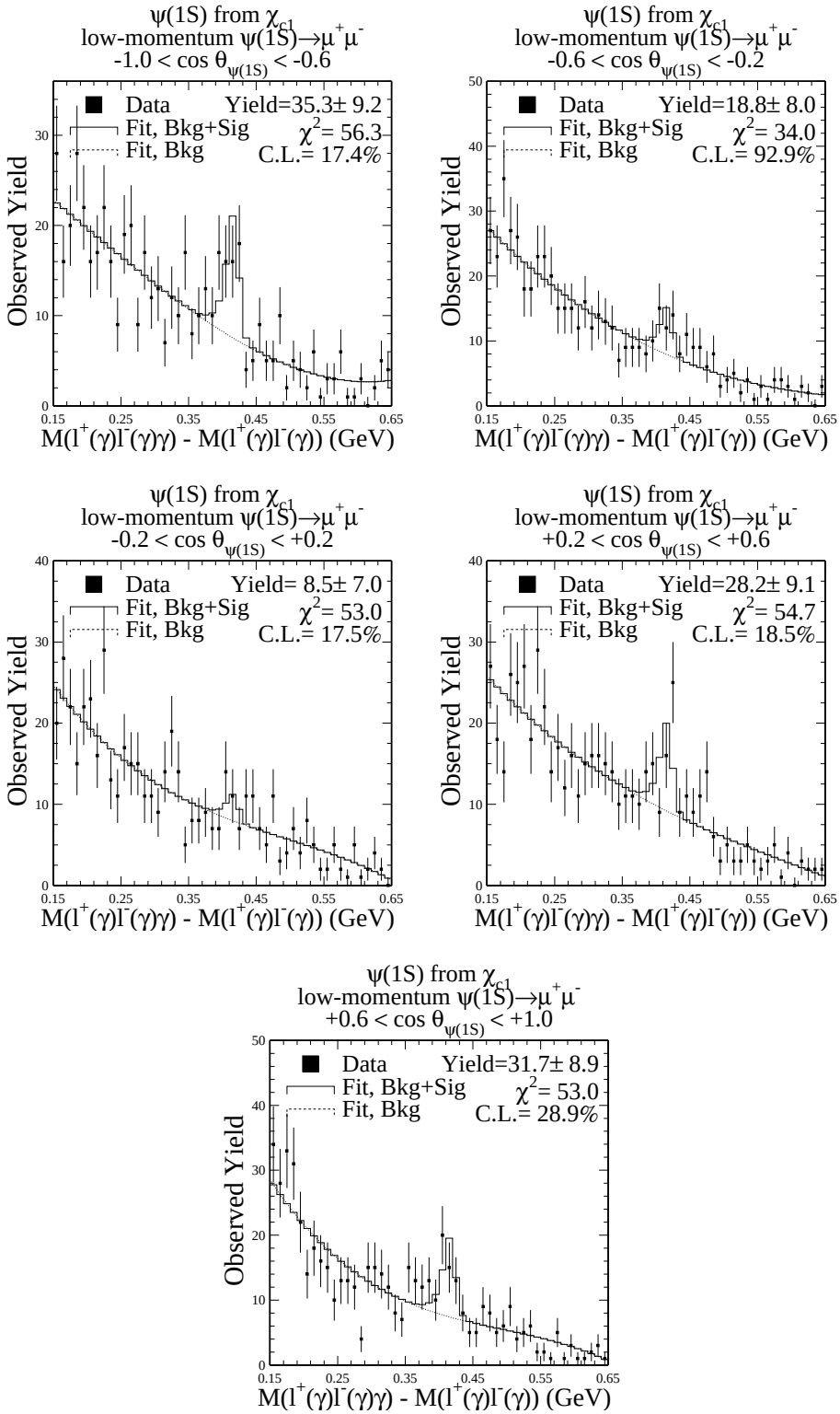

Figure 5.16: Data: Mass fits for $\psi(1S)$ from χ_{c1}, low momentum bin, $\psi(1S) \to \mu\mu$. The yields of these fits are the entries of the bins in the upper center plot of Figure 5.6.

List of References

[1] CLEO Collaboration, D. Andrews *et al.,* Phys. Rev. Lett. **44**, 1108 (1980); CLEO Collaboration, D. Andrews *et al.,* Phys. Rev. Lett. **45**, 219 (1980).

[2] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996)

[3] P. F. Harrison and H. R. Quinn (ed.), SLAC Report SLAC-R-504, Chapter 2, hep-ph/9812217.

[4] G. T. Bodwin, E. Braaten, T. C. Yuan, and G. P. Lepage, Phys. Rev. D **46**, R3703 (1992).

[5] K. G. Wilson and W. Zimmermann, Commun. Math. Phys. **24**, 87 (1972).

[6] E.C.G. Stueckelberg and A. Petermann, Helv. Phys. Acta **26**, 499 (1953); M. Gell-Mann and F.E. Low, Phys. Rev. **95**, 1300 (1954); L. V. Ovsyannikov, Dokl. Acad. Nauk SSSR **109**, 1112 (1956); C. G. Callan, Jr., Phys. Rev. D **2**, 1541 (1970); K. Symanzik, Commun. Math. Phys. **18**, 227 (1970); G. 't Hooft, Nucl. Phys. B **61**, 455 (1973); S. Weinberg, Phys. Rev. D **8**, 3497 (1973).

[7] L. B. Okun, Leptons and Quarks, paperback edition (1984).

[8] `http://www-theory.lbl.gov/~ianh/alpha/alpha.html`

[9] G. 't Hooft and M. Veltman, Nucl. Phys. B **44**, 189 (1972).

[10] D.E. Groom *et al.*, Eur. Phys. J. C **15** 1 (2000), and 2001 off-year partial update for the 2002 edition available on the PDG WWW pages (pdg.lbl.gov)

[11] R. Barbierei, M. Caffo, R. Gatto and E. Remiddi, Phys. Lett. B **95**, 93 (1980); Nucl. Phys. B **192**, 61 (1981).

[12] CDF Collaboration, F. Abe *et al.,* Phys. Rev. Lett. **69**, 3704 (1992).

[13] CDF Collaboration, F. Abe *et al.,* Phys. Rev. Lett. **79**, 572 (1997).

[14] I. Rothstein, Proceedings of Heavy Flavors 8, Southampton, UK (1999). hep-ph/9911276

[15] CDF Collaboration, T. Affolder *et al.,* Phys. Rev. Lett. **85**, 2886 (2000).

[16] H. Fritzsch, Phys. Lett. B **67**, 217 (1977).

[17] E. Braaten, S. Fleming, and T.C. Yuan, Ann. Rev. Nucl. Part. Sci. **46**, 197 (1996). hep-ph/9602374

[18] CLEO Collaboration, S. Chen *et al.,* Phys. Rev. D **63**, 31102 (2001).

[19] G.T. Bodwin, E. Braaten, and G.P. Lepage, Phys. Rev. D **51**, 1125 (1995). hep-ph/9407339

[20] E. Braaten. Talk given at 3rd International Workshop on Particle Physics Phenomenology, Taipei, Taiwan, Nov 14-17, 1996. hep-ph/9702225

[21] E. Braaten, Nucl. Phys. B (Proc. Suppl.) **71**, 441 (1999)

[22] B. Grinstein, Int. J. Mod. Phys. A **15**, 461 (2000). hep-ph/981264

[23] M. Kramer, Prog. Part. Nucl. Phys. **47**, 141 (2001). hep-ph/0106120

[24] C. Quigg and J.L. Rosner, Phys. Rep. **56**, 167 (1979).

[25] C. T. H. Davies *et al.,* Phys. Rev. D **52**, 6519 (1995). hep-lat/9506026

[26] H. Georgi, Ann. Rev. Nucl. Part. Sci. **43**, 209 (1993).

[27] G. P. Lepage, in From Actions to Answers ed. by T. DeGrand and D. Toussaint (World Scientific, Singapore, 1989)

[28] C. T. H. Davies *et al.,* Phys. Rev. D **50**, 6963 (1994).

[29] M. Beneke, F. Maltoni, and I.Z. Rothstein, Phys. Rev. D **59**, 054003 (1999). hep-ph/9808360

[30] M. Neubert, Phys. Rept. **245**, 259 (1994). hep-ph/9306320

[31] G. Altarelli, N. Cabibbo, G. Corbo, L. Maiani, and G. Martinelli, Nucl. Phys. B **208**, 365 (1982).

[32] P. Ko, J. Lee, and H.S. Song, Phys. Rev. D **53**, 1409 (1995).

[33] J. P. Ma, Phys. Lett. B **488**, 55 (2000).

[34] E. Eichten and C. Quigg, Phys. Rev. D **52**, 1726 (1995).

[35] CLEO Collaboration, R. Balest *et al.*, Phys. Rev. D **52**, 2661 (1995). CLNS 94/1315, CLEO 94-26.

[36] M. Beneke and M. Kramer, Phys. Rev. D **55**, 5269 (1997).

[37] W. Palmer, E. Paschos, and P. Soldan, Phys. Rev. D **56**, 5794 (1997). hep-ph/9701328

[38] M. Beneke, G. A. Schuler, and S. Wolf, Phys. Rev. D **62**, 034004 (2000).

[39] C.-H. V. Chang and W.-S. Hou, Phys. Rev. D **64**, 071501(R) (2001).

[40] BELLE Collaboration, talk given by S.E. Schrenk, 30th Int'l Conf. on High Energy Phys., 2000, Osaka, Japan.

[41] BaBar Collaboration, M. Milek, talk given at PHENO 2001, May 7, 2001.

[42] H. Yamamoto, private communication.

[43] W. Palmer, E. Paschos, and P. Soldan, hep-ph/9602376

[44] C. Peterson, D. Schlatter, J. Schmitt, and P.M. Zerwas, Phys. Rev. D **27**, 105 (1983);

[45] S. Fleming, O. F. Hernandez, I. Maksymyk, and H. Nadeau, Phys. Rev. D **55**, 4098 (1997). hep-ph/9608413

[46] J. P. Ma, Phys. Rev. D **62**, 054012 (2000).

[47] S.W. Herb *et al.,* Phys. Rev. Lett. **39**, 252 (1977).

[48] R. Littauer, CLEO internal document CBN 85-7.

[49] D. Cinabro, K. Korbiak, R. Ehrlich, S. Henderson, and N. Mistry, CLNS-00-1706, physics/0011075, submitted to Nucl. Instrum. Meth. A (2000).

[50] CLEO Collaboration, D. Cinabro *et al.*, Phys. Rev. E **57**, 1193 (1998).

[51] CLEO Collaboration, Y. Kuboda *et al.*, Nucl. Instrum. Meth. Phys. A **320**, 66 (1992).

[52] F. Wurthwein and J.P. Alexander, CLEO internal document CBX 93-39.

[53] P. Hopman, Ph.D. Thesis, Cornell University (2000).

[54] D. Asner, Ph.D. Thesis, University of California at Santa Barbara (2000).

[55] BaBar Collaboration, B. Aubert *et al.* Phys. Rev. Lett. **87**, 162002 (2001). hep-ex/0106044

[56] A. Ershov, Ph.D. Thesis, Harvard University (2001); CLEO internal documents CBX 00-53, CLEO 00-17, CLEO 99-31.

[57] A. Weinstein, CLEO internal document CSN 96-347.

[58] G. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[59] MARK III Collaboration, D. Coffman *et al.*, Phys. Rev. Lett. **68**, 282 (1992).

[60] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994). `www.thep.lu.se/~torbjorn/Pythia.html`

[61] Application Software Group, Computing and Networks Division, CERN, CERN Program Library Long Writeup W5013, Geneva, Switzerland (1993). `wwwinfo.cern.ch/asdoc/geant_html3/geantall.html`

[62] T. Pham, B. Pire, and T. Truong, Phys. Lett. B **61**, 183 (1976).

[63] SPEAR Collaboration, E. Hilger *et al.*, Phys. Rev. Lett. **35**, 625 (1975).

[64] BaBar Collaboration, V. Brigljevic, talk given at XXXVIth Rencontres de Moriond, session QCD and High Energy Hadronic Interactions (2001). hep-ex/0109043

[65] BaBar Collaboration, B. Aubert *et al.*, Phys. Rev. D **65**, 031101 (2002). hep-ex/0109004

[66] B. Berger, CLEO internal documents CBX 00-32 (2000) and CBX 98-36 (1998).

[67] A. Gritsan, CLEO internal document CBX 99-72 (1999).

[68] B. Heltsley, CLEO internal document CBX 95-35 (1995).

[69] B. Heltsley, CLEO internal documents CBX 95-35 (1995) and CBX 92-117 (1992); CLEO Collaboration, G. Wei *et al.*, CLEO internal document CBX 94-72 (1994).

[70] B. Gittelman *et al.*, CLEO internal document CBX 91-82 (1991).

[71] B. Gittelman, CLEO internal document CBX 99-32 (1999).

[72] D. Kim, CLEO internal document CBX 01-49 (2001).

[73] S. Lee, CLEO internal document CBX 01-57 (2001).

[74] S. Anderson, private communication.

[75] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **87**, 161601 (2001).

[76] CLEO Collaboration, S.J. Richichi *et al.*, Phys. Rev. **D63**, 031103 (2001)

[77] V. Fadeyev and R. Stroynowski, CLEO internal document CBX 00-43 (2000).

[78] K. Honscheid, talk given at Recontres de Moriond (1992).

[79] S. Ichizawa, Ph.D. Thesis, Graduate School of Science and Engineering, Tokyo Institute of Technology (2000).