Semileptonic $B$ Decays to Light Neutral Hadrons:

$B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \eta \ell \nu$

Shoshanna Beth Cole

A thesis submitted in fulfilment of the requirements for
the degree of Master of Science

Falkiner High Energy Physics
School of Physics
University of Sydney

May, 2007
Abstract

An analysis of $B \to \pi^0 \ell \nu$ and $B \to \eta \ell \nu$ decays using the neutrino reconstruction technique is presented. The dataset consists of 535 million $B \bar{B}$ pairs in 492 fb$^{-1}$ of data collected with the Belle detector at the KEKB asymmetric $e^+e^-$ collider. The measured $B \to \pi^0 \ell \nu$ and $B \to \eta \ell \nu$ branching fractions are $\mathcal{B}(B \to \pi^0 \ell \nu) = (0.68 \pm 0.09 \pm 0.11 \pm 0.04) \times 10^{-4}$ and $\mathcal{B}(B \to \eta \ell \nu) = (0.42 \pm 0.13) \times 10^{-4}$. The errors on the $\pi^0$ measurement are statistical, experimental systematic, and due to $b \to u \ell \nu$ modelling, respectively; that on the $\eta$ is statistical only. The $B \to \pi^0 \ell \nu$ branching fraction is measured in three $q^2$ bins: $q^2 < 8 \text{ GeV}^2$, $8 \text{ GeV}^2 \leq q^2 < 16 \text{ GeV}^2$, and $16 \text{ GeV}^2 \leq q^2$. The Cabibbo-Kobayashi-Maskawa quark-mixing matrix element $|V_{ub}|$ is extracted from the $B \to \pi^0 \ell \nu$ branching fraction using a Light-Cone Sum Rules form factor extrapolated to the full $q^2$ range, and is found to be $|V_{ub}| = (3.29 \pm 0.23 \pm 0.27 \pm 0.05) \times 10^{-3}$, where the errors are statistical, experimental systematic, and theoretical, respectively.
Contribution of the Author

The analysis described within this thesis was performed on data collected by the Belle Collaboration, of which the author is a member. The author contributed to the collection of the data. The Monte Carlo data sets were generated by the Belle Collaboration.

The analysis code was written by the author. It made use of local software libraries and interfaced with the Belle Analysis Software Framework (BASF). The local software libraries were developed by members of the Falkiner High Energy Physics research group, including the author. BASF was developed by the Belle Collaboration.

The analysis cuts, other than the $\cos \theta_{\gamma\gamma}$ cut, are standard selection criteria in $B$ physics in general and in exclusive semileptonic $B$ decays in particular. Introduction of the $\cos \theta_{\gamma\gamma}$ cut was at the author’s initiative. All selection criteria ultimately employed were evaluated and optimised by the author in the context of this analysis.

The author conducted the entire analysis (developing analysis code, analysing the data, optimising the cuts, conducting most of the systematic studies, etc.). The neutrino reconstruction systematic study referred to in Section 5.6.1 was conducted by Nicholas Parslow.

The author is responsible for all tables and figures unless otherwise indicated.
Acknowledgements

“Yeah...I am a scientist
'cause I can live on science alone.”
– The Dandy Warhols

Actually, I can’t. This thesis comes to you thanks to the toils and tribulations of many people. My thanks to everyone, especially:

Kevin Varvell, for being a wonderful supervisor, and for his patience with my code;

Lawrence Peak, for being a wonderful associate supervisor, and for his guidance in writing this thesis;

Anne Green, for her support in my formal Australian education;

The FHEP team over the past few years: Dr. Kevin Varvell, Dr. Andy Bakich, Assoc. Prof. Lawrence Peak, Dr. Bruce Yabsley, Dr. Aldo Saavedra, Nick Parslow, Sam McOnie, Yang Shi, Noel Hanna, Anthony Waugh, Jason Lee, Anmol Mishra, James Dunne, Clem Schlegel, Jules Morel, and Michaela Freeman;

The gang from my old office, Room 362: Daniel Andrzyczek, Nick Parslow, Oded Shrier, Michael Fitzgerald, Katherine Newton-McGee, and Aiping Zeng;

Jeremy Dalseno, Bruce Yablsey, Nick Hastings, Aki Nakahara, Jan Schueman, Jasna Dragic, Tony Limosani, Tim Gershon, Karina Arinstein, Seema Bahinipati, Manmohan Dash, the Shimada family, and everyone at Goldrush, for helping me survive in Japan;

Everyone at Morning and Afternoon Tea, for the conversation and entertainment over the past few years;

The Thomas family and all of my friends in Sydney, for being there for me and keeping me sane;

Mom, Dad, Ethan, Grandma, Grandpa, and all my friends and family who aren’t in Australia, for putting up with me being so far away;

Grandpa, for encouraging me to pursue a postgraduate degree;

Len, for encouraging me to stick with it when I wanted to quit;

Ceeto, for giving me the time to realise what I want to do and who I want to be. You were the best and smartest dog ever, and I miss you every day;

And of course Bez, who provided help, support, food, love, and clean dishes, and was very patient and wonderful with me when I was recovering from a bad concussion in the middle of my degree.
Contents

Contribution of the Author i

Acknowledgements iii

List of Figures ix

List of Tables xv

1 Introduction 1

2 Theory 3

2.1 The Standard Model of Particle Physics 3

2.1.1 Weak Interactions 4

2.1.2 Strong Interactions 5

2.1.3 The Baryon Asymmetry 5

2.2 CP Violation 6

2.3 The Cabibbo Angle 6

2.4 The Cabibbo-Kobayashi-Maskawa (CKM) Quark-mixing Matrix 7

2.5 The Unitarity Triangle 8

2.6 $b \rightarrow u$ Transitions 9

2.6.1 $V_{ub}$ 10

2.6.2 Dependence upon $q^2$ 11

2.7 Theoretical Approaches 11

2.7.1 Form Factors 11

2.7.2 Light-Cone Sum Rules (LCSR) 12

2.7.3 Lattice Calculations 13

2.7.4 Quark Models 14
List of Figures

2.1 The unitarity triangle. Here, the parameter \( \lambda = \frac{V_{cd}}{V_{ud}} \), and \( \frac{V_{tb}^*}{V_{ud}} \approx 1 \). Note: the BaBar Collaboration refers to the angles \( \phi_1, \phi_2 \) and \( \phi_3 \) as \( \beta, \alpha \) and \( \gamma \), respectively. Diagram courtesy of Kevin Varvell. ................................................. 9

2.2 Illustration of the unitarity triangle for the cases in which the interior angles do not add up to 180° (left) or a side extends beyond the vertex (right). ......................... 9

2.3 A Feynman diagram of the decay \( B \to \pi^0(\eta)\ell\nu \). Diagram courtesy of Kevin Varvell. ........................................ 10

2.4 \( q^2 \) distributions for \( B \to \pi\ell\nu \) predicted by [1], [2], [3], and [4] (left); and for \( B \to \rho\ell\nu \), predicted by [5], [6], [3], and [7] (right). From [8]. .................................................... 12

2.5 A Feynman diagram of a \( B \to X_s\ell\nu \) decay, including some possible gluon interactions. Diagram courtesy of Nicholas Parslow. ............................................. 12

2.6 A conceptual picture of “the lattice”. Important lattice parameters include the spatial extent of the lattice, \( L \sim 2 \) fm and the lattice spacing \( a \sim 0.1 \) fm. ................. 14

3.1 The KEKB accelerator (left) and the Belle detector (right). The images were adapted from [9]. ........................................................................................................ 18

3.2 The Belle Collaboration consists of 380 researchers from 55 institutions in 13 countries. The images are from [9] (left) and [10] (right). ............................... 18

3.3 The interaction region, showing the beam crossing angle and magnets. From [11]. ........................................... 19

3.4 Belle’s integrated luminosity, in fb\(^{-1}\). Belle’s luminosity has scaled greatly since commissioning in 1999. From [12]. .............................................. 20

3.5 Schematic cross-section of a silicon strip detector, from [13]. .................................................. 21

3.6 Schematic cross-section diagrams of SVD1 (left) and SVD2 (right), adapted from [14]. The relative scale of the two diagrams is approximately correct. .................... 21

3.7 Cross-sectional view of the CDC (innermost), ACC and TOF (outermost) detectors, from [15]. ................................................................. 24

3.8 Schematic diagram of barrel (left) and endcap (right) ACC modules, from [15]. .... 24

3.9 Mechanical assembly of the ECL counters. .................................................................................. 27

3.10 Schematic diagram of the forward and backward EFCs, from [11]. ................................. 27

3.11 Left: configuration of the superconducting coil within the aluminium stabiliser, adapted from [16]. Right: cross-section view of the solenoid coil and the outer surface of the support structure, adapted from [11]. ................................. 28
3.12 Schematic diagram of the solenoid and iron yoke, from [11].


3.14 Left: the BaBar detector, from [17]. Right: the CLEO detector, from [18].

4.1 The invariant mass distribution, in units of GeV, of photon pairs in $B \rightarrow \eta \ell \nu$ signal MC. The narrow peak at 0.135 GeV corresponds to $\pi^0$s, and the smaller, broader peak at 0.547 GeV corresponds to $\eta$s. In this figure, no restrictions are made on the photons’ energies.

4.2 The invariant mass distribution, in units of GeV, of photon pairs in $B \rightarrow \eta \ell \nu$ signal MC, for photon energy thresholds ranging from 50 to 300 MeV. The S/B of the $\pi^0$ and $\eta$ peaks begin to increase as the threshold is raised. The $\pi^0$ efficiency, however, decreases rapidly after $E_\gamma > 200$ MeV.

4.3 The invariant mass distribution, in units of GeV, of photon pairs in $B \rightarrow \eta \ell \nu$ signal MC, for photon energy thresholds ranging from 350 to 500 MeV.

4.4 Missing momentum ($|\vec{p}_{\text{miss}}|$; left) and missing energy ($E_{\text{miss}}$; right) in generated $B \rightarrow \pi^0 \ell \nu$ signal MC. Both histograms are in units of GeV. The central peak, and therefore the resolution, of the missing momentum distribution is narrower than that of the missing energy distribution. No cuts have been applied to these distributions.

4.5 Neutrino energy resolution, in units of GeV, calculated using the classic neutrino reconstruction (light grey histogram) and the modified CLEO technique (dark grey histogram). The modified CLEO technique dramatically improves the neutrino resolution.

4.6 cms lepton momentum vs. $q^2$ for $B \rightarrow \pi^0 \ell \nu$ signal MC. The cms lepton momentum and $q^2$ axes are in units of GeV and GeV$^2$, respectively. Note the correlation between cms lepton momentum and $q^2$.

4.7 $\cos \theta_{\text{thrust}}$ vs. $q^2$ for $B \rightarrow \pi^0 \ell \nu$ signal MC. The $\cos \theta_{\text{thrust}}$ axis is unitless, and the $q^2$ axis is in units of GeV$^2$. Note that $\cos \theta_{\text{thrust}}$ is not strongly correlated with $q^2$.

4.8 Two-dimensional histograms of the fit region, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. Both the $\Delta E$ and $p_T^\ell$ axes are in units of GeV. The events in these distributions have passed the lepton skim and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.

4.9 $\cos \theta_\nu$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
4.10 Net charge histograms, in units of $|e|$, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The histogram bins are symmetric about integer values, leading to the appearance of a non-integer cut; the actual cut values are $\pm 1e$. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 56

4.11 $M^2_{miss}$ histograms, in units of GeV$^2$, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .................. 57

4.12 cms lepton momentum ($p^*_\ell$) histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised lower cut threshold. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. Note that the lepton_skim includes a cut of $p^*_\ell > 1.2$ GeV, but slight differences between the particle identification algorithms used in the skim and event reconstruction result in a reconstructed cms lepton momentum of slightly less than 1.2 GeV in a small number of events. .......................... 58

4.13 $R_2$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised upper cut threshold. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 59

4.14 $\cos \theta_{BY}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 60

4.15 $\cos \theta_{\gamma\gamma}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised lower cut threshold. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 61

4.16 $\cos \theta_{thrust}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 62

4.17 $\cos \theta_{W\ell}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. .......................... 63
4.18 $m_{bc}$ histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. . . . 64

4.19 $\Delta E$ histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. . . . 65

5.1 The process of analysing $B \rightarrow \pi^0 \ell \nu$. The previous chapter described the steps up to and including the application of analysis cuts. This chapter will describe the fitting procedure and branching fraction calculation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ..
6.3 $\cos \theta_{BY}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined $uds$ and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.

6.4 $\cos \theta_{\gamma\gamma}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined $uds$ and charm continuum (bottom row). The vertical lines indicate the respective lower cut thresholds for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.

6.5 $\cos \theta_{\text{thrust}}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined $uds$ and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.

6.6 $\cos \theta_{W\ell}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined $uds$ and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.

6.7 $\Delta E$ (left) and $p^*_{\ell}$ (right) projections of the 2D fit result for $B \rightarrow \eta \ell \nu$. Both the $\Delta E$ and $p^*_{\ell}$ axes are in units of GeV.
List of Tables

2.1 The complete list of the fundamental particles in the Standard Model of Particle
Physics. The particles in the top three rows are the fermions that make up matter;
those in the bottom row are the gauge bosons that mediate the electromagnetic,
strong and weak interactions. The graviton, which is hypothesised but as of yet
unobserved, is not in the table. .................................................. 3

2.2 Summary of the SM description of the fundamental forces. Note that gravity has
not yet been incorporated into the SM. The relative strength is that between two
protons when just in contact. Values for the range and relative strength are from [19]. 4

3.1 Characteristics of SVD1 and SVD2, from [15]. .......................... 21

3.2 Configurations of the CDC sense wires and cathode strips, from [11]. ............ 22

3.3 Parameters of the TOF and TDC counters, adapted from [11]. .................... 25

3.4 Geometrical parameters of the ECL, from [15]. ................................. 26

3.5 Main parameters of the solenoid coil, from [11]. ............................... 29

3.6 Subdetectors used to identify final-state particles within Belle. The SVD and CDC
record charged particle tracks, while the ACC, TOF, ECL, EFC, and KLM record
energy deposits, as described in the text. ...................................... 30

3.7 A comparison of the Belle, BaBar and CLEO detectors. Information for CLEO-c
is not shown as it is not relevant to B physics. The listed integrated luminosities
include data collected on the $\Upsilon(4S)$ resonance only. ................... 31

4.1 A table representing a generated MC $B^- \rightarrow \pi^0\mu^-\nu_\mu$ event. Each row represents
a particle in the event: ID is an index number, starting with the $\Upsilon(4S)$ at ID=1;
ISTHEP is the particle’s decay status (1 for a final state particle and 2 for a particle
which decays); IDHEP identifies the type of particle using its LUND ID [20]; MOTHER
indicates the index of the particle’s mother (with M0(1) and M0(2) redundant but
retained in the code for historical reasons); the particle’s daughters, if any, have
indices ranging from DA(1) to DA(2); P(1) through P(4) are the particle’s four-
momentum in GeV, with its mass given in P(5), also in GeV; and V(1) through
V(4) are the particle’s production vertex positions and time, in mm and mm/c
respectively. .................................................. 35

4.2 Cross-sections and HadronB efficiencies for various processes in $e^+e^-$ collisions at
$\sqrt{s} = 10.58$ GeV, from [21]. QED refers to Bhabha ($e^+e^- \rightarrow e^+e^-$) and radiative
Bhabha ($e^+e^- \rightarrow e^+e^-\gamma$) processes. .............................. 37
4.3 lepton skim efficiencies for real data and various MC event types. Values, other than those for $b \rightarrow u\ell\nu$, are from [22].

4.4 Number of $\pi^0$s and $\eta$s, and their respective combinatorial backgrounds and figures of merit, for a range of photon thresholds. For the $\pi^0$, signal errors from the fit are typically 3% for photon thresholds up to 300 MeV and rise to 10% for the maximum photon threshold, while background errors from the fit are typically 1% for photon thresholds up to 150 MeV and rise to 9% for the maximum photon threshold. The $\eta$ signal exhibits the opposite trend, with signal errors from the fit typically 7-8% for photon thresholds up to 100 MeV and dropping to 4% for the maximum photon threshold; background errors from the fit are typically 1-3% for the entire range. All numbers were extracted using a fit (ROOFIT) [23]. The minimum photon energy thresholds used in the analysis were 100 MeV for $\pi^0$s and 350 MeV for $\eta$s.

4.5 Current $B \rightarrow \pi\ell\nu$ and $B \rightarrow \eta\ell\nu$ branching fraction (BF) measurements, obtained using neutrino reconstruction, semileptonic tagging, and full-reconstruction tagging. Charge conjugation is implied in all cases. The first, second and third errors are statistical, experimental systematic, and due to model uncertainty, respectively. Studies that used isospin assumptions to combine $\pi^+$ and $\pi^0$ modes include a factor of $2\frac{\tau_{B^0}}{\tau_{B^+}}$ to account for the difference between the $B^0$ and $B^+$ lifetimes. Note the low statistical error resulting from neutrino reconstruction. $B^+ \rightarrow \eta\ell^+\nu$ values and Belle’s full-reconstruction values are from [24]; all others are from [25]. All values are exactly as quoted by the Heavy Flavor Averaging Group (HFAG).

4.6 Primary decay modes of some of the light mesons in the $b \rightarrow u\ell\nu$ MC. For charged mesons, the decay mode(s) listed represent those of the positively charged meson, with charge-conjugation implied in all cases. All values are from [20].

4.7 Number of events and efficiencies in the complete $b \rightarrow u\ell\nu$ MC sample, and after various selection criteria.

4.8 The optimised set of event selection criteria, or “cuts”, for $B \rightarrow \pi^0\ell\nu$. These values were found through the optimisation procedure.

4.9 The effect of each cut on the efficiency of each sample, for $B \rightarrow \pi^0\ell\nu$. The final row shows the effect of all of the cuts put together. Cuts are not cumulative, except for the final row. The efficiencies listed are with respect to each sample after skimming, event reconstruction, and $B$ candidate selection, not with respect to the complete sample. All values are percentages except for S/B.

5.1 Signal, background and on-resonance data efficiencies in the fit region, $|\Delta E| < 1.0$ GeV and $p_{\ell}^* > 1.2$ GeV, after lepton skim and $B$ candidate selection, but before neutrino reconstruction and analysis cuts have been applied. All values listed are percentages.

5.2 Fitted signal and background yields, the Signal to Background ratio (S/B), and the $\chi^2/(\text{number of degrees of freedom})$ goodness-of-fit for the $B \rightarrow \pi^0\ell\nu$ fit.

5.3 Signal efficiency, fitted yields, and the partial branching fraction in each $q^2$ region. The statistical error of the summed region is the quadratic sum of those of the individual regions.

5.4 A summary of the systematic errors in the $B \rightarrow \pi^0\ell\nu$ branching fraction.
5.5 Effect of the use of $5.26 < M_{bc}^{(a)} < 5.29$ and $\cos \theta_{\gamma\gamma} > 0.9$ cuts on the branching fraction, statistical error, signal-to-background ratio (S/B), $\chi^2$ of the 2D fit in the $\Delta E - p_{T}^{\ell}$ plane (given by the fitting program), signal efficiency, and efficiency-corrected signal yield. The error quoted is statistical only. All other cuts were imposed and cut values were not varied from those in Sec. 4.10.2. This study examined a subset (Experiments 7-27) of the data used in the full analysis.

5.6 Variation in the $B \to \pi^0 \ell \nu$ branching fraction and signal efficiency due to the model used to generate $b \to u \ell \nu$ events. The LCSR measurement was summed over three $q^2$ bins, and the ISGW2 measurement was summed over two $q^2$ bins.

5.7 Preliminary $|V_{ub}|$ results, extracted from the $B \to \pi^0 \ell \nu$ branching fractions measured in this analysis. The branching fraction obtained with LCSR MC was used in the Ball’04 $|V_{ub}|$ extraction, and that obtained with ISGW2 MC was used in the ISGW2 $|V_{ub}|$ extraction. The latter is presented to indicate the extent of the theoretical error on the $|V_{ub}|$ measurement.

5.8 The most recent $B(B \to \pi \ell \nu)$ results. Charge conjugation is implied in all cases. The first, second, and third and fourth errors are statistical, experimental systematic, and due to $\pi$ and $\rho$ model uncertainties, respectively. Studies that used isospin assumptions to combine $\pi^+$ and $\pi^0$ modes include a factor of $2 \frac{\tau_0}{\tau_+}$ to account for the difference between the $B^0$ and $B^+$ lifetimes.

5.9 The preliminary $|V_{ub}|$ result obtained in this analysis, compared with recent $|V_{ub}|$ results extracted from the average of multiple $B \to \pi \ell \nu$ branching fractions [24]. The errors on the result from this analysis are statistical, experimental systematic, and due to theoretical modelling, respectively. The errors on the results from [24] are statistical and theoretical, respectively. In all cases, the $\Gamma_{th}$ values were extrapolated to the full $q^2$ region.

6.1 The optimised set of analysis cuts (event selection criteria), for $B \to \eta \ell \nu$. Those in bold use different threshold values from the $B \to \pi^0 \ell \nu$ study.

6.2 The effect of each cut on the efficiency of each sample. The final row shows the effect of all of the cuts put together. Cuts are not cumulative, except for the final row. The efficiencies listed are with respect to each sample after the lepton skim, event reconstruction, and $B$ candidate selection, not with respect to the complete sample. All values are percentages except for S/B.

6.3 Signal, background and on-resonance data efficiencies in the fit region, $|\Delta E| < 1.0$ GeV and $p_{T}^{\ell} > 1.2$ GeV, after the lepton skim, event reconstruction and $B$ candidate selection, but before any neutrino reconstruction or analysis cuts have been applied. All values listed are percentages except for S/B.

6.4 Fitted signal and background yields, the Signal to Background ratio (S/B), and the $\chi^2/(\text{number of degrees of freedom})$ goodness-of-fit for the $B \to \eta \ell \nu$ fit.

6.5 The preliminary $B \to \eta \ell \nu$ branching fraction measured in this analysis, compared with the most recent public $B \to \eta \ell \nu$ results. Charge conjugation is implied in all cases. The first, second and third errors are statistical, experimental systematic, and due to model uncertainty, respectively. The CKM06 results were presented in December 2006, and supercede the HFAG values, which are based on results presented at the ICHEP conference in August 2006.
Chapter 1

Introduction

According to current cosmological theory, when the Universe began with the Big Bang, equal amounts of matter and anti-matter were created. But astronomers see no anti-planets, anti-stars, nor anti-galaxies. As far out as we can see, there is a negligible amount of anti-matter. Where could it have gone?

The cause of this disparity, known as “the baryon asymmetry”, is one of the most intriguing questions in both cosmology and high energy physics. While cosmologists peer earlier and earlier into the Universe’s history with ever larger telescopes, high energy physicists probe tiny constituents of matter (and anti-matter), which they create in particle accelerators.

According to the Standard Model of particle physics, the baryon asymmetry is a result of a broken symmetry known as “$CP$”, the combination of charge-conjugation and parity. $CP$ turns a matter process identically into its anti-matter counterpart (and vice-versa), particle for anti-particle and right for left. It is in those rare weak processes where $CP$ is not conserved that matter could have gained the upper hand in the early universe.

The Standard Model description of $CP$ violation is contained wholly within the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The elements in the CKM matrix are proportional to the coupling constants that allow the three quark generations – up and down, charm and strange, top and bottom – to transition between one another. A bottom quark can decay into a charmed quark, a strange quark can decay into an up quark, and, more rarely but most relevant to the analyses described in this thesis, a bottom quark can decay into an up quark. Not all of these decays involve $CP$ violation; indeed, it is generally assumed that those examined in this analysis do not. But by measuring the rates at which the decays occur, values of the CKM matrix elements can be deduced. If the experimentally determined values differ greatly from theoretical predictions, this may point the way to “new physics”: physics beyond the Standard Model.

This analysis focuses on the semileptonic $B$ meson decays $B \to \pi^0 \ell \nu$ and $B \to \eta \ell \nu$. The CKM matrix element governing these decays is known as $V_{ub}$, because the $b$ quark within the $B$ meson emits a virtual $W$ boson and transitions to a $u$ quark, creating a $\pi^0$ or an $\eta$ meson in place of the $B$. $\pi^0$ and $\eta$ share a common decay mode ($\pi^0, \eta \to \gamma \gamma$), presenting an opportunity to measure two decays with a single analytical process.

The second chapter in this thesis will present the theory behind $B \to \pi^0 \ell \nu$ and $B \to \eta \ell \nu$ decays. In the third, the Belle detector, the experimental apparatus used in this analysis, will be described in detail. The fourth chapter describes the analytical procedure that allows tracks and energy clusters within the detector to be reconstructed as $B$ meson decay events, and that finds the $B \to \pi^0 \ell \nu$ and $B \to \eta \ell \nu$ signals which are buried in the deep ocean of background events. The
fifth and sixth chapters, respectively, present the $B \to \pi^0\ell\nu$ and $B \to \eta\ell\nu$ results. Finally, the results are discussed and possibilities for future work are suggested.
Chapter 2

Theory

2.1 The Standard Model of Particle Physics

The Standard Model (SM) is the current physical and mathematical description of the microscopic (subatomic) world. It encompasses three of the four fundamental forces of nature, but it is not yet a complete theory. It cannot explain gravitation, nor, of relevance to this thesis, does it provide a deep description of CP violation. While the SM’s CP formalism is consistent with experimental results, it does not, for example, explain the extent of the observed baryon asymmetry in the Universe (see Section 2.1.3).

Despite these issues, the SM is a remarkably successful model. It divides the fundamental constituents of matter into twelve spin-$\frac{1}{2}$ fermion fields: six quarks ($u, d, s, c, t, b$) and six leptons ($e^-, \mu^-, \tau^-, \nu_e, \nu_\mu, \nu_\tau$), and their respective antiparticles (see Table 2.1). Quarks interact via the strong, electromagnetic, and weak interactions. The charged leptons do not experience the strong force. Neutrinos “feel” only the weak interaction. The quarks and the charged leptons are point masses, with no finite size within current experimental limits. Neutrinos are considered to be massless within the SM, though neutrino oscillation observations have confirmed that they are (just barely) massive.$^1$

Table 2.1: The complete list of the fundamental particles in the Standard Model of Particle Physics. The particles in the top three rows are the fermions that make up matter; those in the bottom row are the gauge bosons that mediate the electromagnetic, strong and weak interactions. The graviton, which is hypothesised but as of yet unobserved, is not in the table.

<table>
<thead>
<tr>
<th>Strong</th>
<th>Weak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation 1</td>
<td>$u, d$</td>
</tr>
<tr>
<td>Generation 2</td>
<td>$c, s$</td>
</tr>
<tr>
<td>Generation 3</td>
<td>$t, b$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$g$ (gluon)</td>
</tr>
</tbody>
</table>

The fundamental forces are “mediated” by spin-1 gauge bosons: the photon governs electromagnetic encounters, the $Z^0$ and $W^\pm$ together are responsible for all weak interactions, and the strong

$^1$The period of neutrino flavour oscillations depends on the difference between the squares of the masses of the neutrino mass eigenstates, $\Delta m^2 = (m_2^2 - m_1^2)$. Hence, the observation of neutrino oscillations indicates that at least some neutrinos are massive. For a comprehensive review of neutrino mass and flavour oscillations, see [20]. For recent neutrino mixing results, see [20] and [26].
force is the domain of the eight electrically neutral (but colour-charged) gluons. Some extensions of the SM mention gravity in the proposed spin-2 graviton. These gauge bosons couple to the fermions with the following strengths:

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Gauge boson(s)</th>
<th>Range (m)</th>
<th>Relative strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fundamental Strong</td>
<td>gluon</td>
<td>∞</td>
<td>Large and variable</td>
</tr>
<tr>
<td>Residual Strong</td>
<td>pion</td>
<td>≤ 10^{-15}</td>
<td>1</td>
</tr>
<tr>
<td>Electromagnetic</td>
<td>photon</td>
<td>∞</td>
<td>10^{-2}</td>
</tr>
<tr>
<td>Weak</td>
<td>Z^0, W^±</td>
<td>10^{-18}</td>
<td>10^{-7}</td>
</tr>
<tr>
<td>Gravity</td>
<td>graviton (proposed)</td>
<td>∞</td>
<td>10^{-39}</td>
</tr>
</tbody>
</table>

Two fermions interacting via a gauge boson can be visualised with the following (naïve and very simplified) analogy. My dad and I used to play “doggy in the middle” with my dog Ceeto. We would toss a squeaky-toy back and forth, and he would follow it, running between us, hoping to catch it before we did. Sometimes, in the middle of the game, one of us would hide the toy behind our back, but we would pretend we were still throwing it back and forth. Ceeto wouldn’t believe us at first, but since we were acting as if nothing had changed (and clapping our hands to make a noise when we “caught” the toy), he would eventually go back to running between us. He couldn’t see the toy, but he could see us recoiling as the toy passed between us. (After a few volleys without the toy, we would bring it back into play, and let him catch it.)

In a remotely similar interaction, two electrons undergoing an elastic collision exchange virtual photons as they scatter off of each other. The electrons are analogous to the people recoiling as they throw the invisible ball, which represents the virtual photons. The photons can’t be seen; they exist only within the ΔEΔt ≈ ℏ limits of the Heisenberg uncertainty principle. The interactions mediated by virtual photons (and other gauge bosons) can be attractive (e.g., between particles with opposite electrical charges) or repulsive (e.g., between like electrical charges), unlike the solely repulsive “force” in the naïve analogy.

### 2.1.1 Weak Interactions

Most SM interactions obey various conservation laws; the weak interaction is the rebel. It is believed that weak interactions conserve electric charge, and weak isospin, but they do not conserve flavour, parity (P), charge conjugation (C), or the CP combination.

The decays examined in this thesis, \( B^± \rightarrow π^0 l^± \nu \) and \( B^± \rightarrow η l^± \nu \), involve a change of quark flavour, and therefore the weak interaction plays a critical role in this research. The \( b \) quark emits a virtual \( W^± \) boson\(^2\), which quickly decays into a positron (electron) and its associated neutrino, and becomes a \( u \) quark.

The coupling of the weak interaction is literally weaker than that of the electromagnetic and strong interactions. This weaker coupling leads to slower characteristic reaction times: weak decays occur less frequently than strong and electromagnetic decays.

---

\(^2\)The \( W \) is considered to be virtual because its coming into existence violates energy conservation in a classical sense, though for such a short time that the uncertainty equation \( ΔEΔt \approx ℏ \) holds.
2.1.2 Strong Interactions

Interactions between quarks and gluons are “strong”. One of the stranger discoveries of the past few decades is that when two quarks within a meson are drawn apart, the (strong) potential between them increases. At a certain point, the potential energy is great enough for a new quark-antiquark pair to appear out of the vacuum. These quarks join with those in the original meson, creating two new mesons. Free quarks have not been observed.

The eight gluons are the strong interaction’s gauge bosons. There is a trio of charges for the strong interaction (as compared to the positive and negative electric charges). This three-fold charge is known as colour, and the charges themselves are “red”, “green” and “blue”, with corresponding anti-charges of anti-red, anti-green and anti-blue. Quantum Chromodynamics (QCD) is the gauge quantum field theory for the strong interaction. QCD is analogous to QED in some respects, e.g. the massless mediating vector boson. QCD’s gluons, however, carry colour charge, while QED’s photons do not carry electromagnetic charge.

This research focuses on flavour-changing $B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \eta \ell \nu$ decays. The $b \rightarrow u$ transition is a weak interaction, but $B_s, \pi^0$ s and $\eta$s are quark-antiquark bound states and the valence and sea quarks and gluons within the mesons interact, as discussed below in Section 2.7.1.

2.1.3 The Baryon Asymmetry

According to modern cosmology, the Big Bang should have produced equal amounts of matter and antimatter. Today when we look into the sky, we see no evidence of antimatter – no anti-galaxies, no anti-nebulae, no anti-stars. We have not observed a border region, where a steady plane of gamma rays would indicate contact between a possible antimatter section of the Universe and the matter in the adjoining intergalactic medium. Sometime during the past 13.7 billion years, the Universe’s antimatter disappeared.

This lack-of-antimatter conundrum is known as the baryon asymmetry. In a symmetrical universe, the amount of antibaryonic antimatter would balance the amount of baryonic matter. There should either be plenty of antimatter, or only small amounts of both matter and antimatter and a large number of residual photons from their annihilation. Instead, we observe the small residual amount of matter resulting from a bias. The baryon-to-photon ratio is measured to be [27]

$$\eta \equiv \frac{n_B - n_\bar{B}}{n_\gamma} = (6.14 \pm 0.25) \times 10^{-10} \quad (2.1)$$

where $n_B$, $n_\bar{B}$ and $n_\gamma$ are the number densities of baryons, antibaryons and photons, respectively.

In 1966, theoretical physicist (and human rights activist) Andrei Sakharov proposed a mechanism involving three necessary conditions: some process that violates the conservation of baryon number, unequal rates of production and destruction of baryons and anti-baryons (non-equilibrium), and CP violation [28].

Unfortunately, the SM’s CP violation prediction is many orders of magnitude smaller than that required for the observed baryon asymmetry. Furthermore, results from the Belle and BaBar experiments thus far agree very well with the SM predictions [20].
2.2 CP Violation

CP violation is the violation of the combined “charge conjugation + parity” symmetry. The charge conjugation operator, \( \hat{C} \), changes the wavefunction of a particle into that of its antiparticle and vice versa; the effect is to change matter into antimatter (and vice versa), leaving the spatial part of the wavefunction, including momentum, intact. \( \hat{P} \), the parity operator, reflects the spatial part of a wavefunction through the origin. The CP combination converts a matter interaction completely into its antimatter counterpart (and, of course, vice versa).

CP violation can occur in decay (direct CP violation), by mixing (indirect CP violation), or via the interference between decays with and without mixing. In direct CP violation, the rate of a particle’s decay to a final state differs from the decay’s antimatter counterpart. CP violation by mixing (neutral particle oscillation) occurs when a neutral meson and its antiparticle are both observed to decay to the same final CP eigenstate. It can be observed following the production of a pair of neutral mesons, such as \( B^0\overline{B}^0 \). One of the pair is observed to decay to a final state accessible only to, say, a \( B^0 \), at time \( \tau_{tag} \). The other meson must then be a \( \overline{B}^0 \) at \( \tau_{tag} \). The \( \overline{B}^0 \) may or may not then oscillate into a \( B^0 \), and decays at time \( \tau_{decay} \) to a state accessible to both \( B^0 \) and \( \overline{B}^0 \). The \( (\tau_{decay} - \tau_{tag}) \) distribution is slightly different for \( B^0 \) tags and \( \overline{B}^0 \) tags. CP violation in the interference between decays with and without mixing can occur when a neutral meson and its antiparticle can decay to the same final state.

Until 1964, it was thought that CP was an exact symmetry. In that year, however, Christenson et al. observed CP violation in the neutral kaon system [29]. A larger CP violation effect can be seen in the \( B \) system.

CP violation and its effects are not the subject of this thesis\(^3\). It is, however, related to this research inasmuch as the mechanism for CP violation within the Standard Model is contained in the CKM matrix, and the decay rates examined in this research can be used to calculate the magnitude of the CKM matrix element \( V_{ub} \).

2.3 The Cabibbo Angle

As people began to study “strange” particles (those with an \( s \) or \( \overline{s} \) valence quark), they noticed odd things about strangeness-changing semileptonic decays. Strangeness \( S \) is defined such that \( S(s) = -1, S(\overline{s}) = +1, S(u) = S(\overline{u}) = S(d) = S(\overline{d}) = 0 \). Decays with \( \Delta S = 1 \) occurred much less frequently than those with \( \Delta S = 0 \). That is, decays such as

\[ n \rightarrow p + e^- + \overline{\nu}_e \quad (2.2) \]

were observed roughly twenty times more often than those such as

\[ \Sigma^- \rightarrow n + e^- + \overline{\nu}_e \quad (2.3) \]

In 1963, Nicola Cabibbo proposed an explanation [31]. He suggested that the eigenstates corresponding to the weak interaction are linear combinations of those of the strong interaction. Only three quarks were known at the time, so Cabibbo wrote the lepton and quark doublets as

\(^3\)For a review of CP violation in \( B \) mesons, see [30].
lepton doublets \( \left( \nu_e, \nu_e \right), \left( \nu_\mu, \nu_\mu \right) \)

quark doublet \( \left( \begin{array}{c} u \\ d \cos \theta_c + s \sin \theta_c \end{array} \right) = \left( \begin{array}{c} u \\ d' \end{array} \right) \) (2.4)

with \( \theta_c \) being the Cabibbo angle. With this quark doublet, the weak-interaction neutral-current matrix element, which governs the decay rate, in an \( s \to u \) transition would include a factor of \( \sin^2 \theta_c \), and that of a \( d \to u \) transition would have a factor of \( \cos^2 \theta_c \) [19]:

\[
\begin{align*}
    u\bar{u} + (d\bar{d} \cos^2 \theta_c + s\bar{s} \sin^2 \theta_c) & \quad (\Delta S = 0) \\
    + (s\bar{d} + d\bar{s}) \sin \theta_c \cos \theta_c & \quad (\Delta S = 1)
\end{align*}
\]

(2.5)

The Cabibbo angle was experimentally shown to be approximately 12°, modifying the decay rates by factors of \( \sim 0.05 \) and \( \sim 0.95 \), respectively. Strangeness-changing decays were suppressed.

The Cabibbo angle (moderately) successfully described the \( \Delta S \) discrepancy, but it did not explain it. The concept was expanded by Glashow, Iliopoulos and Maiani (GIM) in 1970, who posited the existence of a fourth quark, charm or \( c \) [32]. With a fourth quark, a second doublet was possible:

\[
\left( \begin{array}{c} u \\ d' \end{array} \right) = \left( \begin{array}{c} d \cos \theta_c + s \sin \theta_c \\ c \sin \theta_c - d \cos \theta_c \end{array} \right) \]

with the matrix element proportional to [19]

\[
\begin{align*}
    u\bar{u} + c\bar{c} + (d\bar{d} + s\bar{s} \cos^2 \theta_c + (s\bar{s} + d\bar{d}) \sin^2 \theta_c) & \quad (\Delta S = 0) \\
    + (s\bar{d} + d\bar{s} - s\bar{s} - d\bar{d}) \sin \theta_c \cos \theta_c & \quad (\Delta S = 1)
\end{align*}
\]

(2.7)

and the terms associated with strangeness-changing decays now cancelling out. Charm was found in the \( c\bar{c} \) state \( J/\psi \), discovered in 1974 [33] [34].

2.4 The Cabibbo-Kobayashi-Maskawa (CKM) Quark-mixing Matrix

The CKM matrix is the three-generation extension of the Cabibbo matrix. CP violation within the Standard Model is described completely by the CKM matrix. To date, the Belle, Babar and CLEO experiments have for the most part confirmed the Standard Model predictions of CP violation.

The matrix, \( V_{CKM} \), transforms quark mass eigenstates (\( e.g. \) \( d, s, b \)) into weak interaction eigenstates (\( e.g. \) \( d', s', b' \)):

\[
\left( \begin{array}{c} d' \\ s' \\ b' \end{array} \right) = \left( \begin{array}{ccc} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{array} \right) \left( \begin{array}{c} d \\ s \\ b \end{array} \right)
\]

The current matrix element values and their uncertainties are [20]:
The weak interaction eigenstates are thus linear combinations of the mass eigenstates. Another way of expressing this is that quarks can transmutate: the up-type quarks, those with charge $+\frac{2}{3}e$, can become down-type quarks, with charge $-\frac{1}{3}e$ via a weak transition; and quarks of the same charge can mix into each other:

$$V_{CKM} = \begin{pmatrix}
0.97383^{+0.00024}_{-0.00023} & 0.2272 \pm 0.0010 & (3.96 \pm 0.09) \times 10^{-3} \\
0.2271 \pm 0.0010 & 0.97296 \pm 0.00024 & (42.21^{+0.10}_{-0.80}) \times 10^{-3} \\
(8.14^{+0.32}_{-0.64}) \times 10^{-3} & (41.61^{+0.12}_{-0.78}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004}
\end{pmatrix}$$

The matrix elements get smaller as one moves away from the diagonal, indicating that quarks “prefer” to transform via the weak interaction into their own generation, or else into a close generation (e.g. $b \rightarrow c$) rather than skipping a generation (e.g. $b \rightarrow u$). $b$ quarks can only turn into $c$ or $u$ quarks because there are no direct flavour-changing neutral currents, and thus a $b$ cannot change directly into a $d$, nor a $c$ into a $u$, etc. The CKM matrix element $V_{ub}$ is proportional to the coupling constant for a $b$ quark mixing into a $u$ quark.

The matrix is sometimes approximated with the Wolfenstein parameterisation [35]:

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

where $\lambda = \sin \theta_c = V_{us}$. In this form\(^4\), it is clear that the matrix is nearly diagonal, and that the nine elements reduce to four independent parameters: $\lambda$, $A$, $\rho$ and $\eta$. The complex phase $\eta$ contains the CP violation predicted by the SM, as we will illustrate (literally) in the next section.

### 2.5 The Unitarity Triangle

The CKM matrix is unitary, meaning its adjoint, or Hermitian conjugate, is also its inverse. Hence

$$VV^\dagger = VV^{-1} = V^{-1}V = V^\dagger V = I$$

and the element formed by multiplying the $i^{th}$ row by the complex conjugate of the $j^{th}$ column is given by $\delta_{ij}$, the Kronecker delta. If we apply this to the first row and third column, those which

\(^4\)Remember that $\theta \sim 12^\circ$, and therefore $\sin \theta_c \sim 0.2 = \lambda$, $\sin^2 \theta_c \sim 0.05 = \lambda^2$ and $\sin^3 \theta_c \sim 0.009 = \lambda^3$. 

8
contain the element $V_{ub}$, we get

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$  \hspace{1cm} (2.8)$$

This can be represented by three vectors in the complex plane (see Fig. 2.1).

Figure 2.1: The unitarity triangle. Here, the parameter $\lambda = \frac{V_{cd}}{V_{ud}}$, and $\frac{V_{tb}}{V_{td}} \approx 1$. Note: the BaBar Collaboration refers to the angles $\phi_1, \phi_2$ and $\phi_3$ as $\beta, \alpha$ and $\gamma$, respectively. Diagram courtesy of Kevin Varvell.

One of the Belle Collaboration’s primary goals is to over-constrain the Unitarity Triangle. We want to measure each of the triangle’s interior angles to see if they do actually add up to 180 degrees. Similarly, we wish to know if the length of the sides correspond to their respective angles, or if one (or more!) side extends beyond one of its vertices, or not far enough to connect with an adjacent side, leaving the triangle open (see Fig. 2.2).

Figure 2.2: Illustration of the unitarity triangle for the cases in which the interior angles do not add up to 180° (left) or a side extends beyond the vertex (right).

Hundreds of scientists are working on hundreds of $B$ decay channels, and are extracting values for the sides and angles. The research described in this thesis aims to extract a value for $V_{ub}$.

The rest of this chapter will focus on $V_{ub}$.

### 2.6 $b \rightarrow u$ Transitions

A bottom quark can decay into an up quark by emitting a virtual $W$ boson, which sometimes decays into a charged lepton ($e$ or $\mu$) and an associated neutrino (see Fig. 2.3). The coupling for a quark transition is given by the appropriate CKM matrix element. For $b \rightarrow c$ transitions this is $V_{cb}$, which has a value of approximately 0.042 and is known to $\sim 2\%$ precision. For $b \rightarrow u$ transitions, such as those involved in this research, the matrix element is $V_{ub}$, which is smaller
by a factor of ten. Such transitions can be studied either inclusively, for example by examining semileptonic transitions in which the charged lepton momentum is close to or greater than that allowed in a $b \to c$ transition, or exclusively, by examining specific decay chains. The latter, specifically the exclusive decays of charged $B$ mesons to light, neutral, pseudoscalar mesons, is the focus of this work.

One advantage of studying exclusive decays is that by isolating a particular final-state hadron, one is not limited to using the lepton momentum spectrum to keep $b \to c$ transitions at bay; additional properties, most relating to the reconstructed $B$ meson, can be used to suppress the $b \to c$ background. The trade-off is lower statistics, with other $b \to u$ transitions being a significant source of background.

The current world-average $V_{ub}$ value calculated using inclusive decays differs by only a standard deviation from that calculated using exclusive decays [20]:

$$|V_{ub}| = (4.40 \pm 0.20 \pm 0.27) \times 10^{-3} \quad \text{(inclusive)}$$
$$|V_{ub}| = (3.84^{+0.67}_{-0.49}) \times 10^{-3} \quad \text{(exclusive)}$$

The cause of this disagreement remains unknown, and is not mirrored in $|V_{cb}|$ measurements [20]:

$$|V_{cb}| = (41.7 \pm 0.7) \times 10^{-3} \quad \text{(inclusive)}$$
$$|V_{cb}| = (40.9 \pm 1.8) \times 10^{-3} \quad \text{(exclusive)}$$

![Feynman Diagram](image)

Figure 2.3: A Feynman diagram of the decay $B \to \pi^0(\eta)\ell\nu$. Diagram courtesy of Kevin Varvell.

### 2.6.1 $V_{ub}$

The CKM matrix element $V_{ub}$ is proportional to the coupling constant in $b \to u$ transitions. It is the smallest and least-experimentally-constrained element in the CKM matrix. Mathematically, $|V_{ub}|$ can be extracted from the branching fraction of a semileptonic $B$ decay to a pseudoscalar meson $P$, such as $\pi^0$ and $\eta$, using the equation\(^5\)

$$\frac{dB(B^0 \to P\ell\nu)}{dq^2d(cos\theta_{W\ell})} = |V_{ub}|^2\tau_{B^0} \frac{G_F^2 k_\pi^3}{32\pi^3} \sin^2\theta_{W\ell} |f(q^2)|^2$$

\(^5\)|$V_{ub}| can also be extracted from semileptonic $B$ decays to vector mesons. The equation, which is more complicated as such transitions involve three form factors, is not directly relevant to this analysis and is therefore omitted.
where \( q^2 \) is the square of the four-momentum carried by the virtual \( W \), \( \theta_{Wl} \) is the angle of the lepton’s three-momentum in the \( W \)’s frame with respect to the \( W \)’s three-momentum in the \( B \)’s frame, \( \tau_B \) is the \( B \) lifetime, \( G_F \) is the Fermi constant, \( k_\pi \) is the pion momentum in the \( B \) rest frame, and \( f(q^2) \) is a form factor. The results of most \( b \to u\ell\nu \) form-factor theory papers are given for \( B \to \pi^0\ell\nu \); \(|V_{ub}|\) can be extracted from an experimentally determined \( B \to \pi^0\ell\nu \) branching fraction using the following simplified formula:

\[
|V_{ub}| = \sqrt{\frac{2 \times B(B^+ \to \pi^0\ell\nu)}{\Gamma_{thy}\tau_{B^+}}} \tag{2.10}
\]

where the factor of 2 comes from assumed \( \pi^0 - \pi^+ \) isospin symmetry, \( \tau_{B^+} \) is the \( B^+ \) lifetime, and all other quantities in Eq. 2.9 have been absorbed into \( \Gamma_{thy} \).

### 2.6.2 Dependence upon \( q^2 \)

In \( b \to u \) transitions, strong interaction effects are greatly reduced as the decay of the \( b \) quark is governed by the weak interaction. Interactions between the \( b \) (and the resulting \( u \)) and the “spectator” quark, however, are governed by quantum chromodynamics (QCD). The QCD effects on the hadronisation of the final state meson are described by form factors, which characterise the structure of the mesons. The four-momentum of the virtual \( W \) (and the charged and neutral leptons which result from it) affects the overlap between the wavefunctions of the initial- and final-state mesons. This effect is described in terms of \( q^2 \), which is the square of the “mass”\(^6\) of the virtual \( W \):

\[
q^2 = E_W^2 - (\mathbf{p}_W)^2 = (E_l + E_\nu)^2 - (\mathbf{p}_l + \mathbf{p}_\nu)^2 \tag{2.11}
\]

Different theoretical approaches predict different shapes for the \( q^2 \) distribution (see Fig. 2.4). It is therefore very important for experimentalists to determine the shape of the observed spectrum, and to compare it with those predicted by various theories.

### 2.7 Theoretical Approaches

#### 2.7.1 Form Factors

Form factors describe the strong interaction’s effect in semileptonic decays. These are somewhat independent of the weak decay of the \( b \) quark, as they describe the inter-quark dynamics of the \( b/u \) and the “spectator” \( d \) (see Fig. 2.5). \( B \) decays to light vector mesons, such as the \( \rho \), involve four form factors due to the final state’s helicity. The matter is simpler for pseudoscalar final states, for which there are only two heavy-to-light form factors, and in fact one of these can often be ignored\(^7\). The \( B \) to \( \pi \) hadronic matrix element can be written as \[^{[30]}\]

\[
< \pi|\mathbf{p}_\mu b|B >= f_+(q^2)(p_B + p_\pi)_\mu + f_-(q^2)(p_B - p_\pi)_\mu \tag{2.12}
\]

Due to their small masses, the form factor \( f_- \) can be neglected for \( l = e, \mu \).

---

\(^6\)Since the \( W \) is virtual, it is off the mass shell: \( E^2 - (\mathbf{p})^2 \neq m_{\text{nominal}}^2 \).

\(^7\)The second form factor is greatly suppressed for small lepton masses \((m_e, m_\mu)\).
The quark-quark interactions, and therefore the form factors, depend on the relative momenta of the $b$ and $u$ quarks. Different theoretical approaches are required to deal with different $q^2$ regimes: Light-Cone Sum Rules (LCSR) are appropriate when the light meson’s momentum is large in the $B$’s rest frame (low $q^2$), and lattice QCD calculations can only be achieved when the light meson is nearly at rest in the $B$’s frame (high $q^2$). Quark models depict the (two, valence-only) quarks in a meson as moving independently, and can be applied to a wide $q^2$ range.

Transitions from a heavy quark to another heavy quark, say $b \rightarrow c$, can be described by Heavy Quark Effective Theory (HQET) [36, 37, 38, 39]. Heavy to light meson transitions, such as $B \rightarrow \pi$ and $B \rightarrow \eta$, are not as well understood theoretically. Several approximations and models have been conceived in an attempt to describe them. The current leading theories are now discussed.

\[ ( l = e, \mu ) \quad W \quad V_l \quad l^+ \text{ or } l^- \]

\[ B \quad X_u \quad ( \bar{u} \text{ or } \bar{d} ) \]

Figure 2.5: A Feynman diagram of a $B \rightarrow X_u \ell \nu$ decay, including some possible gluon interactions. Diagram courtesy of Nicholas Parslow.

### 2.7.2 Light-Cone Sum Rules (LCSR)

The strong interaction description of the $B \rightarrow \pi^0 \ell \nu$ transition must include both the large-scale bound states (the $B$ and $\pi^0$ mesons) and the small-scale interactions of those states’ partons (the $b$, $d$ and $u$ valence quarks, as well as any gluons and sea quarks that may appear out of the...
vacuum). Conventional QCD perturbation theory (expansion in $\alpha_s$) is applicable to the latter, but breaks down at the hadron’s large scale.

The two scales can be accommodated by the use of correlation functions, combined with sum rules: non-trivial constraints on sums over hadronic parameters. In the case of the weak-decay form factor calculations for $B \to \pi^0 \ell \nu$, the functions correlate the weak current $V_\mu$ with a current whose quantum numbers are those of the $B$ meson. In LCSR, also known as QCD sum rules on the light-cone, the products of the currents are expanded near the light cone, where the quark’s four-vector has zero interval:

$$x^2 = x_0^2 - \vec{x}^2 \sim 0$$ (2.13)

and QCD is quantised at a fixed light-cone time, $\tau = t + z/c$. The wave function of the relativistic pion becomes [40]

$$|\pi\rangle = |q\bar{q}\rangle \psi_{\pi q} + |q\bar{q}g\rangle \psi_{\pi qg} + \ldots$$ (2.14)

where each $\psi_n$ describes a state corresponding to a particular number of quarks and gluons, at the same given “time” $\tau$.

In the light-cone expansion, the hadron distribution amplitudes $\phi$ are convoluted with process-dependent amplitudes $T$ [41]:

$$\text{correlation function} \sim \sum_n T^{(n)} \otimes \phi^{(n)}$$ (2.15)

where the sum is over contributions with increasing twist.\(^8\) The correlation equation for $B \to \pi^0 \ell \nu$ decays is [41]

$$\Pi_\mu(q,p_B) = i \int d^4x e^{iqy} \langle \pi(p_{\pi})|TV_\mu(x)j_B^\dagger(0)|0\rangle$$ (2.16)

where $\mu$ is the renormalisation scale; $p_B$ and $p_\pi$ are the $B$’s and the $\pi^0$’s momenta, respectively; $q = p_B - p$; $V_\mu$ is the weak current; $j_B = m_B i \gamma_5 b$ is the interpolating field for the $B$ meson; and $T$ is the process-dependent amplitude. Eq. 2.16 represents the $B$’s field sandwiched (i.e., evaluated) between the vacuum and the $\pi^0$ state.

LCSR calculations are limited to the range $0 \leq q^2 \leq 14 \text{ GeV}^2$ because of the restriction

$$m_B^2 - q^2 \geq \mathcal{O}(\Lambda_{QCD} m_b)$$ (2.17)

where $\Lambda_{QCD}$ is the QCD length scale. This corresponds to the regime in which the $\pi^0$ has sufficient energy in the $B$’s rest frame to travel near the light-cone. LCSR form factors must be extrapolated to $q^2 > 14 \text{ GeV}^2$.

### 2.7.3 Lattice Calculations

Lattice QCD is a computational method that numerically evaluates physical quantities from first principles. The lattice itself is a conceptual, discrete version of space-time (an array of space-time points; see Figure 2.6). Quarks are placed onto points on the lattice, and the action [30]:

$$S(A_\mu, \bar{\psi}, \psi) = \int d^4x \left[ -\frac{1}{4} \epsilon^{\mu\nu\rho\sigma} G_{\mu\nu}^{\rho\sigma} + \bar{\psi}(i\gamma_5 \vec{\partial} - g A - m)\psi + \ldots \right]$$ (2.18)

is computed in the discrete volume $L^3$. Typically, there are 16-20 lattice points in each direction, with a spacing $a$ of approximately 0.1 fm between them. The action is evaluated for all possible

---

\(^8\)The twist $t$ of an operator $\mathcal{O}$ with spin $l$ is defined as $t = \dim[\mathcal{O}] - l$. 

13
configurations of the given quark \( (\psi) \) and potential \( (A_\mu) \) fields, in order to calculate the partition function:

\[
Z = \int [dA_\mu][d\bar{\psi}_i][d\psi_i]e^{iS(A_\mu, \bar{\psi}_i, \psi_i)}
\]

(2.19)

The integral is over all fields at each point on the lattice.

![Lattice Diagram](image)

Figure 2.6: A conceptual picture of “the lattice”. Important lattice parameters include the spatial extent of the lattice, \( L \sim 2 \) fm and the lattice spacing \( a \sim 0.1 \) fm.

The inverse of the lattice spacing is related to the quark mass accessible to the calculations. With the current lattice spacing, \( a^{-1} \sim 2 - 4 \) GeV, the \( b \) quark cannot be placed directly on the lattice. Instead, the simulations use a variety of masses around the charm quark mass, and results are extrapolated to realistic \( m_b \) values using Heavy Quark Effective Theory (HQET) scaling rules.

Nor can the light quarks be used on the lattice. Singularities appear in the propagators as quark masses approach 0 \( (m_u \sim 5 \) MeV, \( m_d \sim 10 \) MeV), so a “light” quark of mass of \( \sim 100 \) MeV is used and results are extrapolated to appropriate values.

The precision of these calculations can be improved by decreasing the lattice spacing \( a \) and increasing the size of the lattice \( L^3 \). This can, however, require an extraordinary amount of computer power.

Until 2004, most \( b \to u\ell\nu \) lattice calculations were performed in the quenched or valence approximation. This ignores the possibility of quark (or gluon) loops, resulting in a considerably faster simulation. The downside of the approximation is that errors are introduced by neglecting the loop diagrams, and the extent of these errors are unknown for \( b \to u \).

The finite size of the lattice means that all mesons in a heavy-to-light semileptonic decay must have small momenta, corresponding to high \( q^2 \). Lattice \( b \to u \) form factors, therefore, are most appropriately used at high \( q^2 \), when most of the momentum is transferred to the charged lepton and the neutrino, and the recoiling pion (or other light meson) is nearly at rest.

### 2.7.4 Quark Models

As we have seen, QCD cannot currently be used to directly describe all semileptonic \( B \) decays at all \( q^2 \) values. Years of probing the structure of hadrons has shown us that quarks within mesons act to some extent like independent (though confined) free particles. This is the basis for quark models.
Quark models treat mesons as bound states of two valence quarks. They tend to ignore the possibility of sea quarks. The models are generally based on quantum mechanics rather than field theory and relativity, and are not derived from QCD. Some relativistic effects do have to be included: in $B \rightarrow D$ transitions, where much of the four-momentum is absorbed into the $D$ meson mass ($m_D \sim 1870$ MeV), the velocity of the $D$ in the $B$’s rest frame can be up to $0.7c$. The $D$’s mass is more than 13 times greater than the pion’s; clearly the pion in a $B \rightarrow \pi$ transition can be relativistic.

The quark model used in the systematic study of this analysis is known as ISGW2 [3]. It is a hybrid quark model, meaning it uses a quark model to calculate form factor values at $q_0$, and extrapolates to other $q^2$ values with an ad hoc ansatz. ISGW2 uses $q_0 = q_{\text{max}}$, the maximum possible $q^2$ value for the system. The advantage of this approach is that the initial- ($B$) and final-state ($\pi, \rho, \eta$, etc.) mesons share the same rest frame (remember that high $q^2$ means low meson momentum). The form factor at this “zero recoil” is the overlap between the initial- and final-state meson wave functions.

The problem with quark models is that it is not always possible to know what systematic effect(s) the model may cause, particularly when the model is used to extract physical parameters. Early $b \rightarrow u\ell\nu$ measurements relied on the ISGW2 [3] quark model as an input for signal and crossfeed Monte Carlo. To date, no theoretical $B \rightarrow \pi\ell\nu$ description has been excluded, but quark models tend to be more disfavoured than LCSR or lattice predictions [42].

It should be noted that because models are designed for specific purposes, it is often inappropriate to apply models to situations other than those for which they were designed.
Chapter 3

The KEKB Collider and the Belle Detector

“The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Čerenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter comprised of CsI(Tl) crystals (ECL) located inside a super-conducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect $K^0_L$ mesons and to identify muons (KLM). The detector is described in detail elsewhere [11]. Two inner detector configurations were used. A 2.0 cm beampipe and a 3-layer silicon vertex detector was used for the first sample of $152 \times 10^6 B\bar{B}$ pairs, while a 1.5 cm beampipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining $383 \times 10^6 B\bar{B}$ pairs[14].”

The Belle detector, and to a lesser extent the KEKB accelerator (Fig. 3.1), have been described briefly and succinctly many times by many people (as in the above template excerpt). But what are each of the subdetectors? How do they work? What are the acceptance parameters used for each in this research? This chapter is intended to address these questions.

Most of the information in this chapter comes from References[11] and [16] unless otherwise indicated.

3.1 The Belle Collaboration

The primary purpose of the Belle experiment is to search for CP violation in the $B$ meson system. Not all Belle analyses, however, relate directly to CP violation. Some search for rare $B$ meson decays. Others use Belle’s extensive data set to study tau lepton, two-photon or charm physics. Still others, including this analysis, seek to precisely determine CKM matrix elements and parameters of the unitarity triangle.

1From the Belle 2006 summer conference paper template.
At the time of writing, the Belle Collaboration consisted of 380 physicists from 55 institutions in 13 countries (see Fig 3.2). The Collaboration is divided into research groups such as $\phi_3$, Indirect CP Violation/Rare $B$ Decays, Charm Studies, Tau/Two Photon, and CKM Matrix Elements, of which I am a member. These groups meet at KEK with videoconference support to discuss analyses on a weekly, fortnightly, or monthly basis. The entire Collaboration meets ten times a year (again with videoconference support), with Belle Analysis Meetings and Belle General Meetings alternating approximately every month.
3.2 The KEKB Accelerator

KEK stands for Kō Enerugi Kasokuki Kenkyu Kikō, or the High Energy Accelerator Research Organization. It is located in Tsukuba “Science City”, Ibaraki Prefecture, Japan, amidst crop fields near the foot of Mount Tsukuba. The main ring, which has a circumference of 3016 m [43], first began operation in 1984 for the TRISTAN experiment [44].

KEKB collides 8.0 GeV electrons on 3.5 GeV positrons. The electron beam is known as the High Energy Ring (HER) beam and the positron beam is known as the Low Energy Ring (LER) beam. These run at currents of up to 1401 mA and 2000 mA, respectively [45]. The beam energies are tuned so that the centre-of-mass (CM) energy corresponds to the mass of the Υ(4S) resonance\(^2\). At this energy, approximately one in four \(e^+e^-\) annihilation events produces a Υ(4S). The remaining interactions result in \(q\bar{q}\) pairs \((q = u, d, s, c)\), which are known as the “continuum background” in \(B\) analyses.

The accelerator runs twenty-four hours a day, seven days a week, from September through June (with short a break at the beginning of the calendar year).

3.2.1 Crossing Angle

The colliding beams cross at an angle of \(±11\) milliradians, with the LER beam aligned with the solenoid’s field and the HER beam at a 22mr angle (see Fig. 3.3). The decision to use a non-zero crossing angle was made in order to maximise the luminosity, which is defined as

\[
\mathcal{L} = \frac{\text{number of events per second}}{\text{interaction cross section}} \quad (3.1)
\]

and has dimensions of \([\text{area}]^{-1}[\text{time}]^{-1}\).

With this configuration, all RF buckets can be filled with the beam without increasing the risk of parasitic collisions. Additionally, beam-related backgrounds within the detector are reduced as there is no need for separation-bend magnets.

\(^2m_{\Upsilon(4S)} = (10.5794 \pm 0.0012)\) GeV; \(B(\Upsilon(4S) \rightarrow B\bar{B}) > 96\%\) [20]
3.2.2 Continuous Injection

In September 2003, KEKB began using “continuous injection”, meaning the beams can now be re-populated without pausing the data acquisition (DAQ) system. Prior to this, data collection ceased approximately once an hour for several minutes while the beams were filled. Continuous injection has greatly increased Belle’s integrated luminosity (the total recorded luminosity for a given period of time; in this case, the entire experiment), which currently stands at $710.254 \text{ fb}^{-1}$ [45]. KEKB boasts the highest peak luminosity in the world, having logged $1.7118 \times 10^{-34} \text{ cm}^{-2} \text{s}^{-1}$ [45] (see Fig. 3.4).

The Belle detector will now be discussed from the inside out.

3.3 Silicon Vertex Detector (SVD) and Beampipe

The innermost layer of the Belle Detector is the Silicon Vertex Detector (SVD), which includes the beampipe surrounding the beams’ collision point (Interaction Point, or IP). The portion of the beampipe around the IP is a pair of concentric cylinders separated by 0.5 mm. Paraffin oil is circulated between these layers at a flow rate of 1.15 litres per minute as a cooling liquid. The temperature of the paraffin at the inlet is 14°C.

Semiconductor detectors are the best for providing spatial and energy resolution [13]. They are essentially thin, reverse-biased p-n junction layers, with a metal layer on one side and alternating strips of metal electrode and SiO₂ on the other (see Figure 3.5). The bulk material is doped to be n-type, with thin p-type strips underneath the metal electrode strips. Ionising particles create electron-hole pairs as they pass through the silicon. An applied potential causes the holes to collect at the negative-potential metal layer, and the electrons to collect at the anodic electrode strips. The resulting electrode pulse is a measure of both the energy deposited by the ionising particle (the amount of charge generated) and its position (the particular electrode strip that pulsed).

The SVD itself consists of ladders of $57.5 \times 33.5 \text{ mm}^2$ double-sided silicon strip detectors (DSSDs), reinforced by boron nitride support ribs. In the original design, SVD1, there was a total of 102 DSSDs; an upgrade to SVD2 increased the number to 246. Each DSSD contains 1280 electrode strips and 640 readout pads. The ladders are arranged in a series of cylindrical layers (see Fig. 3.6). These provide a resolution of $\sim 100 \mu \text{m}$. 
Figure 3.5: Schematic cross-section of a silicon strip detector, from [13].

Figure 3.6: Schematic cross-section diagrams of SVD1 (left) and SVD2 (right), adapted from [14]. The relative scale of the two diagrams is approximately correct.

Table 3.1: Characteristics of SVD1 and SVD2, from [15].

<table>
<thead>
<tr>
<th></th>
<th>SVD1</th>
<th>SVD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beampipe radius (mm)</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Number of layers</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Number of DSSD ladders in layers 1/2/3/4</td>
<td>8/10/14/NA</td>
<td>6/12/18/18</td>
</tr>
<tr>
<td>Number of DSSDs in a ladder in layers 1/2/3/4</td>
<td>2/3/4/NA</td>
<td>2/3/5/6</td>
</tr>
<tr>
<td>Radii of layers 1/2/3/4 (mm)</td>
<td>30.0/45.5/60.5/NA</td>
<td>20.0/43.5/70.0/88.8</td>
</tr>
<tr>
<td>Angular coverage (acceptance)</td>
<td>23 $\leq \theta &lt; 140^\circ$ (0.86)</td>
<td>17 $\leq \theta &lt; 150^\circ$ (0.92)</td>
</tr>
<tr>
<td>Total number of channels</td>
<td>81920</td>
<td>110592</td>
</tr>
<tr>
<td>Strip pitch ($\mu$m) for z</td>
<td>84</td>
<td>75 (73 for layer 4)</td>
</tr>
<tr>
<td>Strip pitch ($\mu$m) for r$\phi$</td>
<td>25 (50 for readout)</td>
<td>50 (65 for layer 4)</td>
</tr>
<tr>
<td>DSSD thickness ($\mu$m)</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Total material at $\theta = 90^\circ$ ($%X_0$)</td>
<td>1.85</td>
<td>2.6</td>
</tr>
<tr>
<td>Readout VLSI</td>
<td>VA1</td>
<td>VA1TA</td>
</tr>
<tr>
<td>Radiation tolerance (MRad)</td>
<td>$\sim 1$</td>
<td>$&gt; 20$</td>
</tr>
<tr>
<td>Intrinsic DAQ deadtime/trigger ($\mu$s)</td>
<td>128</td>
<td>25.6</td>
</tr>
</tbody>
</table>
In 2003, the SVD was upgraded from SVD 1.6 to SVD 2.0 and the beampipe around the IP was replaced by one with a smaller radius. The SVD had been replaced three times since Belle was commissioned, but while SVD1.2 and SVD1.4 were improvements upon SVD1.0, the differences between the first four versions were comparatively small. SVD2 was a major redesign, including an additional layer of DSSDs and a small-cell drift chamber between the final layer and the CDC.

### 3.4 Central Drift Chamber (CDC)

Drift chambers measure the paths of charged particles. The gas- or liquid-filled chamber is composed of cathodes, field wires which produce an electric field, and anode sense wires. As a charged particle passes through the chamber, it ionises gas molecules within the chamber. The ionisation electrons drift towards the anodes. Successive “hits” within the drift chamber form the particle’s track.

The CDC detects charged particles after they leave the SVD, and is integral in both charged particle tracking and particle identification. It provides a highly efficient $z$-trigger for identifying tracks produced at the interaction point, and the extremely high momentum resolution required for Belle analyses. Even in the high momentum range of $4 \text{ GeV} < p_t < 5.2 \text{ GeV}$, the transverse momentum resolution is 1.85%; most particles created in $B$ meson decays have a momentum lower than 1 GeV [46].

The CDC is composed of fifty cylindrical layers of sense wires and three cathode strip layers, and is filled with a mixture of 50% He and 50% ethane ($C_2H_6$). Each layer contains three cathode strip layers, and three to six layers which are either axial or small-angle-stereo. The inner three layers are smaller than the rest. There is no inner CDC wall. In total there are 8400 nearly-square drift cells. The CDC is asymmetric in the $z$ direction with an azimuthal range of $17^\circ < \theta < 150^\circ$, and is conical at either end at small $r$. The radius of the CDC extends from $103.5 \text{ mm} < r < 874 \text{ mm}$. The longest wires in the CDC are 2400 mm long.

<table>
<thead>
<tr>
<th>Superlayer type and no.</th>
<th>No. of layers</th>
<th>Signal channels per layer</th>
<th>Radius (mm)</th>
<th>Stereo angle (mrad) and strip pitch (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cathode 1</td>
<td>64(\text{z}) × 8(\text{φ})</td>
<td>83.0</td>
<td>(8.2)</td>
<td></td>
</tr>
<tr>
<td>Axial 1</td>
<td>64</td>
<td>88.0-98.0</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Cathode 1</td>
<td>80(\text{z}) × 8(\text{φ})</td>
<td>103.0</td>
<td>(8.2)</td>
<td></td>
</tr>
<tr>
<td>Cathode 1</td>
<td>80(\text{z}) × 8(\text{φ})</td>
<td>103.5</td>
<td>(8.2)</td>
<td></td>
</tr>
<tr>
<td>Axial 1</td>
<td>64</td>
<td>108.5-159.5</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Stereo 2</td>
<td>80</td>
<td>178.5-209.5</td>
<td>71.46-73.75</td>
<td></td>
</tr>
<tr>
<td>Axial 3</td>
<td>96</td>
<td>224.5-304.0</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Stereo 4</td>
<td>128</td>
<td>322.5-353.5</td>
<td>-42.28- -45.80</td>
<td></td>
</tr>
<tr>
<td>Axial 5</td>
<td>144</td>
<td>368.5-431.5</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Stereo 6</td>
<td>160</td>
<td>450.5-497.5</td>
<td>45.11-49.36</td>
<td></td>
</tr>
<tr>
<td>Axial 7</td>
<td>192</td>
<td>512.5-575.5</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Stereo 8</td>
<td>208</td>
<td>594.5-641.5</td>
<td>-52.68- -57.01</td>
<td></td>
</tr>
<tr>
<td>Axial 9</td>
<td>240</td>
<td>656.5-719.5</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>Stereo 10</td>
<td>256</td>
<td>738.5-785.5</td>
<td>62.10-67.09</td>
<td></td>
</tr>
<tr>
<td>Axial 11</td>
<td>288</td>
<td>800.5-863.0</td>
<td>0.</td>
<td></td>
</tr>
</tbody>
</table>
The field wires are unplated aluminium and have a diameter of 126 $\mu$m. They provide an electric field of less than 20 kV/cm at the surface of the wires (minimising the radiation damage to the subdetector) and high electric fields up to the edge of the cell. This minimises scattering by reducing the material of the chamber, and simplifies the relation between the drift-time and the distance. The sense wires are gold-plated tungsten and are 30 $\mu$m in diameter. The total wire tension of 3.5 tons is supported by aluminium endplates, which contain anodes and cathodes, and carbon-fibre reinforced plastic (CFRP) cylinder structures. The CDC has a $dE/dx$ resolution of about 6% for Bhabha electrons, and a $K/\pi$ separation of up to 0.8 GeV.

### 3.5 Aerogel Čerenkov Counters (ACC)

When a charged particle travels through a dielectric medium with a speed ($\beta = v/c$) which exceeds the speed of light in that medium ($c/n$), excited atoms near the particle become polarised and coherently emit photons at a distinct angle $\theta$ which depends on the particle’s speed and the medium’s index of refraction ($n$):

$$\cos \theta = 1/\beta n$$  \hspace{1cm} (3.2)

where $\beta > 1/n$. This is known as the Čerenkov effect, and the emitted photons are known as Čerenkov light or radiation [47]. Čerenkov detectors discriminate between massive particles which travel comparatively slowly through the detector material, and light particles which travel through the detector faster than the speed of light in that material (and therefore emit Čerenkov radiation). Belle uses silica aerogel threshold Čerenkov counters to distinguish between charged pions and charged kaons. The barrel counter modules have indices of refraction ranging from 1.010 to 1.028, depending on their polar angle with respect to the beamline. The endcap counters were designed for flavour tagging; their index of refraction is 1.030.

There are 960 counter modules in sixty cells in the $\phi$ direction of the barrel, and 228 counter modules arranged in five concentric layers in the forward endcap (see Fig. 3.7). Each module consists of a stack of five aerogel tiles contained within a $\sim 12 \times 12 \times 12$ cm$^3$, 0.2mm-thick aluminium box. The aerogels are coupled to one or two fine-mesh 2-, 2.5- or 3-inch diameter photomultiplier tubes (FM-PMT; see Fig. 3.8). The FM-PMT diameter varies with the refractive index of the aerogels (and therefore the angle of the Čerenkov radiation). The ACC can detect particles with momentum in the range $1.2$ GeV $< p < 3.5$ GeV.

---

$^3$Yes, the PMT diameters are measured in inches!
Figure 3.7: Cross-sectional view of the CDC (innermost), ACC and TOF (outermost) detectors, from [15].

Figure 3.8: Schematic diagram of barrel (left) and endcap (right) ACC modules, from [15].
3.6 Time-of-Flight Counters (TOF)

As a charged particle passes through a scintillating material, atoms in the material are excited into higher energy states. In plastic scintillators, some of the excitation energy is dissipated into the lattice, and some is radiated as light (scintillation). A light guide is generally used to transmit the light to a photomultiplier tube (PMT) which converts the photons into an electronic signal.

Two scintillation counters separated by a distance $L$ can be used to identify particle masses. A relativistic particle with momentum $p$ travelling between two counters separated by a distance $L$ will do so at a rate of ($c = 1$) [48]:

$$t/L = (3333/p)(p^2 + m^2)^{1/2} \text{ ps/m}$$

(3.3)

Two particles with different masses but the same momentum travelling between the counters will have a time difference per unit length of [48]:

$$(t_1 - t_2)/L \simeq 1667(m_1^2 - m_2^2)/p^2 \text{ ps/m}$$

(3.4)

The fast time resolution ($\sim 100$ ps) of TOF counters can also be used to initiate trigger signals to other subdetectors.

Belle uses pairs of plastic (Bicron BC408) time-of-flight (TOF) scintillation counters to measure charged particle time-of-flight. These are augmented by plastic (Bicron BC412) thin trigger scintillation counters (TSC) to provide fast trigger signals to other subdetectors.

Each module in the TOF subdetector consists of one TSC sandwiched between two trapezoidal TOF counters, with a 1.5 cm gap between the counters. Belle’s TOF and TSC scintillation counters are mounted directly onto fine-mesh-dynode photomultiplier tubes (FM-PMT). The time resolution is increased by not using light guides. Instead, each counter is wrapped in a single layer of 45 $\mu$m-thick polyvinyl film (Tedlar) to protect its surface and to contain the light within the counter. The 64 modules are individually mounted to the inside of the barrel Electromagnetic Calorimeter (ECL), and provide a polar angular coverage of $34^\circ < \theta < 120^\circ$. The TOF can measure particles with transverse momenta ranging from 0.28 to 1.2 GeV.

In addition to its function as part of Belle’s particle identification (PID) system, the TOF provides fast timing information to the trigger system. It indicates when to send gate signals for the analog-to-digital converters (ADC) and stop signals for the time-to-digital converters (TDC). The trigger system works most efficiently if the TOF signals are kept below 70 kHz; signals pile up in the trigger queue above this rate. The TSC layers reduce backgrounds due to photon conversion, and are used to keep the fast trigger rate below 70 kHz.

<table>
<thead>
<tr>
<th>Counter</th>
<th>Dimensions ($T \times W \times L$, cm)</th>
<th>$z$ coverage (cm)</th>
<th>$r$ (cm)</th>
<th>$\phi$ segments</th>
<th>No. of PMTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOF</td>
<td>4.0 $\times$ 6.0 $\times$ 255.0</td>
<td>-72.5 to +182.5</td>
<td>122.0</td>
<td>128</td>
<td>2</td>
</tr>
<tr>
<td>TSC</td>
<td>0.5 $\times$ 12.0 $\times$ 263.0</td>
<td>-80.5 to +182.5</td>
<td>117.5</td>
<td>64</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.3: Parameters of the TOF and TDC counters, adapted from [11].

25
3.7 Electromagnetic Calorimeter (ECL) and Extreme Forward Calorimeter (EFC)

The ECL and EFC are homogeneous shower counters that detect energy deposits in scintillation crystals. Electrons shed energy within the calorimeter material by producing photons through bremsstrahlung. The photon undergoes pair production (ECL and EFC detection of photons begins at this step), and the daughter electron and positron also release bremsstrahlung photons. This process, known as electromagnetic showering, continues until the end particles' energy is so low ($\sim 100$ MeV) that the primary mechanism of energy loss is ionisation rather than bremsstrahlung [48]. Electromagnetic showers tend to be highly collimated, and the number of secondary particles is proportional to the energy of the original incident electron or photon. The characteristic length scale of an electromagnetic calorimeter is the radiation length, which is the distance an electron travels in the time that its energy is reduced by a factor of $e$ due to radiation loss only [49]. The radiation length is given approximately by [48]

$$X_{rad} = 180A/Z^2 \text{g/cm}^2$$  \hspace{1cm} (3.5)

The energy loss in a material due to bremsstrahlung is inversely proportional to the material’s radiation length.

The ECL covers a polar angle of $17^\circ < \theta < 150^\circ$. The angular coverage is extended to $6.4^\circ < \theta < 11.5^\circ$ in the forward direction and $163.3^\circ < \theta < 171.2^\circ$ in the background direction by the EFC.

3.7.1 ECL

Belle primarily detects electromagnetic energy, in the form of photons and electrons, in a cylindrical array of 8736 CsI(Tl) crystals. The crystals are “tower” shapes (see Fig. 3.9): barrel crystals are typically $(55\text{mm})^2$ on the front face and $(65\text{mm})^2$ at the rear face, and endcap modules range from $(44.5\text{mm})^2$ to $(70.8\text{mm})^2$ on the front face and from $(54\text{mm})^2$ to $(82\text{mm})^2$ on the rear face. The size of the crystal faces vary slightly across the barrel; the exact size is chosen such that 80% of the total energy produced by a photon in the center of the crystal is contained within that crystal. The crystals are 30 cm long, which corresponds to 16.2 radiation lengths. The sides of each crystal are wrapped in a 200 $\mu$m-thick sheet of Goretex teflon to enhance the collection of scintillation light, and 25 $\mu$m-thick sheets of aluminium and mylar to provide light and electrical shielding, respectively (see Fig. 3.9).

The ECL has an inner radius of 1.25m and extends outward to 1.64m. The barrel is 3.0m long, and is asymmetrically distributed about the interaction point, ranging from $+2.0\text{m}$ in the $e^-$ direction to $-1.0\text{m}$ in the $e^+$ direction. Its total solid-angle coverage is 91% of $4\pi$.

The ECL weighs 43 tons.

<table>
<thead>
<tr>
<th>Item</th>
<th>$\theta$ coverage</th>
<th>$\theta$ seg.</th>
<th>$\phi$ seg.</th>
<th>No. of crystals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward end-cap</td>
<td>$12.4 - 31.4^\circ$</td>
<td>13</td>
<td>48 - 144</td>
<td>1152</td>
</tr>
<tr>
<td>Barrel</td>
<td>$32.2 - 128.7^\circ$</td>
<td>46</td>
<td>144</td>
<td>6624</td>
</tr>
<tr>
<td>Backward end-cap</td>
<td>$130.7 - 155.1^\circ$</td>
<td>10</td>
<td>64 - 144</td>
<td>960</td>
</tr>
</tbody>
</table>

Table 3.4: Geometrical parameters of the ECL, from [15].
3.7.2 EFC

The EFC is composed of radiation-hard bismuth germanate (Bi$_4$Ge$_3$O$_{12}$), or “BGO”, crystals, which have a radiation length of 1.12 cm and a refractive index of 2.15. Each crystal is a trapezoidal tower (see Fig 3.10) wrapped in a 100 μm-thick teflon sheet and a 20 μm-thick aluminised mylar sheet. There are two EFC subdetectors, one at each end of the ECL, each of which is composed of five concentric truncated cones of thirty-two crystals (see Fig 3.10). The nearest surfaces of the forward and rear EFCs are 60 and 43.5 cm from the interaction point, respectively. The crystals in the forward subdetector are 12 radiation lengths long and those in the rear are 11 radiation lengths long. The EFC’s inner bore radius is 6.5 cm. A 1mm-thick stainless steel housing attaches the inner bore to the beampipe.

Belle analyses do not tend to use information from the EFC, though its data is used for determining the luminosity collected by the detector. It also reduces backgrounds in the CDC.
3.8 The Superconducting Solenoid

A superconducting solenoid provides the 1.5 Tesla magnetic field required for charged-particle identification and momentum measurement. The magnetic field is approximately constant in a cylindrical volume 4.4m long and 3.4m in diameter. The solenoid itself is composed of a single layer of NbTi/Cu wire within a 99.99% pure Al stabiliser, and is contained within a cylindrical aluminium support structure (see Fig. 3.11). The coil is wrapped around the inner surface of the support structure, and is cooled by liquid helium flowing through a tube welded to the structure’s outer surface (see Fig. 3.11). The solenoid’s main properties are listed in Table 3.5.

The flux from the solenoid is returned through the iron of the KLM, which surrounds it (see Fig. 3.12).
Table 3.5: Main parameters of the solenoid coil, from [11].

<table>
<thead>
<tr>
<th>Items</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryostat</td>
<td></td>
</tr>
<tr>
<td>Radius: outer/inner</td>
<td>2.00 m/1.70 m</td>
</tr>
<tr>
<td>Central field</td>
<td>1.5 T</td>
</tr>
<tr>
<td>Total weight</td>
<td>23 tonnes</td>
</tr>
<tr>
<td>Effective cold mass</td>
<td>∼ 6 tonnes</td>
</tr>
<tr>
<td>Length</td>
<td>4.41 m</td>
</tr>
<tr>
<td>Coil</td>
<td></td>
</tr>
<tr>
<td>Effective radius</td>
<td>1.8 m</td>
</tr>
<tr>
<td>Length</td>
<td>3.92 m</td>
</tr>
<tr>
<td>Conductor dimensions</td>
<td>$3 \times 33 \text{ mm}^2$</td>
</tr>
<tr>
<td>Superconductor</td>
<td>NbTi/Cu</td>
</tr>
<tr>
<td>Stabilizer</td>
<td>99.99% aluminium</td>
</tr>
<tr>
<td>Nominal current</td>
<td>4400 A</td>
</tr>
<tr>
<td>Inductance</td>
<td>3.6 H</td>
</tr>
<tr>
<td>Stored energy</td>
<td>35 MJ</td>
</tr>
<tr>
<td>Typical charging time</td>
<td>0.5 h</td>
</tr>
<tr>
<td>Liquid helium cryogenics</td>
<td>Forced flow two phase</td>
</tr>
<tr>
<td>Cool down time</td>
<td>$\leq 6 \text{ day}$</td>
</tr>
<tr>
<td>Quench recovery time</td>
<td>$\leq 1 \text{ day}$</td>
</tr>
</tbody>
</table>

3.9 $K_{Long}^0$ and Muon Detector (KLM)

The KLM is constructed of alternating layers of charged particle detectors and 4.7cm-thick iron plates. There are fifteen detector layers and fourteen iron layers in the barrel and fourteen detector layers and fourteen iron layers in each endcap, encompassing a total angular coverage of $20^\circ - 155^\circ$. The iron both limits the effect of the solenoid’s magnetic field upon the accelerator by providing a return path for the majority of the magnetic flux, and dramatically reduces the energy of any charged pions that penetrate into the KLM. The detector layers are pairs of glass-electrode-resistive plate counters (RPCs) sandwiched between readout electronics (see Fig. 3.13, left).

The RPCs are essentially parallel-plate capacitors (see Fig. 3.13, right). The glass electrode plates in each counter are separated by a 1.9mm gap filled with a gas mixture of 62% HFC-134a, 30% argon, and 8% butane-silver (composed of 70% n-butane and 30% iso-butane). As an ionising particle (or, in the case of a $K_L$, a shower of ionising particles resulting from interactions within the iron layers) crosses the RPC, it leaves a “streamer” in the gas which causes a localised discharge in the plates. The copper strips of the readout pickups are arranged such that the strips above the RPC pair are perpendicular to those below the pair, enabling the position, though not the energy, of the traversing particle to be determined. The use of two RPCs provides redundancy and enhances the efficiency to $\geq 98\%$.

The barrel RPCs range in size from $2.2 \times 1.5$ to $2.2 \times 2.7$ m$^2$. The resistivity of the glass ($10^{12} - 10^{13} \Omega \text{cm}$ bulk resistivity) is large enough that a conductive coating is required in order to ensure a uniform potential across each plate’s surface. An india ink mixture (Koh-i-noor 3080F, 30% black and 70% white by weight) was chosen to reduce the surface resistivity to $10^6 - 10^7 \Omega$/square, which allows for uniform surface potential without shielding the streamer discharge from the pickups.
In terms of particle identification, $K_L$s are first detected in the KLM, where they leave “tracks” in concentric layers RPCs. Muon tracks begin in the SVD or CDC and continue through the KLM. They penetrate deeper into the KLM than charged hadrons, and tend to have fewer deflections.

### 3.10 Particle Identification

Particles of different types are identified by the tracks and energy deposits they leave in the subdetectors, as shown in Table 3.6.

<table>
<thead>
<tr>
<th>Final-state Particle</th>
<th>Tracks</th>
<th>Energy deposits</th>
</tr>
</thead>
<tbody>
<tr>
<td>photon ($\gamma$)</td>
<td>–</td>
<td>ECL, EFC</td>
</tr>
<tr>
<td>$e^{\pm}$</td>
<td>SVD, CDC, TOF</td>
<td>ECL, EFC, EFC</td>
</tr>
<tr>
<td>$\mu^{\pm}$</td>
<td>SVD, CDC</td>
<td>TOF, KLM</td>
</tr>
<tr>
<td>$\pi^{\pm}$</td>
<td>SVD, CDC</td>
<td>TOF, ACC</td>
</tr>
<tr>
<td>$K^{\pm}$</td>
<td>SVD, CDC</td>
<td>TOF</td>
</tr>
<tr>
<td>$K_L^0$</td>
<td>–</td>
<td>KLM</td>
</tr>
<tr>
<td>$p^{\pm}$</td>
<td>SVD, CDC</td>
<td>TOF</td>
</tr>
</tbody>
</table>

Neutral pions and etas (in the $\eta \rightarrow \gamma\gamma$ mode) are detected as pairs of photons, where the invariant mass of the pair is close to the nominal pion or eta mass. Invariant mass, $M_{inv}$, is defined as

$$M_{inv} = \sqrt{(\sum E_i)^2 - (\sum p_{x_i})^2 - (\sum p_{y_i})^2 - (\sum p_{z_i})^2} \quad (3.6)$$
where \( i = 1, 2 \) refers to the daughter photons. Other short-lived particles, such as \( K_0^0 \) and \( D \) mesons, are similarly reconstructed from their final products. Belle is not able to detect neutrinos; they must be reconstructed from an event’s missing four-momentum (see Section 4.6).

Particles are occasionally misidentified. For example, if a charged kaon passes through the aluminium wall of an ACC module, it may emit Čerenkov radiation and therefore be misidentified as a pion.

### 3.11 Comparison of the Belle, BaBar and CLEO Detectors

Belle, BaBar and CLEO are the only \( e^+e^- \) experiments to have studied \( B \) mesons in depth. Belle and BaBar are quite similar, while CLEO’s parameters and integrated luminosity are rather different (see Table 3.7). In particular, due to its symmetric beam energies, CLEO was not able to measure time-dependent \( CP \) violation.

#### Table 3.7: A comparison of the Belle, BaBar and CLEO detectors. Information for CLEO-c is not shown as it is not relevant to \( B \) physics. The listed integrated luminosities include data collected on the \( \Upsilon(4S) \) resonance only.

<table>
<thead>
<tr>
<th>Detector (Accelerator)</th>
<th>Beam Energies (GeV)</th>
<th>Int. Lum. (10^6 fb(^{-1}))</th>
<th>Ang. coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belle (KEK-B)</td>
<td>8.0 ( e^- ) on 3.5 ( e^+ )</td>
<td>630.23 [45]</td>
<td>17(^0) &lt; ( \theta ) &lt; 150(^0) (CDC, ECL) [11]</td>
</tr>
<tr>
<td>BaBar (PEP-II)</td>
<td>9.0 ( e^- ) on 3.1 ( e^+ )</td>
<td>148.55 [50]</td>
<td>15.8(^0) &lt; ( \theta ) &lt; 141.8(^0) (EMC) [17]</td>
</tr>
<tr>
<td>CLEO-II (CESR)</td>
<td>5.3 ( e^- ) on 5.3 ( e^+ )</td>
<td>3.2 [51]</td>
<td>95% of 4(\pi) [8]</td>
</tr>
<tr>
<td>CLEO-II.V and III (CESR)</td>
<td>5.3 ( e^- ) on 5.3 ( e^+ )</td>
<td>5.9 [51]</td>
<td>93% of 4(\pi) [8]</td>
</tr>
</tbody>
</table>

#### 3.11.1 BaBar

BaBar is Belle’s main competitor. The two B-Factories began operation in 1999. Belle and BaBar have similar designs (see Figure 3.14), as do their accelerators.

BaBar is situated at the interaction point of the PEP-II \( e^+e^- \) storage ring at the Stanford Linear Accelerator Center (SLAC) in Palo Alto, California. Both KEK-B and PEP-II are tuned to a cms energy of 10.58 GeV, the mass of the \( \Upsilon(4S) \).

#### 3.11.2 CLEO

The Cornell Electron Storage Ring (CESR) in Ithaca, New York, with its CLEO detector, is the progenitor of the B-Factories. The CLEO Collaboration first detected pairs of \( B \) mesons in 1980, and ceased running on the \( \Upsilon(4S) \) resonance in 2001. As the first experiment to study \( B \) mesons in-depth, it cleared the path for Belle and BaBar. Various incarnations of the detector have been used for various experiments since 1979. The current experiment, CLEO-c, operates at a cms energy ranging from 3 to 5 GeV in its study of charm physics.

CLEO only collected a small fraction of the \( B\bar{B} \) pairs that Belle and BaBar have, but, due to the detector’s long history, its behaviour is well understood and many systematic effects can be
accounted for. Furthermore, the detector’s excellent hermiticity (see Table 3.7), means that CLEO is particularly well-suited to the study of charmless semileptonic $B$ decays using the neutrino reconstruction technique (see Sec. 4.6). CLEO remains competitive in $b \to u\ell\nu$ measurements of some decay modes thanks to its high neutrino reconstruction efficiency and comparatively very low systematic errors.
Chapter 4

Analysis

This research measured the branching fractions of $B^+ \rightarrow \pi^0 \ell^+ \nu$ and $B^+ \rightarrow \eta \ell^+ \nu$ decays\(^1\) (and charge conjugates), using the missing momentum in each event to reconstruct the neutrino. The data was collected electronically by the Belle detector and reconstructed with a C++ program. Signal events were preferentially selected using ROOT\(^{[52]}\), and the signals were extracted using a fitting program, which is described in more detail below.

4.1 Data Set

This analysis was performed on 492 fb\(^{-1}\) of Belle data, corresponding to $(534.586 \pm 7.044)$ million $B\bar{B}$ events, recorded between 1999 and 2006. The signal and background events were estimated using Monte Carlo (MC) simulated data, which was generated using the EvtGen generator\(^{[53]}\) and processed through a simulated Belle detector using GEANT3.21\(^{[54]}\). A MC sample corresponding to approximately 810 fb\(^{-1}\) was used for backgrounds due to common $B$ decays and non-$B$ “continuum” events (described in Sec. 4.3.1). The $b \rightarrow u \ell \nu$ signal and crossfeed MC sample contained approximately 23 million events, including 802,255 $B \rightarrow \pi^0 \ell \nu$ events and 936,860 $B \rightarrow \eta \ell \nu$ events. In this sample, one $B$ decays to a common mode based on known branching fractions, and the other decays semileptonically to a $u \ell \nu$ mode. Separate LCSR models were used to generate the $\pi$\(^{[2]}\) and $\rho$ and $\omega$\(^{[6]}\) modes, and the ISGW2\(^{[3]}\) model was used for $\eta$, $\eta'$, $a_2$, $f_2$, $b_1$, $h_1$, $a_1$, $f_1$, $a_0$, and $f_0$ modes. MC events were reconstructed in the same manner as experimental data.

4.2 Event Reconstruction

Each $e^+e^-$ event detected within Belle is recorded electronically as occupancies in the silicon wafers making up the SVD ladders, hits in the CDC’s sense wires, electron showers in the photomultiplier tubes coupled to the aerogel tiles in the ACC, energy deposits in ECL crystals, and discharges in the KLM’s detection layers. The raw data is processed by the reconstruction farm computers in the Belle control room, within metres of the detector itself. After processing, the data is represented by a series of “tables” describing the properties of the particles in the event. Table 4.1 is an example of a table representing a generated event. The ID column is an index, with a parent particle’s index appearing before the indices of its daughters. The IDs of a given particle’s daughters always occur consecutively (e.g. 4–6, 17–19, etc.). The first particle listed in Table 4.1 (ID of 1) is the $\Upsilon(4S)$.

\(^{1}\)The decays are often abbreviated as $B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \eta \ell \nu$ in this thesis.
followed by its daughters, the generated $B_s$ (IDs 2–3), the $B_s$’s daughters (IDs 4–6 and 7–9), and so on. Each particle’s description includes its decay status ($\text{ISTHEP} = 1$ for final state particles and 2 for particles which decay), a “LUND ID” ($\text{IDEHEP}$) which identifies the type of particle using the Monte Carlo Numbering Scheme [20], the ID index of the particle’s mother ($\text{MOTHER}$, with $\text{MO}(1)$ and $\text{MO}(2)$ redundant for historical reasons), the range of IDs of the particle’s daughters in the case of non-final-state particles ($\text{DA}(1)$ through $\text{DA}(2)$), momentum in GeV ($\text{P}(1)$–$\text{P}(3)$), energy in GeV ($\text{P}(4)$), mass in GeV ($\text{P}(5)$), and production vertex position in mm ($\text{V}(1)$–$\text{V}(3)$) and time in mm/c ($\text{V}(4)$). Information about the data is stored in dozens of different tables, some of which relate the reconstructed particles to the original SVD, CDC, ACC, and KLM tracks and ECL cluster information. Researchers can choose which tables they wish to use in their analysis code.

Events in this study were reconstructed from tracks and clusters using a C++ analysis program which interfaced with the Belle Analysis Software Framework, or BASF. Photons were required to have energies greater than 50 MeV if they were detected in the barrel region of the ECL, and energies greater than 100 MeV if detected in the ECL endcaps. The also had to pass a set of reconstruction criteria designed to strain out low-quality and poorly resolved ECL showers.

Reconstructed electrons and muons (and their charge-conjugates) were identified by a likelihood which was based on information from the ECL, CDC and ACC, and from the KLM in the case of muons. In order to be considered an electron, a particle’s lab-frame momentum was required to be greater than 0.25 GeV, and its likelihood was required to be greater than 0.6. Muons were required to have lab-frame momentum greater than 0.5 GeV and a likelihood greater than 0.9. Charged tracks which did not pass the electron and muon identification criteria were further examined to see if they were consistent with protons ($\text{Prob}(P) = \frac{L_P}{L_P + L_\pi} > 0.6$) or charged kaons ($\text{Prob}(K) = \frac{L_K}{L_P + L_K} > 0.6$). If a particle passed both the proton and the kaon selections, it was taken to be a proton if ($\frac{L_P}{L_P + L_K} > \frac{L_K}{L_P + L_K}$), and a kaon otherwise. Those which failed both the proton and kaon checks were considered to be charged pions.

The analysis program reconstructed $\pi^0$, $\eta$ and $B$ meson candidates as described in Sections 4.4 and 4.7.
Table 4.1: A table representing a generated MC $B^- \rightarrow \pi^0\mu^-\bar{\nu}_\mu$ event. Each row represents a particle in the event: ID is an index number, starting with the $\Upsilon(4S)$ at ID=1; ISTHEP is the particle’s decay status (1 for a final state particle and 2 for a particle which decays); IDHEP identifies the type of particle using its LUND ID [20]; MOTHER indicates the index of the particle’s mother (with MO(1) and MO(2) redundant but retained in the code for historical reasons); the particle’s daughters, if any, have indices ranging from DA(1) to DA(2); P(1) through P(4) are the particle’s four-momentum in GeV, with its mass given in P(5), also in GeV; and V(1) through V(4) are the particle’s production vertex positions and time, in mm and mm/c respectively.

Table: GEN_HEPEVT (#entities = 133)

<table>
<thead>
<tr>
<th>ID</th>
<th>ISTHEP</th>
<th>IDHEP</th>
<th>MOTHER</th>
<th>MO(1)</th>
<th>MO(2)</th>
<th>DA(1)</th>
<th>DA(2)</th>
<th>P(1)</th>
<th>P(2)</th>
<th>P(3)</th>
<th>P(4)</th>
<th>P(5)</th>
<th>V(1)</th>
<th>V(2)</th>
<th>V(3)</th>
<th>V(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0.18</td>
<td>-0.00</td>
<td>4.49</td>
<td>11.50</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.97</td>
<td>-0.50</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>521</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>-0.06</td>
<td>0.08</td>
<td>1.91</td>
<td>5.61</td>
<td>5.28</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.97</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-521</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>9</td>
<td>0.24</td>
<td>-0.08</td>
<td>2.59</td>
<td>5.88</td>
<td>5.28</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.97</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>-423</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>10</td>
<td>11</td>
<td>-0.11</td>
<td>-0.58</td>
<td>1.06</td>
<td>2.35</td>
<td>2.01</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>413</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>13</td>
<td>-0.69</td>
<td>0.68</td>
<td>0.38</td>
<td>2.27</td>
<td>2.01</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>311</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>14</td>
<td>14</td>
<td>0.73</td>
<td>-0.02</td>
<td>0.47</td>
<td>1.00</td>
<td>0.50</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>111</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>15</td>
<td>16</td>
<td>-0.01</td>
<td>1.18</td>
<td>0.73</td>
<td>1.40</td>
<td>0.13</td>
<td>-0.05</td>
<td>-0.00</td>
<td>-7.91</td>
<td>-0.46</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>13</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>-0.67</td>
<td>-2.47</td>
<td>1.09</td>
<td>2.78</td>
<td>0.11</td>
<td>-0.05</td>
<td>-0.00</td>
<td>-7.91</td>
<td>-0.46</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>-14</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0.93</td>
<td>1.21</td>
<td>0.77</td>
<td>1.71</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.00</td>
<td>-7.91</td>
<td>-0.46</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>-421</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>17</td>
<td>19</td>
<td>-0.12</td>
<td>-0.55</td>
<td>1.03</td>
<td>2.20</td>
<td>1.86</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>111</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>20</td>
<td>21</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.14</td>
<td>0.13</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>421</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>22</td>
<td>23</td>
<td>-0.64</td>
<td>0.67</td>
<td>0.33</td>
<td>2.11</td>
<td>1.86</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>211</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>34</td>
<td>34</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.16</td>
<td>0.14</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>310</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>35</td>
<td>36</td>
<td>0.73</td>
<td>-0.02</td>
<td>0.47</td>
<td>1.00</td>
<td>0.50</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>22</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>37</td>
<td>38</td>
<td>-0.02</td>
<td>0.57</td>
<td>0.43</td>
<td>0.72</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.00</td>
<td>-7.91</td>
<td>-0.46</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>22</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>39</td>
<td>42</td>
<td>0.01</td>
<td>0.61</td>
<td>0.30</td>
<td>0.68</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.00</td>
<td>-7.91</td>
<td>-0.46</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>321</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>43</td>
<td>45</td>
<td>-0.29</td>
<td>-0.40</td>
<td>0.71</td>
<td>1.00</td>
<td>0.49</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-7.83</td>
<td>-0.39</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>-213</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>24</td>
<td>25</td>
<td>0.30</td>
<td>-0.24</td>
<td>0.41</td>
<td>0.97</td>
<td>0.77</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-7.83</td>
<td>-0.39</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>111</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>26</td>
<td>27</td>
<td>-0.14</td>
<td>0.10</td>
<td>-0.10</td>
<td>0.24</td>
<td>0.13</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-7.83</td>
<td>-0.39</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>22</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0.04</td>
<td>-0.08</td>
<td>0.00</td>
<td>0.09</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>22</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>46</td>
<td>47</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.03</td>
<td>0.06</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.00</td>
<td>-7.94</td>
<td>-0.47</td>
</tr>
</tbody>
</table>
Belle collects upwards of 500 gigabytes (GB) of data each day it operates; each 1 fb\(^{-1}\) corresponds to roughly 600-800 GB. It would take a prohibitively long time for analyses to run over data if the datasets were not reduced through preselection. The primary preselection used by Belle B-decay analyses is known as the HadronB “skim”. This analysis further reduced the dataset by using the lepton skim. These are now described.

### 4.3.1 HadronB

Approximately 75% of \(e^+e^-\) annihilations at the \(\Upsilon(4S)\) energy produce lighter quark pairs (\(u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}\)) rather than the \(\Upsilon(4S)\) bound \(b\bar{b}\) state. These are known as continuum events, as opposed to \(B\bar{B}\) events. Many \(\tau\)-pair, QED process (\(e^+e^- \rightarrow e^+e^-(\gamma)\)), and two-photon (\(\gamma\gamma \rightarrow q\bar{q}\)) events also occur (see Table 4.2). Since most (though by no means all) events of interest to the Belle Collaboration involve \(B\) mesons, a standard set of criteria, known within the Collaboration as “HadronB”, has been developed [21]. HadronB rejects almost all \(\tau\)-pair, QED, and two-photon events and over 95% of continuum events, while retaining 99.1% of “hadronic” \((B\bar{B})\) events. HadronB also removes nearly all “beam-gas” events, which result from interactions between the beam and residual gas in the beampipe.

The HadronB criteria are as follows:

**Track multiplicity** Each event is required to contain at least three charged tracks.

**Visible energy** The sum of the good track momenta and good photon energies in an event must satisfy \(E_{\text{vis}} \geq 0.2\sqrt{s}\), where \(\sqrt{s}\) is the energy in the centre-of-mass frame (10.58 GeV for on-resonance events).

**Conditional calorimeter energy sum** The sum of cluster energies in the ECL must satisfy \(E_{\text{sum}} > 0.18\sqrt{s}\) or \(M_{HJ} > 1.8\) GeV, where \(M_{HJ}\), the heavy jet mass, is essentially equal to the \(\tau\) invariant mass. This removes \(\tau\)-pair, beam-gas, QED, and two-photon events while allowing continuum events to pass if they are not consistent with a \(\tau\)-pair event.

**Conditional normalised heavy jet mass** The ratio \(M_{HJ}/E_{\text{vis}}\) must be greater than 0.25, or \(M_{HJ}\) must be greater than 1.8 GeV. \(^2\) This combination removes \(\tau\) pairs.

**Calorimeter cluster multiplicity and average cluster energy** There must be at least one ECL cluster, and the average cluster energy must be less than 1 GeV. The cluster multiplicity requirement removes QED events, beam-gas and two-photon events, which tend to be directed down the beampipe and are therefore outside the ECL coverage. The average energy requirement suppresses QED events.

**Momentum balance** The sum of the \(z\)-components (those along the beam direction) of the good track and good photon momenta must be within the range \(|P_z| < 0.5\sqrt{s}|\). This removes events that are very jetlike along the beam axis while retaining a high efficiency of hadronic events.

**Event primary vertex** The best possible vertex of all tracks whose transverse momentum is greater than 0.1 GeV is calculated. This vertex must lie within the \(z\) range \(|z| < 3.5\) cm, \(^3\)Note that if \(E_{\text{sum}} > 0.18\sqrt{s}\), the previous criterion will allow events with \(M_{HJ} < 1.8\) GeV. Similarly, if \(M_{HJ}/E_{\text{vis}} > 0.25\) GeV, this criterion will allow such events.
and within the radial range \( r < 1.5 \text{cm} \). The allowed radial distance is smaller than the beampipe’s radius, and therefore removes events in which spent beam particles collide with the beam pipe.

Table 4.2: Cross-sections and \texttt{HadronB} efficiencies for various processes in \( e^+e^- \) collisions at \( \sqrt{s} = 10.58 \text{ GeV} \), from [21]. \textit{QED} refers to Bhabha (\( e^+e^- \rightarrow e^+e^- \)) and radiative Bhabha (\( e^+e^- \rightarrow e^+e^-\gamma \)) processes.

<table>
<thead>
<tr>
<th>Process</th>
<th>( \sigma ) (nb)</th>
<th>Eff. (%)</th>
<th>( \sigma ) after \texttt{HadronB} (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b\bar{b} )</td>
<td>1.1</td>
<td>0.991</td>
<td>1.09</td>
</tr>
<tr>
<td>( q\bar{q}(q = u, d, s, c) )</td>
<td>3.3</td>
<td>0.795</td>
<td>2.62</td>
</tr>
<tr>
<td>( \tau^+\tau^- )</td>
<td>0.93</td>
<td>0.049</td>
<td>0.05</td>
</tr>
<tr>
<td>( QED(25.551^0 &lt; \theta &lt; 159.94^0) )</td>
<td>37.8</td>
<td>0.00002</td>
<td>0.001</td>
</tr>
<tr>
<td>( \gamma\gamma \rightarrow q\bar{q}(w &gt; 500 \text{ MeV}) )</td>
<td>11.1</td>
<td>0.004</td>
<td>0.04</td>
</tr>
</tbody>
</table>

### 4.3.2 lepton\_skim

\texttt{HadronB} dramatically decreases the size of the data to run over, but it would still take an unacceptably long time to run over the dataset. An additional set of requirements, known within the Collaboration as the \texttt{lepton\_skim}, was used to further reduce the data, with the intent of retaining as many signal events as possible while rejecting as many background events as possible. In order to pass the \texttt{lepton\_skim}, an event must pass \texttt{HadronB} and have one or more leptons with centre-of-mass frame (cms) momentum greater than 1.2 GeV. To be considered an electron or positron, a reconstructed particle must have an electron probability greater than 0.5; muon candidates must have a muon probability greater than 0.8. MC events must pass an additional selection which mimics Belle’s online trigger. \texttt{lepton\_skim} efficiencies can be seen in Table 4.3.

Table 4.3: \texttt{lepton\_skim} efficiencies for real data and various MC event types. Values, other than those for \( b \rightarrow u\ell\nu \), are from [22].

<table>
<thead>
<tr>
<th>Event type</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{HadronB} (data)</td>
<td>10.5</td>
</tr>
<tr>
<td>( b \rightarrow u\ell\nu ) MC</td>
<td>66</td>
</tr>
<tr>
<td>mixed/charged MC</td>
<td>21</td>
</tr>
<tr>
<td>uds MC</td>
<td>3.5</td>
</tr>
<tr>
<td>charm MC</td>
<td>8</td>
</tr>
</tbody>
</table>

### 4.4 \( \pi^0 \) and \( \eta \) Reconstruction

The \( \pi^0 \)s and \( \eta \)s used in this research were reconstructed in the analysis program; pre-reconstructed (by BASF) \( \pi^0 \)s were not used. Only the \( \pi^0 \rightarrow \gamma\gamma \) and \( \eta \rightarrow \gamma\gamma \) decay chains were reconstructed. \( \pi^0 \)s decay to \( \gamma\gamma \) with a branching fraction of \( (98.798 \pm 0.032)\% \), and the branching fraction of \( \eta \rightarrow \gamma\gamma \) is \( (39.38 \pm 0.26)\% \) [20]. \( B \rightarrow \pi^0\ell\nu \) has now been measured by other experiments and can be used as a control: the sources of experimental systematic uncertainty for \( B \rightarrow \eta\ell\nu \) (\( \eta \rightarrow \gamma\gamma \)) should be similar.
Pairs of photons in each event were combined, and in the cases where the invariant mass of the pair, defined as

\[ M_{\text{inv}} = \sqrt{(E_{\gamma_1} + E_{\gamma_2})^2 - (\vec{p}_{\gamma_1} + \vec{p}_{\gamma_2})^2} \]  

fell within the range \(119 < m_{\gamma\gamma} < 151\) MeV (510 < \(m_{\gamma\gamma}\) < 570 MeV), the pair was considered to be a \(\pi^0\) (\(\eta\)). The \(\pi^0\)s and \(\eta\)s were then “mass-constrained”: the daughter gammas’ four-momenta were slightly modified such that the reconstructed meson’s invariant mass was equal to its nominal mass (134.976 MeV for \(\pi^0\) and 547.300 MeV for \(\eta\)).

The \(\pi^0 \to \gamma\gamma\) and \(\eta \to \gamma\gamma\) modes are characterised by a low signal-to-background (S/B) ratio due to a large combinatorial background. Belle records a vast number of photons in each event, and with so many photons, there is a high probability that random pairs can fall within the \(\pi^0\) and \(\eta\) mass windows as defined above (see Figure 4.1). The combinatorial background can be reduced, and the S/B improved, by reconstructing \(\pi^0\)s and \(\eta\)s using only photons that exceed an energy threshold (see Figures 4.2 and 4.3).

![Figure 4.1: The invariant mass distribution, in units of GeV, of photon pairs in \(B \to \eta\ell\nu\) signal MC. The narrow peak at 0.135 GeV corresponds to \(\pi^0\)s, and the smaller, broader peak at 0.547 GeV corresponds to \(\eta\)s. In this figure, no restrictions are made on the photons’ energies.](image)

The S/B improvement comes at a price: as the threshold rises, the reconstruction efficiency falls. A figure of merit, \(S/\sqrt{S + B}\), was found to be optimised with a minimum photon energy threshold of 50 MeV for \(\pi^0\)s and 400 MeV for \(\eta\)s (see Table 4.4). By using the nearby thresholds of 100 MeV for \(\pi^0\)s and 350 MeV for \(\eta\)s, the \(\pi^0\)’s S/B was greatly improved, and additional efficiency was gained for the \(\eta\).
Table 4.4: Number of $\pi^0$s and $\eta$s, and their respective combinatorial backgrounds and figures of merit, for a range of photon thresholds. For the $\pi^0$, signal errors from the fit are typically 3% for photon thresholds up to 300 MeV and rise to 10% for the maximum photon threshold, while background errors from the fit are typically 1% for photon thresholds up to 150 MeV and rise to 9% for the maximum photon threshold. The $\eta$ signal exhibits the opposite trend, with signal errors from the fit typically 7-8% for photon thresholds up to 100 MeV and dropping to 4% for the maximum photon threshold; background errors from the fit are typically 1-3% for the entire range. All numbers were extracted using a fit (ROOFIT) [23]. The minimum photon energy thresholds used in the analysis were 100 MeV for $\pi^0$s and 350 MeV for $\eta$s.

<table>
<thead>
<tr>
<th>Min. photon threshold (MeV)</th>
<th>$\pi^0$ Num. Signal $(10^3)$</th>
<th>$\eta$ Num. Signal $(10^3)$</th>
<th>$S/\sqrt{S+B} \pi^0$</th>
<th>$S/\sqrt{S+B} \eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.77</td>
<td>10.08</td>
<td>34.8</td>
<td>23.3</td>
</tr>
<tr>
<td>50</td>
<td>8.12</td>
<td>10.07</td>
<td>44.8</td>
<td>22.7</td>
</tr>
<tr>
<td><strong>100</strong></td>
<td><strong>4.57</strong></td>
<td><strong>8.57</strong></td>
<td><strong>44.1</strong></td>
<td><strong>25.5</strong></td>
</tr>
<tr>
<td>150</td>
<td>2.95</td>
<td>6.33</td>
<td>41.4</td>
<td>29.8</td>
</tr>
<tr>
<td>200</td>
<td>1.80</td>
<td>4.23</td>
<td>34.8</td>
<td>35.0</td>
</tr>
<tr>
<td>250</td>
<td>1.16</td>
<td>2.58</td>
<td>29.1</td>
<td>39.6</td>
</tr>
<tr>
<td>300</td>
<td>0.74</td>
<td>1.41</td>
<td>23.8</td>
<td>50.3</td>
</tr>
<tr>
<td><strong>350</strong></td>
<td><strong>0.53</strong></td>
<td><strong>0.93</strong></td>
<td><strong>20.5</strong></td>
<td><strong>53.5</strong></td>
</tr>
<tr>
<td>400</td>
<td>0.33</td>
<td>0.65</td>
<td>16.3</td>
<td>54.5</td>
</tr>
<tr>
<td>450</td>
<td>0.23</td>
<td>0.49</td>
<td>13.6</td>
<td>54.1</td>
</tr>
<tr>
<td>500</td>
<td>0.15</td>
<td>0.40</td>
<td>10.7</td>
<td>51.3</td>
</tr>
</tbody>
</table>
Figure 4.2: The invariant mass distribution, in units of GeV, of photon pairs in $B \rightarrow \eta \ell \nu$ signal MC, for photon energy thresholds ranging from 50 to 300 MeV. The S/B of the $\pi^0$ and $\eta$ peaks begin to increase as the threshold is raised. The $\pi^0$ efficiency, however, decreases rapidly after $E_\gamma > 200$ MeV.
Figure 4.3: The invariant mass distribution, in units of GeV, of photon pairs in $B \to \eta \ell \nu$ signal MC, for photon energy thresholds ranging from 350 to 500 MeV.
4.5 Methods for Studying Exclusive $B \to X_u \ell \nu$ Decays

The Belle, BaBar and CLEO Collaborations are currently involved in studying semileptonic $b \to u$ transitions. These Collaborations make use of the “clean” (few particle) environments of their $e^+e^-$ colliders; hadronic collisions such as those in the upcoming LHC result in too many particles to measure these decays. Most analyses combine charged and neutral modes using isospin constraints; this analysis examines the $B \to \pi^0 \ell \nu$ (and, separately, $B \to \eta \ell \nu$) only.

The primary difficulty in reconstructing $B \to X_u \ell \nu$ events is one of statistics: roughly one in ten thousand $\Upsilon(4S)$ decays results in $B \to \pi^0 \ell \nu$. The neutrino compounds the problem. Neutrinos can be detected in large specialised detectors which are often located deep in mines, but not in multipurpose spectrometers such as Belle, BaBar and CLEO. Furthermore, exclusive charmless semileptonic $B$ decays do not produce a clean signal such as the $M_{K\pi}$ spectrum in the $B \to D\ell\nu$ decay.

Three techniques have traditionally been used to study exclusive semileptonic $b \to u$ transitions with data from $e^+e^-$ colliders:

**Neutrino reconstruction** This is the original technique. It was pioneered by the CLEO Collaboration in 1996 [55] and has since been used by CLEO [56] [8], BaBar [57] [58] [42], Belle (for $B \to \omega \ell \nu$) [59], and this analysis. The neutrino’s four-momentum is reconstructed by comparing the measured energy and momentum in an event with the input energy of the electron and positron beams. Neutrino reconstruction is described in detail in the following section. This technique is also known as the “untagged” method.

**Semileptonic tagging** This method uses the kinematics of events in which both $B_s$ decay semileptonically, one (the signal side) to a $\pi$ or a $\rho$, and the other (the tag side) to a $D$ or a $D^*$. No attempt is made to reconstruct the neutrinos. Signal events are identified by relative angles between the observed particles. This produces a higher signal-to-background (S/B) ratio than neutrino reconstruction, but at the price of a larger statistical uncertainty. BaBar [60] [61] [62] and Belle [63] [64] have both presented results using semileptonic tagging.

**Full reconstruction** This is the up-and-coming technique for measuring charmless semileptonic decays. One $B$ is reconstructed cleanly in a common mode such as $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\rho^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, $B^- \to D^{(*)0}\pi^-$, or $B^0 \to D^{(*)0}\pi^-$, if the other $B$ decayed semileptonically to a light meson, that meson and the charged lepton will be the only other measured particles in the event. This technique requires a very large data sample to produce a statistically significant signal, but the signal will have a very high S/B. The Belle [65] and BaBar [60] [66] data samples are just starting to be large enough for this technique to be useful for measuring charmless semileptonic decays.

The neutrino reconstruction technique, used in this analysis, results in a low statistical error, but can be limited by systematics affecting the reconstructed neutrino. Full reconstruction provides a high signal-to-background and correspondingly small systematic error, but requires a large amount of data to tame its statistical error. Semileptonic tagging falls between the two: its systematic error is lower than that achieved with neutrino reconstruction but higher than that from full reconstruction, and its statistical error is lower than that from full reconstruction but higher than that from neutrino reconstruction.

Current results from Belle, BaBar and CLEO, organised by technique, can be seen in Table 4.5. There are three subtle but noteworthy things in this table. The first is that the statistical errors on
the neutrino reconstruction measurements are considerably smaller than those on the semileptonic (SL) and full-reconstruction tagged measurements. Second, none of these results show a pure $B \rightarrow \pi^0 \ell \nu$ measurement; at best, $\pi^0$ results are determined from a $\pi^\pm$ or combined $\pi^+ - \pi^0$ measurement, and use isospin symmetry to give a $\pi^0$ result. The present analysis is unique in that it looks specifically at $B \rightarrow \pi^0 \ell \nu$ decays. Third, neutrino reconstruction is the only method that has been used to obtain a $B \rightarrow \eta \ell \nu$ branching fraction.

Table 4.5: Current $B \rightarrow \pi \ell \nu$ and $B \rightarrow \eta \ell \nu$ branching fraction (BF) measurements, obtained using neutrino reconstruction, semileptonic tagging, and full-reconstruction tagging. Charge conjugation is implied in all cases. The first, second and third errors are statistical, experimental systematic, and due to model uncertainty, respectively. Studies that used isospin assumptions to combine $\pi^+$ and $\pi^0$ modes include a factor of $2\pi^\pm$ to account for the difference between the $B^0$ and $B^+$ lifetimes. Note the low statistical error resulting from neutrino reconstruction. $B^+ \rightarrow \eta \ell^+ \nu$ values and Belle’s full-reconstruction values are from [24]; all others are from [25]. All values are exactly as quoted by the Heavy Flavor Averaging Group (HFAG).

<table>
<thead>
<tr>
<th>Mode</th>
<th>BF ($\times 10^4$)</th>
<th>BF ($\times 10^4$)</th>
<th>BF ($\times 10^4$)</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>($q^2 &lt; 16 \text{ GeV}^2$)</td>
<td>($q^2 &gt; 16 \text{ GeV}^2$)</td>
<td></td>
</tr>
<tr>
<td>Neutrino Recon.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- l^+ \nu$</td>
<td>$1.33 \pm 0.18 \pm 0.11$</td>
<td>$1.08 \pm 0.16 \pm 0.10$</td>
<td>$0.25 \pm 0.09 \pm 0.05$</td>
<td>CLEO</td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- l^+ \nu$</td>
<td>$1.46 \pm 0.07 \pm 0.08$</td>
<td>$1.09 \pm 0.06 \pm 0.07$</td>
<td>$0.38 \pm 0.04 \pm 0.03$</td>
<td>BaBar</td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>$0.84 \pm 0.31 \pm 0.16 \pm 0.09$</td>
<td>CLEO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>$0.84 \pm 0.27 \pm 0.21$</td>
<td>BaBar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL tag</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^0 l^+ \nu$</td>
<td>$0.77 \pm 0.14 \pm 0.08$</td>
<td>$0.57 \pm 0.12 \pm 0.06$</td>
<td>$0.20 \pm 0.08 \pm 0.02$</td>
<td>Belle</td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^0 l^+ \nu$</td>
<td>$0.73 \pm 0.18 \pm 0.08$</td>
<td>$0.63 \pm 0.16 \pm 0.06$</td>
<td>$0.10 \pm 0.12 \pm 0.04$</td>
<td>BaBar</td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- l^+ \nu$</td>
<td>$1.38 \pm 0.19 \pm 0.15$</td>
<td>$1.02 \pm 0.16 \pm 0.11$</td>
<td>$0.36 \pm 0.10 \pm 0.04$</td>
<td>Belle</td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- l^+ \nu$</td>
<td>$1.12 \pm 0.25 \pm 0.10$</td>
<td>$0.83 \pm 0.22 \pm 0.08$</td>
<td>$0.29 \pm 0.15 \pm 0.04$</td>
<td>BaBar</td>
</tr>
<tr>
<td>Full-recon. tag</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^0 \ell^+ \nu \times 2\frac{2\pi^0}{\pi^+}$</td>
<td>$1.60 \pm 0.32 \pm 0.11$</td>
<td>Belle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^0 \ell^+ \nu \times 2\frac{2\pi^0}{\pi^+}$</td>
<td>$1.53 \pm 0.41 \pm 0.21$</td>
<td>$1.05 \pm 0.35 \pm 0.15$</td>
<td>$0.49 \pm 0.22 \pm 0.11$</td>
<td>BaBar</td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^0 \ell^+ \nu \times 2\frac{2\pi^0}{\pi^+}$</td>
<td>$0.82 \pm 0.22 \pm 0.11$</td>
<td>$0.56 \pm 0.19 \pm 0.08$</td>
<td>$0.26 \pm 0.12 \pm 0.06$</td>
<td>BaBar</td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- \ell^+ \nu$</td>
<td>$1.49 \pm 0.26 \pm 0.06$</td>
<td>Belle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^0 \rightarrow \pi^- \ell^+ \nu$</td>
<td>$1.07 \pm 0.27 \pm 0.15$</td>
<td>$0.42 \pm 0.18 \pm 0.05$</td>
<td>$0.65 \pm 0.20 \pm 0.13$</td>
<td>BaBar</td>
</tr>
</tbody>
</table>

Belle has reported measurements of semileptonic $B$ decays to charged and neutral pions using semileptonic and full-reconstruction tagging, but not using neutrino reconstruction. The results presented in this thesis are the first Belle measurements of the $B \rightarrow \pi^0 \ell \nu$ mode using the neutrino reconstruction technique, and Belle’s first $B \rightarrow \eta \ell \nu$ measurement.

### 4.6 The Art of Neutrino Reconstruction I: the Original CLEO Method

The neutrino reconstruction technique used in this analysis was originally developed by the CLEO Collaboration [55]. It requires events that have one and only one charged lepton; conservation of lepton number means the event must then have involved a single neutrino.

The assumption is made that all particles (other than the neutrino) in the event are detected and
identified correctly. The measured energy and three-momentum is summed and compared to the energy and three-momentum of the electron and positron that initiated the event. The event’s missing three-momentum

$$\vec{p}_{\text{miss}} = - \sum_i \vec{p}_i$$

(4.2)

and its missing energy

$$E_{\text{miss}} = 2E_{\text{beam}} - \sum_i E_i$$

(4.3)

in the cms are assumed to come from the undetectable, massless neutrino. Here the sum is over all detected particles in the event\(^3\). \(E_{\text{beam}}\) is the (equal) cms energy of the electron and positron beams.

Belle’s momentum resolution is better than its energy resolution (see Fig 4.4). In accordance with this, the neutrino’s four-momentum is assigned as:

$$P_\nu = P_{\text{miss}} = (|\vec{p}_{\text{miss}}|, \vec{p}_{\text{miss}})$$

(4.4)

\(\text{Figure 4.4: Missing momentum (}|\vec{p}_{\text{miss}}|; \text{left) and missing energy (}E_{\text{miss}}; \text{right) in generated }B \rightarrow \pi^0\ell\nu\text{ signal MC. Both histograms are in units of GeV. The central peak, and therefore the resolution, of the missing momentum distribution is narrower than that of the missing energy distribution. No cuts have been applied to these distributions.}\)

Restrictions must be placed on events in order to correctly reconstruct a neutrino. These are not merely designed to strain out the backgrounds; their primary purpose is to remove badly-reconstructed events in order to improve the neutrino’s resolution.

- **One lepton** The event must contain one and only one charged lepton. In this analysis, that lepton could be an electron (or positron) or a muon (or anti-muon). The presence of additional leptons would likely imply a multiple-neutrino event, which would lead to an incorrectly-reconstructed neutrino.

\(^3\pi^0\)s and \(\eta\)s were not included in this sum as they themselves cannot be “detected”; their daughter photons were used instead.
• **cos θν** Belle covers a solid angle ranging from 17° to 150°. Particles whose momenta carry them outside this region are not detected. The momentum and energy of such particles can contribute to the missing four-momentum and hence to the four-momentum attributed to the reconstructed neutrino. This effect was suppressed in this analysis by rejecting events in which the angle of the neutrino’s momentum with respect to the Υ(4S)’s boost vector fell outside the range (−0.87 < cos θν < 0.96), corresponding to 17−150°.

• **Net charge** The net electrical charge was required to be within the range |Qnet| ≤ 1e, again to remove events with undetected particles (other than the neutrino).

• **Missing mass** The square of the event’s missing invariant mass, defined as

\[ M_{miss}^2 = E_{miss}^2 - \vec{p}_{miss}^2 \]  

was required to be small (−2.5 GeV² < M_{miss}² < 4 GeV²), again to limit events with undetected particles.

If an event satisfied these conditions, it was retained for further analysis.

### 4.7 B Meson Reconstruction

After the neutrino has been reconstructed, work can begin on the B. The four-momenta of the charged lepton (e±/µ±), the neutrino, and each π⁰ (or η) in the event are combined into B meson candidates, in the same way as pairs of photons were combined to form π⁰s and ηs. A π⁰(η)ℓν combination was preselected as a B candidate if its invariant mass fell within the range 4 GeV < M_{inv} < 6 GeV.

Two further variables are used to determine if the B meson candidate is likely to be a reconstruction of an actual B meson. The first is the beam-constrained mass of the B candidate (see Fig. 4.18), given by

\[ M_{bc} = \sqrt{E_{beam}^* - |\vec{p}_{\pi^0(\eta)} + \vec{p}_\ell + \vec{p}_\nu|^2} \]  

where \( E_{beam}^* \) is the energy of the electron and positron beams, and all quantities are measured in the cms. B-Factory analyses typically use the beam-constrained mass rather than the B’s invariant mass, because the former takes into account information known from experimental conditions rather than only information from reconstructed particles: in the cms, each B carries an amount of energy equal to \( E_{beam}^* \). Ideally, \( M_{bc} \) would be the nominal B meson mass of 5.279 GeV.

The second variable is the event’s missing energy (see Fig. 4.19):

\[ \Delta E = (E_{\pi^0(\eta)}^* + E_\ell^* + E_\nu^*) - E_{beam}^* \]  

If the B candidate’s π⁰ (or η), charged lepton, and neutrino are actually the decay products of a B meson, their energies should sum up to the cms beam energy and \( \Delta E \) should equal zero. The actual value is smeared by the experimental resolution.

Due to combinatorial effects, an event may have multiple B meson candidates. Various criteria were examined for selecting the “best B candidate” in such events: the angle between the π⁰’s or
η’s daughter gammas, the π^0’s or η’s mass (without constraining), the π^0’s or η’s momentum, the reconstructed B’s momentum angle with respect to Υ(4S) boost vector, and the reconstructed B’s beam-constrained mass. An MC study showed that choosing the highest momentum π^0 or η resulted in the highest signal purity, and was used as the “best B” selection criteria.

4.8 The Art of Neutrino Reconstruction II: the Modified CLEO Method

In 2003, the CLEO Collaboration published an analysis that used a modified neutrino reconstructed method, which improves the neutrino resolution (See Fig. 4.5). This method is described in Ref. [8]. In essence, two additional kinematic constraints are introduced for the neutrino after a “best B” meson candidate has been selected. The neutrino’s momentum vector is “stretched” (or “compressed”) along its flight direction so that the energy difference, ΔE, is identically 0. The neutrino’s momentum vector is then rotated through the minimum angle such that $M_{bc} = M_B = 5.279$ GeV.

The ΔE constraint can be applied to all events with a “best B candidate”. The same cannot be said for the $M_{bc}$ rotation. In cases where the cosine of the angle between the three-momentum of the B and that of the combined $Y = \pi^0(\eta) + \ell$ system (see Fig. 4.14), given by

$$\cos \theta_{BY} = \frac{2E_BE_Y - M_{bc}^2 - M_Y^2}{2 \lvert \vec{p}_B \rvert \lvert \vec{p}_Y \rvert}$$

is not physical, i.e. where $\lvert \cos \theta_{BY} \rvert > 1$, no angle of rotation can give $M_{bc} = 5.279$ GeV. In these cases, the “stretched” neutrino was used.

It should be noted that the modified neutrino, rotated or not, was used for calculations such as $q^2$, but the original neutrino was used to calculate ΔE and $M_{bc}$, which together define the plane in which the signal region lies.

![Figure 4.5: Neutrino energy resolution, in units of GeV, calculated using the classic neutrino reconstruction (light grey histogram) and the modified CLEO technique (dark grey histogram). The modified CLEO technique dramatically improves the neutrino resolution.](image-url)
4.9 Background Composition

The main backgrounds to $B \to \pi^0(\eta)\ell\nu$ are continuum ($q\bar{q}$) events, charmed semileptonic $B$ decays ($B \to X_c\ell\nu$), and other charmless semileptonic $B$ decays ($b \to u$ crossfeed).

4.9.1 Continuum Events

The main background to $B \to \pi^0(\eta)\ell\nu$ events is the $q\bar{q}$ continuum. These are events in which no $\Upsilon(4S)$, and therefore no $B$ mesons (nor $b$ quarks), are created. Instead, the interaction between the electron and positron results in light ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$) or charmed ($c\bar{c}$) quark pairs. Charm is produced slightly less often than the light flavours\(^4\). The difference in production cross-section and in the dynamics involved in the considerably heavier $c$ is reflected in the separate $c$ and ($u, d, s$) MC samples.

Continuum events result in markedly different lepton momentum spectra and event topologies as compared to $B$ meson events. The end products of $B$ decays tend to be distributed isotropically, with one $B$ decaying in each hemisphere, while the continuum events exhibit quark jets.

Many continuum events are strained out by the HadronB selections mentioned in Section 4.3.1, but a very large number of these events remain.

4.9.2 Common $B$ Decays

Of the many common $B\bar{B}$ decay modes, it is charmed semileptonic decays that dominate the $B\bar{B}$ background in this analysis. These events occur approximately fifty times more frequently than charmless semileptonic $B$ decays. Like semileptonic decays to light hadrons, charmed semileptonic decays contain one charged lepton, allowing them to pass lepton skim and neutrino reconstruction criteria. Charmed mesons, however, have a very large mass compared to the $\pi^0$ or the $\eta$. Much of the four-momentum in a charmed semileptonic decay is taken up by the hadronic daughter’s mass, resulting in a lepton momentum spectrum with a lower mean as compared to signal events. The MC for this background is divided into mixed ($B^0\bar{B}^0$, with allowances made for mixing between the neutral states) and charged ($B^+B^-$) samples.

4.9.3 $b \to u$ Crossfeed

Crossfeed from other charmless semileptonic decays is the most difficult background to suppress, as these events are kinematically very similar to signal events. Additionally, neutral pions are amongst the decay products of many of these mesons (see Table 4.6). Such a neutral pion will have less energy than a true $B \to \pi^0\ell\nu$ signal pion, but if the neutrino’s energy is overestimated in reconstruction, the invariant mass of the pion-lepton-neutrino combination can closely mimic a signal $B$ candidate.

\(^4\)Light quark pairs are produced 1.6 times more often than charmed quark pairs at the $\Upsilon(4S)$ energy ($\sqrt{s} = 10.58$ GeV).
Table 4.6: Primary decay modes of some of the light mesons in the $b \to u\ell\nu$ MC. For charged mesons, the decay mode(s) listed represent those of the positively charged meson, with charge-conjugation implied in all cases. All values are from [20].

<table>
<thead>
<tr>
<th>Meson</th>
<th>Mass (MeV)</th>
<th>Decay mode(s)</th>
<th>Branching fraction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi^\pm$</td>
<td>139.6</td>
<td>$\mu^+\nu_\mu$</td>
<td>($99.98770 \pm 0.00004$)</td>
</tr>
<tr>
<td>$\pi^0$</td>
<td>135.0</td>
<td>$2\gamma$</td>
<td>($98.798 \pm 0.032$)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>547.5</td>
<td>$3\pi^0$</td>
<td>($32.51 \pm 0.28$)</td>
</tr>
<tr>
<td>$\eta'$</td>
<td>957.8</td>
<td>$\pi^+\pi^-\eta$</td>
<td>($44.5 \pm 1.4$)</td>
</tr>
<tr>
<td>$\rho^\pm$</td>
<td>763-775</td>
<td>$\pi^\pm\pi^0$</td>
<td>$\sim 100$</td>
</tr>
<tr>
<td>$\rho^0$</td>
<td>769-776</td>
<td>$\pi^+\pi^-$</td>
<td>$\sim 100$</td>
</tr>
<tr>
<td>$\omega$</td>
<td>782.7</td>
<td>$\pi^+\pi^-\pi^0$</td>
<td>($89.1 \pm 0.7$)</td>
</tr>
<tr>
<td>$\pi^0\gamma$</td>
<td></td>
<td>($8.90^{+0.27}_{-0.23}$)</td>
<td></td>
</tr>
<tr>
<td>$\pi^+\pi^-\eta$</td>
<td></td>
<td>($44.5 \pm 1.4$)</td>
<td></td>
</tr>
<tr>
<td>$\eta^*$</td>
<td>957.8</td>
<td>$\pi^+\pi^-\eta$</td>
<td>($20.8 \pm 1.2$)</td>
</tr>
</tbody>
</table>

4.10 Background Suppression

Now that the cacophonous environment that is the expected Belle data has been described, let us explore the phase-space restrictions by which we separate the signal needle from the background haystack. While HadronB (see Sec. 4.3.1) is optimised for hadronic events in general, it is not optimised for $B \to \pi^0\ell\nu$ and $B \to \eta\ell\nu$ events. The lepton skim (see Sec. 4.3.2) reduces the backgrounds further, but not enough for the final analysis. Backgrounds can be tamed by a careful choice of selection criteria, or “analysis cuts”.

4.10.1 Analysis Cuts

Analysis cuts were introduced one at a time. For each variable to be cut upon, the signal and background MC spectra were examined, and a preliminary threshold was chosen. The selected threshold was a balance between retaining signal efficiency and reducing the backgrounds as much as possible. Cuts were applied until the signal-to-background ratio (S/B) had been increased dramatically, but each background sample still retained enough events for shape discrimination by the signal yield fitting program, which is discussed in the following chapter.

The values listed in this chapter are for the neutral pion case. The $\eta$ analysis, described in Chapter 6, followed the same procedure.

Cuts designed to improve neutrino reconstruction have already been described (see Sec. 4.6). These were:

- **One lepton** Only single-charged-lepton events were retained for analysis.
- **$\cos \theta_\nu$** The angle of the neutrino, calculated in the lab frame, must fall within the range $-0.87 < \cos \theta_\nu < 0.96$, corresponding to 17 – 150° from the beam axis (see Fig. 4.9). This is the acceptance range of the Belle detector. Events whose neutrino’s momentum was closer
to the beam direction were rejected as it was likely that one or more particles had escaped
detection, which would adversely affect the neutrino reconstruction.

- **Net charge** The net electrical charge was required to be within the range \(|Q_{\text{net}}| \leq 1e\). Particles are occasionally misidentified. Furthermore, Belle’s hermiticity isn’t perfect; particles are sometimes lost down the beampipe. Both of these situations can lead to events with non-zero net electrical charge. The increase in efficiency achieved by allowing events with net charge = ±e in addition to those with net charge = 0, however, outweighs the loss in neutrino resolution (see Fig. 4.10).

- **\(M^2_{\text{miss}}\)** The square of the event’s missing invariant mass, \(M^2_{\text{miss}} = E^2_{\text{miss}} - \vec{p}^2_{\text{miss}}\), was required to be relatively small. Ideally this quantity should equal zero GeV\(^2\), but, due to the detector’s resolution, values ranging from -2.5 to 4 GeV\(^2\) were included to increase the signal efficiency (see Fig. 4.11).

Other cuts were introduced in order to selectively remove background events and increase the signal-to-background (S/B):

- **\(p^*_\ell\)** The cms lepton momentum \((p^*_\ell)\) spectrum of charmless semileptonic \(B\) decays has a larger mean than that of the \(B\bar{B}\) and continuum backgrounds. This separation provides a good means of removing backgrounds. However, \(p^*_\ell\) is closely related to \(q^2\), the square of the mass of the virtual \(W\) that decays into the charged lepton and the neutrino (see Eq. 2.11 and Fig. 4.6), and one goal of this analysis is to extract the shape of the signal’s \(q^2\) spectrum. The \textit{lepton skim} cut of 1.2 GeV removes very little signal but a vast number of continuum events and a fair amount of the \(B\bar{B}\) background. The effect of increasing the cut value was investigated (see Sec. 4.10.2). It was decided to leave it at 1.2 GeV to retain as much of the \(q^2\) spectrum as possible. The efficiency of this cut with respect to the \textit{lepton skim}, listed in Table 4.9, differs slightly from 100% due to small differences in particle identification between the skim and the event reconstruction program.

![Figure 4.6: cms lepton momentum vs. \(q^2\) for \(B \to \pi^0 \ell\nu\) signal MC. The cms lepton momentum and \(q^2\) axes are in units of GeV and GeV\(^2\), respectively. Note the correlation between cms lepton momentum and \(q^2\).](image)

- **\(R_2\)** This is the ratio of the second to the zeroth Fox-Wolfram moments [67]. Its power is in its shape discrimination: it tends to have a low value for isotropically distributed events (such as the majority of \(B\) events) and a high value for jetty events (such as most continuum events). Limiting this variable to low values is very effective in suppressing continuum events (see Fig. 4.13).
• **cos θ<sub>BY</sub>**  The cosine of the angle between the three-momentum of the B and that of the combined Y = π<sup>0</sup>(η) + ℓ system, defined as \( \frac{2 E_\text{beam} E_Y - M_Y^2 - M_B^2}{2 |p_Y| |p_B|} \) (see Sec. 4.8), distinguishes between the signal and all backgrounds. It falls primarily in the physical region \(-1 < \cos θ_\text{BY} < 1\) for signal events<sup>5</sup>, but is a smooth, broader distribution for the backgrounds (see Fig. 4.14).

• **cos θ<sub>γγ</sub>**  In the π<sup>0</sup>s (or η's) rest frame, its two daughter photons have back-to-back momenta, following from the conservation of momentum. In most cases (other than high \( q^2 \), and therefore low \( p_π(η) \) events), the B's hadronic daughter has a high momentum in the cms. Due to its Lorentz boost, the cms angle between a signal π<sup>0</sup>s or η's daughter gammas is quite acute compared to that of most other π<sup>0</sup>s and ηs in an event (see Fig. 4.15).

• **cos θ<sub>thrust</sub>**  Continuum events are more “jetty” than \( B\bar{B} \) events; that is to say, most particles in continuum events have momenta aligned in two or three directions (depending on the number of quark jets). The “maximum of the directed momentum”, or “thrust” axis, was proposed by Farhi in 1977 to measure jetlikeness in \( e^+e^- \) annihilation [68]. Thrust is a property of a collection of final-state particles. Their momentum directed along an arbitrary unit vector, \( \hat{r} \), is given by

\[
d(\hat{r}) = \sum_a p_a \cdot \hat{r} u(p_a \cdot \hat{r})
\]

where \( p \) is momentum, \( u \) is the unit step function

\[
u(n) = \begin{cases} 
0, & n < 0 \\
1, & n \geq 0 
\end{cases}
\]

which selects only non-negative values of \( p_a \cdot \hat{r} \), and the sum index \( a \) represents each particle in the collection. The thrust axis is the normalised direction of maximum \( d(\hat{r}) \):

\[
\text{thrust} = \frac{\max_a d(\hat{r})}{\sum_a |p_a|}
\]

The cosine of the angle between the thrust of the Y = π<sup>0</sup>(η) + ℓ system and the thrust of the rest of the event

\[
\cos θ_\text{thrust} = \left(\frac{\text{thrust}_{π^0(η)ℓ}}{|\text{thrust}_{π^0(η)ℓ}|}\right) \cdot \left(\frac{\text{thrust}_{\text{rest of event}}}{|\text{thrust}_{\text{rest of event}}|}\right)
\]

is very effective in discriminating against continuum events (see Fig. 4.16). This variable is strongly peaked at +1 for this background, and less strongly peaked for \( B\bar{B} \) events. Furthermore, it is not strongly correlated with \( q^2 \) (see Fig. 4.7). This is an important characteristic when the data is divided into \( q^2 \) bins (see Sec. 5.5).

• **cos θ<sub>Wℓ</sub>**  This is the angle of the lepton’s three-momentum in the virtual W’s rest frame with respect to the virtual W’s three-momentum in the B’s rest frame. Semileptonic B decays to pseudoscalar mesons are predicted to have a model-independent \( \sin^2 θ_\text{Wℓ} \) (i.e. \( 1 - \cos^2 θ_\text{Wℓ} \)) distribution, whereas background distributions tend to be flat or biased towards \( ±1 \) (see Fig. 4.17). This variable is only well defined in the case where cosθ<sub>Wℓ</sub> is physical.

• **M<sub>bc</sub>**  The reconstructed B candidate’s beam-constrained mass, given by

\[
\sqrt{E_{\text{beam}}^2 - |\vec{p}_{π^0(η)} + \vec{p}_ℓ + \vec{p}_ν|^2}
\]

and described in Sec 4.7, is strongly peaked at the nominal B mass (5.279 GeV) for signal events, and peaked at a slightly lower value for background events (see Fig. 4.18).

---

<sup>5</sup>Due to the detector’s resolution, the \( \cos θ_\text{BY} \) value for signal events occasionally strays outside the range \(-1 < \cos θ_\text{BY} < 1\).
Figure 4.7: $\cos \theta_{\text{thrust}}$ vs. $q^2$ for $B \to \pi^0 \ell \nu$ signal MC. The $\cos \theta_{\text{thrust}}$ axis is unitless, and the $q^2$ axis is in units of GeV$^2$. Note that $\cos \theta_{\text{thrust}}$ is not strongly correlated with $q^2$.

- $\Delta E$ The event’s missing energy, $\Delta E = (E_{\pi^0(\eta)}^* + E_l^* + E_\nu^*) - E_{\text{beam}}^*$, was also described in Sec 4.7. It is strongly peaked at zero for signal events, and at lower values for the backgrounds (see Fig. 4.19).

The $p_T^*$ and $\Delta E$ variables are used to determine the region of interest, or “fit region”, for signal extraction (see Fig 4.8).

<table>
<thead>
<tr>
<th>Selection</th>
<th>Number of Signal Events</th>
<th>Signal Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B \to \pi^0 \ell \nu$</td>
<td>$B \to \eta \ell \nu$</td>
</tr>
<tr>
<td>Complete Sample</td>
<td>802,255</td>
<td>936,860</td>
</tr>
<tr>
<td>lepton skim</td>
<td>546,442</td>
<td>618,980</td>
</tr>
<tr>
<td>$\pi^0/\eta \to \gamma\gamma$</td>
<td>539,397</td>
<td>242,045</td>
</tr>
<tr>
<td>Event reconstruction and Best $B$ selection</td>
<td>147,821</td>
<td>35,183</td>
</tr>
<tr>
<td>Fit region</td>
<td>138,727</td>
<td>25,884</td>
</tr>
</tbody>
</table>
Figure 4.8: Two-dimensional histograms of the fit region, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. Both the $\Delta E$ and $p^*_\ell$ axes are in units of GeV. The events in these distributions have passed the lepton skim and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
After the set of cuts had been selected, their exact values were optimised. One by one, each threshold was varied, such that a figure of merit

\[
\text{Figure of Merit} = \frac{\text{Signal}}{\sqrt{\text{Signal} + \text{Backgrounds}}} \tag{4.13}
\]

where \(\text{Signal}\) represents the number of signal events and \(\text{Backgrounds}\) represents the sum of all background events, was maximised. While one cut was being optimised, all other cut values were held constant. The final set of cuts for \(B \to \pi^0\ell\nu\), resulting from the optimisation, are shown in Table 4.8.

Table 4.8: The optimised set of event selection criteria, or “cuts”, for \(B \to \pi^0\ell\nu\). These values were found through the optimisation procedure.

<table>
<thead>
<tr>
<th>Quantity “(x)” (units)</th>
<th>Analysis cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of leptons</td>
<td>(x = 1)</td>
</tr>
<tr>
<td>(\cos \theta_\nu)</td>
<td>(-0.86 &lt; x &lt; 0.96)</td>
</tr>
<tr>
<td>Net charge (</td>
<td>(e</td>
</tr>
<tr>
<td>(M_{miss}^2 ) (GeV(^2))</td>
<td>(-2.5 &lt; x &lt; 4.0)</td>
</tr>
<tr>
<td>(p_\ell^x) (GeV)</td>
<td>(x &gt; 1.2)</td>
</tr>
<tr>
<td>(R_2)</td>
<td>(x &lt; 0.5)</td>
</tr>
<tr>
<td>(\cos \theta_{BY})</td>
<td>(-0.95 &lt; x &lt; 1.10)</td>
</tr>
<tr>
<td>(\cos \theta_{\gamma\gamma})</td>
<td>(x &gt; 0.9)</td>
</tr>
<tr>
<td>(\cos \theta_{\text{thrust}})</td>
<td>(-0.90 &lt; x &lt; 0.85)</td>
</tr>
<tr>
<td>(\cos \theta_{W\ell})</td>
<td>(-0.7 &lt; x &lt; 0.9)</td>
</tr>
<tr>
<td>(M_{bc}) (GeV)</td>
<td>(5.26 &lt; x &lt; 5.29)</td>
</tr>
<tr>
<td>(\Delta E) (GeV)</td>
<td>(</td>
</tr>
</tbody>
</table>

Signal and background distributions of each cut variable, before any neutrino reconstruction or analysis cuts have been applied, can be seen in Figures 4.9 - 4.19. Vertical lines on these histograms represent the placement of the cuts listed in Table 4.8. The effect of each cut on signal and background efficiencies can be seen in Table 4.9.

Now that the events have been reconstructed, the particles analysed, and the selection criteria imposed, the process of signal extraction can begin. \(B \to \pi^0\ell\nu\) will be discussed in Chapter 5. The selection criteria and signal extraction for the \(B \to \eta\ell\nu\) mode will be described in Chapter 6.
Table 4.9: The effect of each cut on the efficiency of each sample, for $B \to \pi^0 \ell \nu$. The final row shows the effect of all of the cuts put together. Cuts are not cumulative, except for the final row. The efficiencies listed are with respect to each sample after skimming, event reconstruction, and $B$ candidate selection, not with respect to the complete sample. All values are percentages except for $S/B$.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Signal</th>
<th>Crossfeed</th>
<th>$B^0 B^\ell$</th>
<th>$B^+ B^-$</th>
<th>$uds$</th>
<th>charm</th>
<th>Bkgd sum</th>
<th>$S/B \times 10^4$</th>
<th>On-res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no cuts]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>1.09</td>
<td>100</td>
</tr>
<tr>
<td>$\cos \theta_\nu$</td>
<td>94.27</td>
<td>92.24</td>
<td>85.30</td>
<td>85.10</td>
<td>65.16</td>
<td>73.95</td>
<td>80.33</td>
<td>1.28</td>
<td>77.06</td>
</tr>
<tr>
<td>Net Charge</td>
<td>82.62</td>
<td>75.75</td>
<td>65.96</td>
<td>70.98</td>
<td>76.23</td>
<td>76.41</td>
<td>71.35</td>
<td>1.27</td>
<td>69.69</td>
</tr>
<tr>
<td>$M_{miss}^2$</td>
<td>49.53</td>
<td>48.88</td>
<td>40.97</td>
<td>44.60</td>
<td>57.46</td>
<td>48.00</td>
<td>46.07</td>
<td>1.18</td>
<td>43.87</td>
</tr>
<tr>
<td>$p_\ell^*$</td>
<td>99.47</td>
<td>99.67</td>
<td>99.54</td>
<td>99.54</td>
<td>98.35</td>
<td>98.91</td>
<td>99.25</td>
<td>1.10</td>
<td>99.07</td>
</tr>
<tr>
<td>$R_2$</td>
<td>99.09</td>
<td>99.33</td>
<td>99.54</td>
<td>99.53</td>
<td>80.49</td>
<td>84.35</td>
<td>93.88</td>
<td>1.15</td>
<td>92.27</td>
</tr>
<tr>
<td>$\cos \theta_{BY}$</td>
<td>61.33</td>
<td>23.41</td>
<td>10.57</td>
<td>10.89</td>
<td>15.60</td>
<td>15.10</td>
<td>12.57</td>
<td>5.34</td>
<td>12.84</td>
</tr>
<tr>
<td>$\cos \theta_{\gamma\gamma}$</td>
<td>79.11</td>
<td>57.75</td>
<td>61.37</td>
<td>63.96</td>
<td>69.41</td>
<td>70.65</td>
<td>65.09</td>
<td>1.33</td>
<td>63.62</td>
</tr>
<tr>
<td>$\cos \theta_{\text{thrust}}$</td>
<td>85.45</td>
<td>82.16</td>
<td>81.54</td>
<td>81.21</td>
<td>36.89</td>
<td>41.73</td>
<td>67.38</td>
<td>1.39</td>
<td>68.04</td>
</tr>
<tr>
<td>$\cos \theta_{W\ell}$</td>
<td>88.14</td>
<td>82.46</td>
<td>78.57</td>
<td>79.41</td>
<td>63.61</td>
<td>72.50</td>
<td>75.64</td>
<td>1.27</td>
<td>77.19</td>
</tr>
<tr>
<td>$\Delta E$</td>
<td>96.43</td>
<td>88.55</td>
<td>76.95</td>
<td>77.09</td>
<td>85.30</td>
<td>84.02</td>
<td>79.82</td>
<td>1.32</td>
<td>80.92</td>
</tr>
<tr>
<td>All cuts</td>
<td>9.40</td>
<td>0.54</td>
<td>0.06</td>
<td>0.08</td>
<td>0.03</td>
<td>0.03</td>
<td>0.14</td>
<td>153.05</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Figure 4.9: $\cos \theta_\nu$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.10: Net charge histograms, in units of $|e|$, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The histogram bins are symmetric about integer values, leading to the appearance of a non-integer cut; the actual cut values are ±1e. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.11: $M_{\text{miss}}^2$ histograms, in units of GeV$^2$, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), $uds$ (bottom left) and $charm$ (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.12: cms lepton momentum ($p^*_\ell$) histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised lower cut threshold. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied. Note that the lepton skim includes a cut of $p^*_\ell > 1.2$ GeV, but slight differences between the particle identification algorithms used in the skim and event reconstruction result in a reconstructed cms lepton momentum of slightly less than 1.2 GeV in a small number of events.
Figure 4.13: $R_2$ histograms for π⁰ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised upper cut threshold. The events in these distributions have passed the lepton_skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.14: $\cos \theta_{BY}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right),uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.15: \(\cos \theta_{\gamma\gamma}\) histograms for \(\pi^0\) signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical line indicates the optimised lower cut threshold. The events in these distributions have passed the lepton skim, event reconstruction, and \(B\) candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.16: $\cos \theta_{\text{thrust}}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.17: $\cos \theta_{W\ell}$ histograms for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.18: $M_{bc}$ histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 4.19: $\Delta E$ histograms, in units of GeV, for $\pi^0$ signal (top left), crossfeed (top right), mixed (middle left), charged (middle right), uds (bottom left) and charm (bottom right) MC. The vertical lines indicate the optimised cut range. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Chapter 5

$B \to \pi^0 \ell \nu$ Signal Extraction

So far, events have been skimmed with both HadronB and lepton skim, and reconstructed in search of $B \to \pi^0 \ell \nu$ decays. Neutrino reconstruction cuts and analysis cuts have been optimised and applied. Up to this point, all description has been in relation to MC events. The same event reconstruction and selection criteria (skims and cuts) were then applied to real data, and the signal and background MC distributions were fitted to the data in order to estimate their relative weightings. An excess of data over background MC, in regions where the signal was expected, was considered to be a $B \to \pi^0 \ell \nu$ signal. These signal events were used to calculate the $B \to \pi^0 \ell \nu$ branching fraction and extract a preliminary value of $V_{ub}$.

The analysis process is illustrated in Fig. 5.1.

![Figure 5.1](image)

Figure 5.1: The process of analysing $B \to \pi^0 \ell \nu$. The previous chapter described the steps up to and including the application of analysis cuts. This chapter will describe the fitting procedure and branching fraction calculation.

5.1 Fitting MC Backgrounds to On-resonance Data

A binned maximum likelihood fit was performed using ROOT’s TFractionFitter algorithm, which is based on the method developed by Barlow and Beeston [69]. This takes into account the
statistical uncertainties of the data and MC histograms. Signal and background MC and
on-resonance data histograms (after selection criteria had been applied) were inputted into the fitter,
which varied the bin contents of the template histograms within their statistical errors. Because
the number of events in the MC samples was significantly greater than that expected in the data,
the signal and background histograms were normalised in the fitting process. The fitter adjusted
the weights of the signal and background histograms until the shape of their weighted sum best
matched the shape of the data histogram. The program output included the fitted yields and
errors corresponding to each input histogram, the number of degrees of freedom used in the fit,
and a $\chi^2$ goodness-of-fit. The fitting process is illustrated in Fig. 5.2.

A two-dimensional fit in the $\Delta E - p_T^*\ell$ plane was used in this analysis. These two variables
were chosen for the fit as they gave the best discrimination between the signal and background
histogram shapes. The signal and background efficiencies in this region are shown in Table 5.1.
The mixed and charged $B\bar{B}$ MC samples were summed before the fit, as were the $uds$ and $charm$
MC samples. This ensured that these samples were constrained to the ratios dictated by the
physics at $\sqrt{s} = 10.58$ GeV, and reduced the number of parameters in the fit. Projections of the
fitted histograms are shown in Fig. 5.3, and the results of the fit can be seen in Table 5.2.
Table 5.1: Signal, background and on-resonance data efficiencies in the fit region, $|\Delta E| < 1.0$ GeV and $p^*_\ell > 1.2$ GeV, after lepton skim and $B$ candidate selection, but before neutrino reconstruction and analysis cuts have been applied. All values listed are percentages.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>95.96</td>
</tr>
<tr>
<td>Crossfeed</td>
<td>88.27</td>
</tr>
<tr>
<td>$B^0\bar{B}^0$</td>
<td>76.68</td>
</tr>
<tr>
<td>$B^+B^-$</td>
<td>76.80</td>
</tr>
<tr>
<td>uds</td>
<td>84.24</td>
</tr>
<tr>
<td>charm</td>
<td>83.38</td>
</tr>
<tr>
<td>Bkgd sum</td>
<td>79.36</td>
</tr>
<tr>
<td>On-res.</td>
<td>80.34</td>
</tr>
</tbody>
</table>

Figure 5.3: $\Delta E$ (left) and $p^*_\ell$ (right) projections of the 2D fit result. Both the $\Delta E$ and $p^*_\ell$ axes are in units of GeV.

5.2 The $B \rightarrow \pi^0 \ell \nu$ Signal

As can be seen in Fig. 5.3, a $B \rightarrow \pi^0 \ell \nu$ signal is decidedly evident in the Belle data. The fit projection histograms show a clear shoulder above the fitted background predictions, with a substantial fitted yield of $1084.8 \pm 116.8$ signal events. The fact that the signal appears as a shoulder rather than a sharp peak reflects the trade-off between signal efficiency and background suppression when looking for a rare signal in a “dirty” environment requiring neutrino reconstruction.

5.3 Signal Efficiency

The signal efficiency was taken to be the number of events (after all cuts) in the signal histogram which was inputted into the fitter, relative to the number of $B \rightarrow \pi^0 \ell \nu$ events in the full b2ujnu MC sample. Remember from Section 4.1 that in the b2ujnu MC sample it was known that one $B$ meson decayed to a $b \rightarrow u \ell \nu$ mode, and the other $B$ meson decayed to a common mode. The signal efficiency was found to be $(1.72 \pm 0.01)\%$. 
Table 5.2: Fitted signal and background yields, the Signal to Background ratio (S/B), and the \( \chi^2/(\text{number of degrees of freedom}) \) goodness-of-fit for the \( B \to \pi^0\ell\nu \) fit.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Crossfeed</th>
<th>( B\bar{B} )</th>
<th>Continuum</th>
<th>S/B</th>
<th>( \chi^2/\text{ndf} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1084.8 ± 116.8</td>
<td>1932.4 ± 195.1</td>
<td>4977.6 ± 135.1</td>
<td>1225.3 ± 122.1</td>
<td>0.133</td>
<td>289/396</td>
</tr>
</tbody>
</table>

### 5.4 Branching Fraction Calculation

The branching fraction (BF or \( B \)) for an initial particle \( A \) decaying to a specific final state \( X \) is defined as

\[
\frac{A \to X}{A \to \text{everything}}
\]  

(5.1)

In the field of B physics, this is generally expressed as the ratio of the number of events containing the desired final state to the number of \( \Upsilon(4S) \) events,\(^1\) i.e.,

\[
\mathcal{B}(B^+ \to \pi^0\ell^+\nu_{\ell}) = \frac{\text{number of } B^+ \to \pi^0\ell^+\nu_{\ell} \text{ events}}{\text{number of } \Upsilon(4S) \text{ events}}
\]  

(5.2)

In this analysis, the decay \( B^+ \to \pi^0\ell^+\nu_{\ell} \) and its charge-conjugate decay were assumed to have the same branching fraction, and \( \ell \) can represent either an \( e \) or a \( \mu \). The branching fraction denoted by \( \mathcal{B}(B \to \pi^0\ell\nu) \) is the average branching fraction of these four decays. The calculated branching fraction was actually

\[
\mathcal{B}(B \to \pi^0\ell\nu) = \frac{(\text{number of } B^- \to \pi^0\ell^-\nu_{\ell} \text{ events}) + (\text{number of } B^+ \to \pi^0\ell^+\nu_{\ell} \text{ events})}{(\text{Signal Efficiency}) \times 2 \times (\text{number of } B\bar{B} \text{ events})}
\]  

(5.3)

where the “number of \( B^{\pm} \to \pi^0\ell^{\pm}\nu \text{ events} \)” was the signal yield.

The preliminary \( B \to \pi^0\ell\nu \) branching fraction was found to be:

\[
\mathcal{B}(B \to \pi^0\ell\nu) = (0.59 \pm 0.06) \times 10^{-4}
\]

The error is statistical only. The preliminary systematic and model errors are discussed in Secs. 5.6 and 5.7 below.

### 5.5 Branching Fraction in \( q^2 \) Bins

The theoretical techniques commonly used to model \( b \to u\ell\nu \) decays were described in Sec. 2.7. From the experimentalist’s perspective, the main difference between these approaches is visible in the predicted branching fractions’ variation with \( q^2 \) (see Fig. 2.4), the square of the “mass” of the virtual \( W \) boson that decays into the charged lepton and the neutrino. As the signal and crossfeed

\(^1\)The \( \Upsilon(4S) \) decays to \( B\bar{B} \) more than 96% of the time [20].
MC must necessarily be based on theoretical predictions, the model used to generate the MC can affect the signal and crossfeed efficiencies, and the shape of their $q^2$ distributions. By dividing the data into $q^2$ regions, or “bins”, the calculated branching fraction becomes less dependent upon the model used to generate the $b \to u\ell\nu$ decays. While the statistical errors are necessarily higher in each bin than across the entire sample, a binned measurement with enough statistics and fine enough bins can be used to confirm or rule out individual models.

A series of $q^2$ cuts were applied in order to estimate the branching fraction in three $q^2$ bins: $q^2 < 8 \text{ GeV}^2$, $8 \text{ GeV}^2 \leq q^2 < 16 \text{ GeV}^2$, and $16 \text{ GeV}^2 \leq q^2$. The partial branching fractions, signal efficiencies, and fitted yields in each bin are shown in Table 5.3, along with the unbinned results. Each binned yield was efficiency-corrected separately to calculate the partial branching fractions.

The sum of the partial branching fractions differs from the central value of the unbinned branching fraction, though the binned and unbinned results are consistent within errors and the yields are comparable. The difference in central values may indicate that the $q^2$ shapes are playing a role.

Table 5.3: Signal efficiency, fitted yields, and the partial branching fraction in each $q^2$ region. The statistical error of the summed region is the quadratic sum of those of the individual regions.

<table>
<thead>
<tr>
<th>$q^2$ region (GeV$^2$)</th>
<th>Signal Eff. (%)</th>
<th>Fitted Yield</th>
<th>Partial BF ($\times10^4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 8</td>
<td>0.57 ± 0.01</td>
<td>219.4 ± 48.6</td>
<td>0.36 ± 0.08</td>
</tr>
<tr>
<td>8 − 16</td>
<td>2.53 ± 0.03</td>
<td>532.7 ± 81.3</td>
<td>0.20 ± 0.03</td>
</tr>
<tr>
<td>&gt; 16</td>
<td>2.31 ± 0.04</td>
<td>305.3 ± 80.0</td>
<td>0.12 ± 0.03</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td></td>
<td>0.68 ± 0.09</td>
</tr>
<tr>
<td>all $q^2$</td>
<td>1.72 ± 0.01</td>
<td>1084.8 ± 116.8</td>
<td>0.59 ± 0.06</td>
</tr>
</tbody>
</table>

The branching fraction obtained by summing the partial branching fractions is less model-dependent than that calculated globally, and is considered the primary result of this analysis.

A comparison of the partial branching fractions and LCSR and ISGW2 predictions can be seen in Fig. 5.4. The measurement favours the ISGW2 prediction over that of LCSR, though neither model can be ruled out.

The fit projections can be seen in Fig. 5.5. The $B \to \pi^0\ell\nu$ signal is evident in each bin, with the S/B highest in the $q^2 < 8 \text{ GeV}^2$ bin. There is more crossfeed in the higher $q^2$ bins, where the S/B is correspondingly poorer.
5.6 Preliminary Systematic Study

A preliminary systematic study was conducted. The primary focus was on the neutrino reconstruction, particle identification, the number of $B\bar{B}$ pairs in the 492 fb$^{-1}$ dataset, event selection criteria, and the assumed $b \to u\ell\nu$ and $b \to c\ell\nu$ branching fractions in the MC backgrounds. The neutrino reconstruction systematics were examined by Nicholas Parslow in a previous study written jointly with this author [70]. A full systematic study is left for further work.

As can be seen in Table 5.4, the systematic error is dominated by the neutrino reconstruction, event selection criteria, and $\pi^0$ reconstruction efficiency. As all detected particles are used to
reconstruct the neutrino, there is a correlation between the systematic uncertainty of the lepton
and that of the neutrino.\(^2\) Accordingly, the lepton track-finding and identification errors were
added linearly to the neutrino reconstruction error; all others were added in quadrature.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>Assigned systematic error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrino reconstruction eff.</td>
<td>10.0</td>
</tr>
<tr>
<td>Lepton track finding eff.</td>
<td>1.0</td>
</tr>
<tr>
<td>Lepton identification</td>
<td>2.1</td>
</tr>
<tr>
<td>(\pi^0) reconstruction eff.</td>
<td>4.6</td>
</tr>
<tr>
<td>(N(B\bar{B}))</td>
<td>1.7</td>
</tr>
<tr>
<td>Event selection criteria</td>
<td>6.8</td>
</tr>
<tr>
<td>(b \to u\ell\nu) branching fractions</td>
<td>2.0</td>
</tr>
<tr>
<td>(b \to c\ell\nu) branching fractions</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>15.7</td>
</tr>
</tbody>
</table>

5.6.1 Neutrino Reconstruction

The difference in the effect of neutrino reconstruction cuts on real data and MC is the largest
systematic error in exclusive \(b \to u\ell\nu\) analyses which use the neutrino reconstruction technique.
The systematic study examined a \(B \to D^*\ell\nu\) control sample, which has very clean signature.
Yields obtained with all of the analysis cuts were compared with yields obtained with the cuts
involving only the pion and charged lepton. This ratio was approximately 10% higher in the MC
than in the data. The systematic error due to neutrino reconstruction was therefore estimated to
be 10%.

5.6.2 Particle Identification

Charged lepton and \(\pi^0\) systematics were estimated to be 1% for lepton track finding efficiency,
2.1% for lepton identification, and 4.6% for \(\pi^0\) reconstruction efficiency, using tools provided by
the Belle Tracking Group [71], the Belle Particle Identification Joint Group [72], and a Belle
study of \(\pi^0\) systematics [73, 74], respectively. These systematic errors are not correlated and are
therefore added in quadrature. Effects due to charged hadron identification would only appear in
the neutrino reconstruction, and are included in that systematic error.

5.6.3 \(N(B\bar{B})\)

The number of \(B\bar{B}\) pairs in the data sample was varied within the statistical error determined by
the Belle Collaboration [75]. The number of \(B\bar{B}\) events in Experiments 7-49 is

\[
N(B\bar{B}) = (534.586 \pm 7.044) \times 10^6
\] (5.4)

\(^2\)Remember from Sec. 4.6 that it was the daughter photons, not the \(\pi^0\)'s themselves, that were used to reconstruct
the neutrino. Therefore, the \(\pi^0\) systematic uncertainty is not correlated with that of the neutrino.
The systematic error in the $B \to \pi^0 \ell \nu$ branching fraction due to the number of $B\bar{B}$ pairs in 492 fb$^{-1}$ of Belle data was estimated to be 1.7%.

5.6.4 Event Selection Criteria

The choice of selection criteria was examined. In particular, the effect of selecting with the original $M_{bc}$ variable, the “stretched” variable $M_{\alpha}^{bc}$, or both, was investigated, as was the use (or not) of the cut on the angle between the signal $\pi^0$’s daughter photons. During this investigation, all other cuts were imposed and cut values were not varied from those quoted in Sec. 4.10.2.

$M_{bc}$ is calculated in a similar fashion to $M_{bc}$, but the “stretched” neutrino mentioned in Sec. 4.8 is used in place of the original neutrino:

$$M_{\alpha}^{bc} = \sqrt{E_{beam}^* - p_{\pi^0(\eta)}^* + p_l^* + \alpha p_{\nu}^*}$$  \hspace{1cm} (5.5)

The use of $M_{\alpha}^{bc}$ resulted in a higher signal efficiency but a lower S/B, while the use of $\cos \theta_{\gamma\gamma}$ resulted in a higher S/B but a lower signal efficiency (see Table 5.5). A conservative estimate of 6.8% was assigned for the uncertainty due to event selection criteria.

<table>
<thead>
<tr>
<th>$\cos \theta_{\gamma\gamma}$ cut</th>
<th>$B(B \to \pi^0 \ell \nu)$ $(\times 10^4)$</th>
<th>Stat. Error (%)</th>
<th>S/B</th>
<th>$\chi^2$ (ndf=396)</th>
<th>Sig. Eff. (%)</th>
<th>Eff.-cor. Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{bc}$</td>
<td>0.59 ± 0.06</td>
<td>10.2</td>
<td>0.165</td>
<td>289</td>
<td>1.72</td>
<td>1084.8 ± 116.8</td>
</tr>
<tr>
<td>$M_{bc}^{\alpha}$</td>
<td>0.56 ± 0.06</td>
<td>10.7</td>
<td>0.148</td>
<td>388</td>
<td>2.05</td>
<td>1224.3 ± 128.1</td>
</tr>
<tr>
<td>$M_{bc}$ and $M_{bc}^{\alpha}$</td>
<td>0.60 ± 0.06</td>
<td>10.0</td>
<td>0.182</td>
<td>284</td>
<td>1.68</td>
<td>1071.3 ± 115.6</td>
</tr>
</tbody>
</table>

| No $\cos \theta_{\gamma\gamma}$ cut     | $M_{bc}$                                 | 0.63 ± 0.06    | 9.5  | 0.151              | 261          | 1.81           | 1215.7 ± 125.1 |
|                                         | $M_{bc}^{\alpha}$                        | 0.57 ± 0.06    | 10.5 | 0.129              | 355          | 2.16           | 1320.1 ± 132.2 |
|                                         | $M_{bc}$ and $M_{bc}^{\alpha}$           | 0.63 ± 0.07    | 11.1 | 0.165              | 269          | 1.77           | 1198.0 ± 123.4 |

5.6.5 $b \to u \ell \nu$ and $b \to c \ell \nu$ Normalisations

The $b \to u \ell \nu$ and $b \to c \ell \nu$ MC samples contain a mix of individual exclusive $B$ decay modes, with the weighting of each mode in the sample reflecting its measured (or predicted) branching fraction. These branching fractions have errors which can affect normalisations within the samples. This is a different effect from the $b \to u \ell \nu$ modelling, which affects the modes’ distributions in phase space and is discussed in Sec. 5.7. The $B \to \rho \ell \nu$, $B \to \omega \ell \nu$, $B \to \eta \ell \nu$, $B \to D^* \ell \nu$ and $B \to D \ell \nu$ branching fractions were individually varied within their errors. This resulted in a systematic error of order 2% for $b \to u \ell \nu$ crossfeed and 0.5% for $b \to c \ell \nu$. 

74
5.7 Preliminary $b \to u\ell\nu$ Model Error

Errors on the branching fraction due to the uncertainty in the $b \to u$ form factors in the signal and $b \to u$ crossfeed were estimated by varying the models used to generate the $b \to u\ell\nu$ MC. The primary $b \to u\ell\nu$ MC sample, used throughout this thesis, was generated using LCSR models for the $\pi$ (Ball'01; [2]) and $\rho$ and $\omega$ (Ball'98; [6]) modes, and the ISGW2 [3] model for $\eta$, $\eta'$, $a_2$, $f_2$, $b_1$, $h_1$, $a_1$, $f_1$, $a_0$, and $f_0$ modes (see Sec. 4.1). A second MC sample, generated using the ISGW2 model for most modes, was also available. This sample consisted of approximately 4 million events, including 171,666 $B \to \pi^0\ell\nu$ decays.

Due to the poorer statistics in the ISGW2 sample (171,666 $B \to \pi^0\ell\nu$ signal events as compared to the LCSR sample’s 802,255), it was not possible to extract a $B \to \pi^0\ell\nu$ branching fraction in three $q^2$ bins. Instead, two bins were used: $q^2 < 16$ GeV$^2$ and $q^2 \geq 16$ GeV$^2$. The partial branching fractions were found to be $(0.60 \pm 0.06) \times 10^{-4}$ in the lower bin and $(0.12 \pm 0.03) \times 10^{-4}$ in the higher bin, with a global efficiency of 1.55%. The sum of the partial branching fractions in these two bins was compared to the branching fraction extracted in three bins using the LCSR MC, and the difference in central values was taken as the uncertainty due to $b \to u\ell\nu$ modelling.

The effect of the model used to generate the $b \to u\ell\nu$ MC on the $B \to \pi^0\ell\nu$ branching fraction and signal efficiency can be seen in Table 5.6.

Table 5.6: Variation in the $B \to \pi^0\ell\nu$ branching fraction and signal efficiency due to the model used to generate $b \to u\ell\nu$ events. The LCSR measurement was summed over three $q^2$ bins, and the ISGW2 measurement was summed over two $q^2$ bins.

<table>
<thead>
<tr>
<th>Model</th>
<th>$B(B \to \pi^0\ell\nu) \times 10^4$</th>
<th>Sig. Eff. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCSR</td>
<td>$0.68 \pm 0.09$</td>
<td>1.72</td>
</tr>
<tr>
<td>ISGW2</td>
<td>$0.72 \pm 0.07$</td>
<td>1.55</td>
</tr>
</tbody>
</table>

The LCSR and ISGW2 models predict quite different $q^2$ distributions (see Fig. 2.4 in Sec. 2.6.2, and Fig 5.6 below). A notable feature of the ISGW2 quark model is that it predicts a particularly large amount of signal at low $q^2$. This study had a comparatively low efficiency at low $q^2$, as can be seen from Fig. 5.5. This combination leads to the lower efficiency, and therefore higher calculated branching fraction, extracted using the ISGW2 MC.

Figure 5.6: $q^2$ distributions for $B \to \pi\ell\nu$ predicted by [1], [2], [3], and [4]. Adapted from [8].

The preliminary $B \to \pi^0\ell\nu$ branching fraction including statistical, preliminary experimental
systematic and preliminary model errors is

\[ B(B \to \pi^0 \ell \nu) = (0.68 \pm 0.09 \pm 0.11 \pm 0.04) \times 10^{-4} \] (5.6)

5.8 Preliminary \( |V_{ub}| \) Measurement

The \( B \to \pi^0 \ell \nu \) branching fraction can be used to extract a value for \( |V_{ub}| \) using Eq. 2.10:

\[ |V_{ub}| = \sqrt{\frac{2 \times B(B \to \pi^0 \ell \nu)}{\Gamma_{thy} \tau_{B^+}}} \] (5.7)

with \( \tau_{B^+} = (1.638 \pm 0.011) \) ps [20]. \( |V_{ub}| \) was measured using \( \Gamma_{thy} \) values from two theories, Ball’04 LCSR [41] and ISGW2 [3], both of which were extrapolated to the full \( q^2 \) region. The sum of the partial branching fractions (those calculated in \( q^2 \) bins) was used to reduce model dependency. The branching fraction calculated using the LCSR MC in three \( q^2 \) bins was used to extract \( |V_{ub}| \) with the Ball’04 \( \Gamma_{thy} \). As the ISGW2 branching fraction could only be calculated in two \( q^2 \) bins, the \( |V_{ub}| \) extraction using the ISGW2 branching fraction and \( \Gamma_{thy} \) is shown to give a sense of the theoretical error on the \( |V_{ub}| \) measurement.

The preliminary \( |V_{ub}| \) results can be seen in Table 5.7.

Table 5.7: Preliminary \( |V_{ub}| \) results, extracted from the \( B \to \pi^0 \ell \nu \) branching fractions measured in this analysis. The branching fraction obtained with LCSR MC was used in the Ball’04 \( |V_{ub}| \) extraction, and that obtained with ISGW2 MC was used in the ISGW2 \( |V_{ub}| \) extraction. The latter is presented to indicate the extent of the theoretical error on the \( |V_{ub}| \) measurement.

| Model Type       | Measured BF \( \times 10^4 \) | \( \Gamma_{thy} \) (ps\(^{-1}\)) | Model        | Measured \( |V_{ub}| \) \( \times 10^3 \) |
|------------------|-------------------------------|---------------------------------|--------------|----------------------------------|
| LCSR             | 0.68 ± 0.09 ± 0.11 ± 0.04     | 7.67 ± 1.84                     | Ball’04 [41] | 3.29 ± 0.23 ± 0.27 ± 0.05        |
| Quark Model      | 0.72 ± 0.07 ± 0.11 ± 0.04     | 9.6 ± 4.8                       | ISGW2 [3]    | 3.0 ± 0.2 ± 0.2 ± 0.8            |

5.9 Comparison with Other Results

Several \( B \to \pi^0 \ell \nu \) results have now been presented by other investigators (see Tables 5.8 and 5.9). Most make use of both \( B \to \pi^0 \ell \nu \) and \( B \to \pi^+ \ell \nu \) measurements, combined using the isospin relation

\[ B(B^0 \to \pi^- \ell^+ \nu) = B(B^+ \to \pi^0 \ell^+ \nu) \times \frac{2 \tau_0}{\tau_+} \] (5.8)

where \( \tau_0(+) \) is the lifetime of the \( B^0(+) \) meson. The ratio \( \frac{\tau_0}{\tau_+} \) is close to unity.
Table 5.8: The most recent $B(B \to \pi \ell \nu)$ results. Charge conjugation is implied in all cases. The first, second, and third and fourth errors are statistical, experimental systematic, and due to $\pi$ and $\rho$ model uncertainties, respectively. Studies that used isospin assumptions to combine $\pi^+$ and $\pi^0$ modes include a factor of $2^{\frac{\Delta m}{\tau}}$ to account for the difference between the $B^0$ and $B^+$ lifetimes.

<table>
<thead>
<tr>
<th>Mode</th>
<th>BF ($\times 10^3$)</th>
<th>Experiment</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu$</td>
<td>$0.68 \pm 0.09 \pm 0.11 \pm 0.04$</td>
<td>Belle</td>
<td>$\nu$ Recon.</td>
<td>This analysis</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu$</td>
<td>$0.77 \pm 0.14 \pm 0.08$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([25])</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu \times 2^{\frac{\Delta m}{\tau}}$</td>
<td>$1.44 \pm 0.26 \pm 0.15$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([25])</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu \times 2^{\frac{\Delta m}{\tau}}$</td>
<td>$1.60 \pm 0.32 \pm 0.11$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([24])</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu$</td>
<td>$0.73 \pm 0.18 \pm 0.08$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([25])</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu$</td>
<td>$0.82 \pm 0.22 \pm 0.11$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([25])</td>
</tr>
<tr>
<td>$B^+ \to \pi^0 \ell^+ \nu \times 2^{\frac{\Delta m}{\tau}}$</td>
<td>$1.36 \pm 0.34 \pm 0.15$</td>
<td>BaBar</td>
<td>Full Recon. tag</td>
<td>HFAG ([25])</td>
</tr>
<tr>
<td>$B^0 \to \pi^- \ell^+ \nu$</td>
<td>$1.33 \pm 0.17 \pm 0.11$</td>
<td>BaBar</td>
<td>Combined SL</td>
<td>CKM06 ([77])</td>
</tr>
<tr>
<td>&amp; $B^+ \to \pi^0 \ell^+ \nu \times 2^{\frac{\Delta m}{\tau}}$</td>
<td>&amp; Full Recon. Tag</td>
<td>&amp;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B^0 \to \pi^- \ell^+ \nu$</td>
<td>$1.33 \pm 0.18 \pm 0.11 \pm 0.01 \pm 0.07$</td>
<td>CLEO</td>
<td>$\nu$ Recon. tag</td>
<td>[8]</td>
</tr>
</tbody>
</table>

Table 5.9: The preliminary $|V_{ub}|$ result obtained in this analysis, compared with recent $|V_{ub}|$ results extracted from the average of multiple $B \to \pi \ell \nu$ branching fractions [24]. The errors on the result from this analysis are statistical, experimental systematic, and due to theoretical modelling, respectively. The errors on the results from [24] are statistical and theoretical, respectively. In all cases, the $\Gamma_{\ell \ell \nu}$ values were extrapolated to the full $q^2$ region.

| $|V_{ub}| (\times 10^3)$ | Model | Reference          |
|------------------------|-------|--------------------|
| $3.29 \pm 0.23 \pm 0.27 \pm 0.05$ | Ball'04 [41] |       |
| $3.40 \pm 0.11^{+0.06}_{-0.07}$ | Ball'04 [41] |       |
| $3.86 \pm 0.13^{+0.83}_{-0.50}$ | HPQCD [78] |       |
| $3.79 \pm 0.13^{+0.87}_{-0.52}$ | FNAL [79] |       |
| $3.58 \pm 0.12^{+0.57}_{-0.52}$ | APE [80] |       |

The $B \to \pi^0 \ell \nu$ branching fraction measured in this analysis was

$$B(B \to \pi^0 \ell \nu) = (0.68 \pm 0.09 \pm 0.11 \pm 0.04) \times 10^{-4}$$

This result was used to extract a preliminary $|V_{ub}|$ measurement:

$$|V_{ub}| = (3.29 \pm 0.23 \pm 0.27 \pm 0.05) \times 10^{-3}$$

For both results, the errors are statistical, experimental systematic, and due to $b \to ul\nu$ modelling, respectively. As can be seen in Tables 5.8 and 5.9, the preliminary $B \to \pi^0 \ell \nu$ branching fraction and $|V_{ub}|$ results presented in this chapter are competitive, and in agreement, with publicly presented results. The statistical error on the $B \to \pi^0 \ell \nu$ branching fraction obtained in this analysis is significantly smaller than that achieved by any of the other studies. The method used in this analysis has thus been shown to be both valid and effective, and is applied to $B \to \eta \ell \nu$ in the following chapter.
Chapter 6

\[ B \rightarrow \eta \ell \nu \] Signal Extraction

The properties of the \( \eta \) are akin to those of the \( \pi^0 \). Both are electrically neutral, light pseudoscalar mesons. Their quark content is similar: \( \frac{u_u + d_d}{\sqrt{2}} \) for the \( \pi^0 \) and \( \frac{u_u + d_d - 2s_s}{\sqrt{6}} \) for the \( \eta \). Due to these properties, \( B \rightarrow \eta \ell \nu \) and \( B \rightarrow \pi^0 \ell \nu \) decays are each governed by one form factor; some of the kinematics, however, such as the signal meson’s velocity, differ due to the \( \eta \)’s greater mass. \( \pi^0 \) and \( \eta \) can both decay to two photons, though their branching fractions to this common state are quite distinct: \( B(\pi^0 \rightarrow \gamma \gamma) = 99.798\% \), while \( B(\eta \rightarrow \gamma \gamma) = 39.38\% \) [20]. As the end products are the same (two photons, a charged lepton, and an undetected neutrino) and the branching fractions are expected to be similar, the method used to measure the \( B \rightarrow \pi^0 \ell \nu \) branching fraction, described in the previous two chapters, was applied to \( B \rightarrow \eta \ell \nu \).

6.1 Event Reconstruction and Selection

The \( B \rightarrow \eta \ell \nu \) study followed the same procedure as the \( B \rightarrow \pi^0 \ell \nu \) analysis. Candidate \( B \) mesons were reconstructed in the \( B \rightarrow \eta \ell \nu \) mode, with \( \eta s \) being reconstructed in the \( \eta \rightarrow \gamma \gamma \) mode only. The reconstruction efficiency of \( B \rightarrow \eta \ell \nu \), is therefore considerably lower than that of \( B \rightarrow \pi^0 \ell \nu \), and when cuts were introduced, the desire for a high S/B needed to be balanced with the number of events required by the fitter to provide shape discrimination. Neutrino reconstruction and analysis cuts were applied and optimised for the \( B \rightarrow \eta \ell \nu \) signal. Signal and background MC histograms were then fitted to the on-resonance data histogram, and a preliminary \( B \rightarrow \eta \ell \nu \) branching fraction was calculated.

The neutrino reconstruction and analysis cuts were applied to the same quantities as for \( B \rightarrow \pi^0 \ell \nu \). The list of cuts and their values can be seen in Table 6.1. Many are identical to those used in the previous chapter. The values of six cuts, shown in bold font in Table 6.1, differed. The \( B \rightarrow \eta \ell \nu \) histograms of these spectra are shown in Figs. 6.1-6.6. The equivalent \( B \rightarrow \pi^0 \ell \nu \) histograms are shown in the same figure for comparison. In these figures, \( B \rightarrow \eta \ell \nu \) histograms are shown on the left and \( B \rightarrow \pi^0 \ell \nu \) histograms are shown on the right. The mixed \( (B^0 \overline{B}^0) \) and charged \( (B^+ B^-) \) MC samples have been combined, as have the uds \((u\bar{u}, d\bar{d}, s\bar{s})\) and charm \((c\bar{c})\) samples. The vertical lines indicate the value of the cuts described in the respective studies.
Figure 6.1: $M^2_{miss}$ histograms, in units of GeV$^2$, for $B \to \eta \ell \nu$ (left) and $B \to \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined uds and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 6.2: $R_2$ histograms for $B \to \eta \ell \nu$ (left) and $B \to \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined uds and charm continuum (bottom row). The vertical lines indicate the respective upper cut thresholds for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 6.3: $\cos \theta_{BY}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\bar{B}$ (third row), and combined uds and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 6.4: $\cos \theta_{\gamma\gamma}$ histograms for $B \rightarrow \eta \ell \nu$ (left) and $B \rightarrow \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B\overline{B}$ (third row), and combined uds and charm continuum (bottom row). The vertical lines indicate the respective lower cut thresholds for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 6.5: $\cos \theta_{\text{thrust}}$ histograms for $B \to \eta \ell \nu$ (left) and $B \to \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $B \bar{B}$ (third row), and combined uds and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
Figure 6.6: $\cos \theta_{W\ell}$ histograms for $B \to \eta \ell \nu$ (left) and $B \to \pi^0 \ell \nu$ (right) MC. The histograms are signal (top row), crossfeed (second row), combined mixed and charged $BB$ (third row), and combined $uds$ and charm continuum (bottom row). The vertical lines indicate the respective cut ranges for each study. The events in these distributions have passed the lepton skim, event reconstruction, and $B$ candidate selection, but neutrino reconstruction and analysis cuts have not been applied.
The optimised value for the upper \( \cos \theta \) is the \( \eta \) for the \( B \) difference in the thresholds increased the Figure of Merit for \( B \) to momentum conservation. Therefore, \( \cos \theta \) is in some ways a measure of the signal hadron’s momentum in the cms. The \( \eta \) has considerably more mass than the \( \pi^0 \) (548 MeV as compared to 135 MeV for the \( \pi^0 \)), and a correspondingly softer momentum spectrum. In both the \( B \) → \( \eta \ell \nu \) and the \( B \) → \( \pi^0 \ell \nu \) analyses, the “best \( B \)’ candidate was chosen to be that whose hadronic daughter \( (\eta \) or \( \pi^0 \)) had the greatest momentum. The \( e^+e^- \) events tend to contain many more \( \pi^0 \)’s than \( \eta_s \). This combinatorial background, combined with the \( \pi^0 \)’s lower mass, leads to a significant difference in the \( B \) → \( \eta \ell \nu \) and \( B \) → \( \pi^0 \ell \nu \) \( \cos \theta_\gamma \) spectra. The optimisation study indicated that a lower threshold of 0.7 GeV would maximise the Figure of Merit \((\Sigma Signal)\) for \( B \) → \( \eta \ell \nu \).

For five of the analysis cut variables \( (M^2_{\text{miss}}, R_2, \cos \theta_{\text{BY}}, \cos \theta_{\text{Wf}}, \text{and } \cos \theta_{\text{thrust}}) \), the \( B \) → \( \eta \ell \nu \) and \( B \) → \( \pi^0 \ell \nu \) spectra were slightly different, and the optimisation study showed that changing the thresholds increased the Figure of Merit for \( B \) → \( \eta \ell \nu \).

The optimised value for the upper \( \cos \theta_{\text{BY}} \) threshold was found to be 1.0. The value used in the final analysis was 1.3. The discrepancy between these values is a due to the requirements of the fitting program. The fitter adjusts the input MC histograms’ normalisations until their sum matches the shape of the input data histogram (see Sec. 5.1). In order for the fit to converge, the fitter must be able to discriminate between the shapes of the input histograms. The optimisation study maximised the Figure of Merit, but in doing so it removed so much background as to prevent the fit from converging. Several variations on the optimised cuts were examined. Releasing the upper \( \cos \theta_{\text{BY}} \) cut from 1.0 to 1.3 was found to provide the fitter with enough shape discrimination for the fit to converge, while retaining a high enough S/B to extract a branching fraction.

The fit region used for \( B \) → \( \eta \ell \nu \) was the same as that used for \( B \) → \( \pi^0 \ell \nu \): \(|\Delta E| < 1.0 \text{ GeV} \) and \( p^*_\ell > 1.2 \text{ GeV} \). The efficiencies of the \( B \) → \( \eta \ell \nu \) MC and data samples for each cut can be seen in Table 6.2, and those for the fit region are shown in Table 6.3. The S/B is slightly better for each cut than that achieved in the \( \pi^0 \) analysis (compare Table 6.2 and Table 4.9), but these tables show efficiencies with respect to each sample after skimming and \( B \) candidate selection. At that point, the \( B \) → \( \pi^0 \ell \nu \) signal efficiency is 17% with respect to the unskimmed \( B \) → \( \pi^0 \ell \nu \) sample, while the \( B \) → \( \eta \ell \nu \) signal efficiency is 3.8% with respect to the unskimmed \( B \) → \( \eta \ell \nu \) sample (remember

---

Table 6.1: The optimised set of analysis cuts (event selection criteria), for \( B \) → \( \eta \ell \nu \). Those in bold use different threshold values from the \( B \) → \( \pi^0 \ell \nu \) study.

<table>
<thead>
<tr>
<th>Quantity (units)</th>
<th>Analysis cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of leptons</td>
<td>( x = 1 )</td>
</tr>
<tr>
<td>( \cos \theta_\ell )</td>
<td>(-0.86 &lt; x &lt; 0.96)</td>
</tr>
<tr>
<td>Net charge (</td>
<td>( e</td>
</tr>
<tr>
<td>( M^2_{\text{miss}} ) (GeV(^2))</td>
<td>(-4.0 &lt; x &lt; 6.5)</td>
</tr>
<tr>
<td>( p^*_\ell ) (GeV)</td>
<td>( x &gt; 1.2)</td>
</tr>
<tr>
<td>( R_2 )</td>
<td>( x &lt; 0.45)</td>
</tr>
<tr>
<td>( \cos \theta_{\text{BY}} )</td>
<td>(-1.05 &lt; x &lt; 1.30)</td>
</tr>
<tr>
<td>( \cos \theta_{\gamma \gamma} )</td>
<td>( x &gt; 0.7)</td>
</tr>
<tr>
<td>( \cos \theta_{\text{thrust}} )</td>
<td>(-0.80 &lt; x &lt; 0.85)</td>
</tr>
<tr>
<td>( \cos \theta_{\text{Wf}} )</td>
<td>(-0.70 &lt; x &lt; 0.85)</td>
</tr>
<tr>
<td>( M_{bc} ) (GeV)</td>
<td>( 5.26 &lt; x &lt; 5.29)</td>
</tr>
<tr>
<td>( \Delta E ) (GeV)</td>
<td>(</td>
</tr>
</tbody>
</table>
that the $\eta \to \gamma\gamma$ branching fraction is only 39%; see Table 4.7).

Table 6.2: The effect of each cut on the efficiency of each sample. The final row shows the effect of all of the cuts put together. Cuts are not cumulative, except for the final row. The efficiencies listed are with respect to each sample after the lepton skim, event reconstruction, and $B$ candidate selection, not with respect to the complete sample. All values are percentages except for S/B.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Signal</th>
<th>Crossfeed</th>
<th>$B^0\bar{B}^0$</th>
<th>$B^+B^-$</th>
<th>uds</th>
<th>charm</th>
<th>Bkgd sum</th>
<th>S/B $\times 10^3$</th>
<th>On-res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no cuts]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>1.50</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>$\cos \theta_\nu$</td>
<td>92.87</td>
<td>92.28</td>
<td>87.59</td>
<td>87.39</td>
<td>69.09</td>
<td>79.66</td>
<td>83.65</td>
<td>1.66</td>
<td>81.54</td>
</tr>
<tr>
<td>Net Charge</td>
<td>82.87</td>
<td>77.92</td>
<td>68.43</td>
<td>73.54</td>
<td>77.14</td>
<td>78.66</td>
<td>73.56</td>
<td>1.69</td>
<td>72.08</td>
</tr>
<tr>
<td>$M^2_{miss}$</td>
<td>72.34</td>
<td>72.35</td>
<td>67.24</td>
<td>70.78</td>
<td>79.16</td>
<td>72.62</td>
<td>71.17</td>
<td>1.52</td>
<td>69.86</td>
</tr>
<tr>
<td>$p_T^*$</td>
<td>99.09</td>
<td>99.54</td>
<td>99.50</td>
<td>99.52</td>
<td>98.01</td>
<td>98.86</td>
<td>99.19</td>
<td>1.50</td>
<td>98.92</td>
</tr>
<tr>
<td>$R_2$</td>
<td>98.52</td>
<td>98.82</td>
<td>99.44</td>
<td>99.39</td>
<td>73.93</td>
<td>76.44</td>
<td>91.48</td>
<td>1.61</td>
<td>89.80</td>
</tr>
<tr>
<td>$\cos \theta_{BY}$</td>
<td>81.81</td>
<td>30.61</td>
<td>17.95</td>
<td>17.60</td>
<td>19.29</td>
<td>17.54</td>
<td>18.18</td>
<td>2.06</td>
<td>18.06</td>
</tr>
<tr>
<td>$\cos \theta_{\gamma\gamma}$</td>
<td>57.57</td>
<td>14.83</td>
<td>15.62</td>
<td>15.94</td>
<td>33.07</td>
<td>37.99</td>
<td>22.49</td>
<td>3.84</td>
<td>22.27</td>
</tr>
<tr>
<td>$\cos \theta_{thrust}$</td>
<td>80.63</td>
<td>73.30</td>
<td>73.85</td>
<td>73.66</td>
<td>28.08</td>
<td>28.96</td>
<td>58.80</td>
<td>2.06</td>
<td>58.46</td>
</tr>
<tr>
<td>$\cos \theta_{W\ell}$</td>
<td>93.37</td>
<td>86.24</td>
<td>84.47</td>
<td>85.02</td>
<td>63.47</td>
<td>77.31</td>
<td>80.36</td>
<td>1.74</td>
<td>80.57</td>
</tr>
<tr>
<td>$M_{bc}$</td>
<td>33.67</td>
<td>9.56</td>
<td>9.63</td>
<td>9.83</td>
<td>7.26</td>
<td>6.13</td>
<td>8.68</td>
<td>5.81</td>
<td>7.82</td>
</tr>
<tr>
<td>$\Delta E$</td>
<td>98.45</td>
<td>91.71</td>
<td>81.11</td>
<td>80.58</td>
<td>85.69</td>
<td>86.06</td>
<td>82.74</td>
<td>1.78</td>
<td>83.51</td>
</tr>
<tr>
<td>All cuts</td>
<td>10.77</td>
<td>0.21</td>
<td>0.05</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
<td>0.07</td>
<td>323.81</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 6.3: Signal, background and on-resonance data efficiencies in the fit region, $|\Delta E| < 1.0$ GeV and $p_T^* > 1.2$ GeV, after the lepton skim, event reconstruction and $B$ candidate selection, but before any neutrino reconstruction or analysis cuts have been applied. All values listed are percentages except for S/B.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>97.61</td>
</tr>
<tr>
<td>Crossfeed</td>
<td>91.30</td>
</tr>
<tr>
<td>$B^0\bar{B}^0$</td>
<td>80.75</td>
</tr>
<tr>
<td>$B^+B^-$</td>
<td>80.23</td>
</tr>
<tr>
<td>uds</td>
<td>84.23</td>
</tr>
<tr>
<td>charm</td>
<td>85.23</td>
</tr>
<tr>
<td>Bkgd sum</td>
<td>82.15</td>
</tr>
<tr>
<td>On-res.</td>
<td>82.74</td>
</tr>
</tbody>
</table>
6.2 The $B \rightarrow \eta \ell \nu$ Signal

The $B \rightarrow \eta \ell \nu$ signal efficiency was defined in the same way as that of $B \rightarrow \pi^0 \ell \nu$: the number of events in the signal histogram which was inputted into the fitting program, relative to the number of $B \rightarrow \eta \ell \nu$ events in the full b2ulnu sample \(^1\). After all selection criteria had been applied, the $B \rightarrow \eta \ell \nu$ signal efficiency was \((0.31 \pm 0.01)\)%. This is rather lower than the $B \rightarrow \pi^0 \ell \nu$ efficiency of \((1.72 \pm 0.01)\)%.

The branching fraction of $\eta \rightarrow \gamma \gamma$ is 39%; that of $\pi^0 \rightarrow \gamma \gamma$ is 99% \([20]\). As this analysis only reconstructed the $\eta$ in its $\gamma \gamma$ mode, the reconstruction efficiency can already be expected to be roughly 40% of the efficiency obtained for $B \rightarrow \pi^0 \ell \nu$. Additionally, the $\gamma \gamma$ combinatorial background is considerably worse for reconstructed $\eta$s than for $\pi^0$s (see Sec. 4.4). In order to battle this background, the minimum photon energy threshold was set at 350 MeV for $\eta$ reconstruction. This sacrificed some efficiency, but dramatically increased the signal to noise (see Table 4.4).

The same fitting procedure and algorithm were used for $B \rightarrow \eta \ell \nu$ as were used in the $B \rightarrow \pi^0 \ell \nu$ analysis, and the branching fraction was calculated in the same manner. The $B \rightarrow \eta \ell \nu$ fit result projections and fitted yields are presented in Fig. 6.7 and Table 6.4, respectively. The preliminary $B \rightarrow \eta \ell \nu$ branching fraction was found to be

$$B(B^+ \rightarrow \eta \ell^+ \nu) = (0.42 \pm 0.13) \times 10^{-4}$$

The error is statistical only.

The $b \rightarrow u \ell \nu$ crossfeed has a fitted yield of $0.0 \pm 734.7$, and its histogram (black) is absent from the fit projection histograms. As was described in Sec. 5.1, the TFractionFitter algorithm adjusts the weights of the input signal and background histograms, attempting to match the shape of their weighted sum to the shape of the input data histogram. In this case, the fitting program determined that the crossfeed histogram's shape did not help the signal and other background histograms to match the on-resonance data histogram. The large error on the crossfeed yield indicates that the shape discrimination for this background was poor.

\(^1\)There were 936,860 $B \rightarrow \eta \ell \nu$ events in the sample, of which 369,877 were $B \rightarrow \eta \ell \nu$, $\eta \rightarrow \gamma \gamma$. 

Figure 6.7: $\Delta E$ (left) and $p^*_\ell$ (right) projections of the 2D fit result for $B \rightarrow \eta \ell \nu$. Both the $\Delta E$ and $p^*_\ell$ axes are in units of GeV.
Table 6.4: Fitted signal and background yields, the Signal to Background ratio (S/B), and the $\chi^2$/ndf goodness-of-fit for the $B \to \eta \ell \nu$ fit.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Crossfeed</th>
<th>$BB$</th>
<th>Continuum</th>
<th>S/B</th>
<th>$\chi^2$/ndf</th>
</tr>
</thead>
<tbody>
<tr>
<td>137.0 ± 44.5</td>
<td>0.0 ± 734.7</td>
<td>641.0 ± 54.0</td>
<td>208.0 ± 41.0</td>
<td>0.161</td>
<td>224/396</td>
</tr>
</tbody>
</table>

The statistics for $B \to \eta \ell \nu$ are considerably lower than those for $B \to \pi^0 \ell \nu$, and the $\cos \theta_{\gamma\gamma}$ cut, which dramatically increased the S/B, removed almost all events in the region $q^2 > 14$ (high $q^2$ corresponds to low hadron momentum, as does low $\cos \theta_{\gamma\gamma}$). Accordingly, it was not possible to extract a $B \to \eta \ell \nu$ branching fraction in $q^2$ bins.

6.3 Discussion of Systematic Uncertainties

The procedure used in the $B \to \pi^0 \ell \nu$ analysis was applied to $B \to \eta \ell \nu$ because the systematics are expected to be similar for the $\eta \to \gamma \gamma$ decay mode. In particular, systematic errors due to the neutrino reconstruction efficiency, lepton track finding efficiency, and lepton identification should be the same for $B \to \eta \ell \nu$ and $B \to \pi^0 \ell \nu$. The Belle $\pi^0$ reconstruction efficiency study compared data and MC $\pi^0$ to data and MC $\eta$s reconstructed in the three dominant modes ($\gamma\gamma$, $\pi^+\pi^-\pi^0$, $\pi^0\pi^0\pi^0$) [73, 74]. The $\eta \to \gamma \gamma$ reconstruction systematics cannot be estimated from that study, and would need to be examined in its own right. This is left for future work.

The effects of event selection criteria and assumed $b \to u \ell \nu$ and $b \to c \ell \nu$ branching fractions are expected to be at least as great for $B \to \eta \ell \nu$ as for $B \to \pi^0 \ell \nu$. The $B \to \eta \ell \nu$ efficiency is considerably lower, and the $\cos \theta_{\gamma\gamma}$ histograms in Fig. 6.4 in particular illustrate that $B \to \eta \ell \nu$ is sensitive to variations in analysis cut thresholds. The absence of the crossfeed histogram in the fit projections (Fig. 6.7) and the very large error on its signal yield (see Table 6.4) indicate that this background is not well controlled, and the $B\overline{B}$ background’s prominence in the fit results suggests that a change in $b \to c \ell \nu$ branching fraction normalisations would perceptibly affect the $B \to \eta \ell \nu$ measurement.

The uncertainty in the number of $B\overline{B}$ events in 492 fb$^{-1}$ of Belle data was easily calculated, and resulted in a systematic error of $\pm 0\%$ in the $B \to \eta \ell \nu$ branching fraction.

6.4 Comparison with Other Results

The CLEO Collaboration presented a $B \to \eta \ell \nu$ branching fraction result at the ICHEP conference in August 2006 [24]. At the more recent CKM06 workshop, however, CLEO superceded the ICHEP result with an upper 90% confidence limit, while BaBar presented a branching fraction with a nearly $3\sigma$ precision (see Table 6.5). There are no published $|V_{ub}|$ extractions from $B \to \eta \ell \nu$ branching fractions.

\footnote{This effect is not as important in the $B \to \pi^0 \ell \nu$ study, as the $\cos \theta_{\gamma\gamma}$ cut removes so little $B \to \pi^0 \ell \nu$ signal (see Fig. 6.4).}

\footnote{This situation would likely be improved by an increase in MC statistics.}
Table 6.5: The preliminary $B \rightarrow \eta \ell \nu$ branching fraction measured in this analysis, compared with the most recent public $B \rightarrow \eta \ell \nu$ results. Charge conjugation is implied in all cases. The first, second and third errors are statistical, experimental systematic, and due to model uncertainty, respectively. The CKM06 results were presented in December 2006, and supercede the HFAG values, which are based on results presented at the ICHEP conference in August 2006.

<table>
<thead>
<tr>
<th>Mode</th>
<th>BF (×10^4)</th>
<th>Experiment</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>0.42 ± 0.13</td>
<td>Belle</td>
<td>Neutrino Recon.</td>
<td>This analysis</td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>0.84 ± 0.27 ± 0.21</td>
<td>BaBar</td>
<td>Neutrino Recon.</td>
<td>CKM06 ([77])</td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>0.84 ± 0.27 ± 0.21</td>
<td>BaBar</td>
<td>Neutrino Recon.</td>
<td>HFAG ([24])</td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>0.44 ± 0.23 ± 0.11 ± 0.00 (≤ 1.01 × 10^{-4} at 90% CL)</td>
<td>CLEO</td>
<td>Neutrino Recon.</td>
<td>CKM06 ([81])</td>
</tr>
<tr>
<td>$B^+ \rightarrow \eta \ell^+ \nu$</td>
<td>0.84 ± 0.31 + 0.16 ± 0.09</td>
<td>CLEO</td>
<td>Neutrino Recon.</td>
<td>HFAG ([24])</td>
</tr>
</tbody>
</table>

The preliminary $B \rightarrow \eta \ell \nu$ branching fraction measured in this analysis was

$$B(B \rightarrow \eta \ell \nu) = (0.42 \pm 0.13) \times 10^{-4}$$

The error is statistical only.

The experimental determination of $B \rightarrow \eta \ell \nu$ is clearly not well established. As can be seen in Table 6.5, the $B \rightarrow \eta \ell \nu$ branching fraction obtained in this analysis is competitive, and in agreement, with publicly presented results. The statistical error on the branching fraction is lower by far than that achieved by any of the other studies.

The results presented in this chapter have shown that it is possible to measure the $B \rightarrow \eta \ell \nu$ branching fraction using Belle data. Prospects for a complete $B \rightarrow \eta \ell \nu$ analysis are discussed in the following chapter.
Chapter 7

Discussion and Further Work

These analyses found the preliminary $B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \eta \ell \nu$ branching fractions to be

$$B(B \rightarrow \pi^0 \ell \nu) = (0.68 \pm 0.09 \pm 0.11 \pm 0.04) \times 10^{-4}$$

and

$$B(B \rightarrow \eta \ell \nu) = (0.42 \pm 0.13) \times 10^{-4}$$

with signal efficiencies of $(1.72 \pm 0.01)\%$ and $(0.31 \pm 0.01)\%$, respectively. The $B \rightarrow \pi^0 \ell \nu$ branching fraction was used to extract a preliminary value for the magnitude of the CKM matrix element $V_{ub}$:

$$|V_{ub}| = (3.29 \pm 0.23 \pm 0.27 \pm 0.05) \times 10^{-3}$$

The errors are statistical only for the $\eta$ branching fraction, and statistical, preliminary experimental systematic, and preliminary theoretical, respectively, for the $\pi^0$ measurements.

The central values of the $B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \eta \ell \nu$ branching fractions are in agreement with published results, as was shown in Tables 5.8 and 6.5. The 15% statistical error on the $B \rightarrow \pi^0 \ell \nu$ branching fraction and the 31% statistical error on the $B \rightarrow \eta \ell \nu$ branching fraction are considerably lower than those of any publicly presented or published results by other investigators. The $B \rightarrow \pi^0 \ell \nu$ branching fraction measurement is now systematics-limited.

These analyses had their limitations. The $B \rightarrow \pi^0 \ell \nu$ systematic study was preliminary, and the $b \rightarrow u \ell \nu$ modelling study examined only two models, LCSR and ISGW2, as these were the only $b \rightarrow u \ell \nu$ MC samples available. These could be expanded by, for example, investigating the systematic effects of each of the analysis cuts, and generating and examining $b \rightarrow u \ell \nu$ MC samples based on other model predictions. The introduction of a $p$ veto (vetoing signal $\pi^0$s if the combined invariant mass of the $\pi^0$ and any charged pion in the event is consistent with the $\rho^+$ mass) might enhance the signal-to-background of the $B \rightarrow \pi^0 \ell \nu$ measurement.

The $B \rightarrow \eta \ell \nu$ analysis demonstrated that there is a definite $B \rightarrow \eta \ell \nu$ signal in the Belle dataset, but no detailed systematic nor modelling study was conducted. Additionally, the use of the cos $\theta_{\gamma\gamma}$ analysis cut removed most of the high-$q^2$ $B \rightarrow \eta \ell \nu$ events. Because of this, it was not possible to measure the $B \rightarrow \eta \ell \nu$ branching fraction in $q^2$ bins. Perhaps a future study could develop an analysis cut which would increase the signal-to-background as the cos $\theta_{\gamma\gamma}$ does, but without affecting the $q^2$ distribution.

Possibly the most obvious expansion of the $\eta$ analysis is to include the $\eta \rightarrow \pi^0 \pi^0 \pi^0$ and $\eta \rightarrow \pi^+ \pi^- \pi^0$ modes, which have branching fractions of 32.5% and 22.7%, respectively. These modes will have different systematics than $\eta$ (or $\pi^0$) $\rightarrow \gamma \gamma$, which will need to be investigated. For the
\[ \eta \rightarrow \gamma \gamma \] mode, the use of a \( \pi^0 \) veto (vetoing signal \( \eta s \) if any of their daughter photons, when combined with another photon in the event, would be consistent with a \( \pi^0 \)) should be examined.

Alternatively (or additionally), the proposed Super-B-Factory with its design luminosity of order \( 10^{35} \text{cm}^{-2}\text{s}^{-1} \), and the “SuperBelle” upgrade to the Belle detector, will produce an on-resonance data sample of order ten times the size of that used in this analysis [15]. With enhanced statistics, it should be possible to extract the \( B \rightarrow \eta \ell \nu \) branching fraction in three or possibly more \( q^2 \) bins using the neutrino reconstruction technique described in this thesis, reducing the model-dependency of the \( B \rightarrow \eta \ell \nu \) measurement. Though the measured \( B \rightarrow \pi^0 \ell \nu \) branching fraction is systematics-limited, additional statistics would improve the measurement in the \( q^2 < 8 \text{ GeV}^2 \) bin, and would allow for finer binning.

This work has examined two instances of the rare, weak \( b \rightarrow u \) transition, and has provided a constraint on the CKM matrix element \( V_{ub} \). Rare decays such as \( B \rightarrow \pi^0 \ell \nu \) and \( B \rightarrow \eta \ell \nu \) can only be measured with large statistical datasets, and Belle has achieved the highest peak and integrated luminosities of any particle accelerator to date. High luminosity experiments such as Belle and BaBar are pushing the bounds of the “luminosity frontier” ever farther, enabling precision measurements of rarer and rarer decays. The further we probe into this exotic realm, the closer we come to seeing what lies beyond the Standard Model.
Bibliography


B. Aubert et al. (BaBar Collab.) Study of $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho \ell \nu$ Decays and Determination of $|V_{ub}|$. *Phys. Rev. D*, 72:051102, 2005.


S. Cole, N. Parslow, and K. Varvell. A Study of $B \rightarrow \pi^0 \ell \nu$ and $B \rightarrow \pi^+ \ell \nu$ Decays Using Belle Data. *Belle Note 701*, 2004.


[77] B. Viaud. Exclusive $B \to X_u\ell\nu$ Decays with BaBar. Presented at the 4th International Workshop on the CKM Unitarity Triangle (CKM2006). Available at http://agenda.hepl.phys.nagoya-u.ac.jp/getFile.py/access?contribId=12&amp;sessionId=4&amp;resId=0&amp;materialId=sldes &amp;confId=2.


[79] J. Shigemitsu et al. Semileptonic $D \to \pi/K$ and $B \to \pi/D$ Decays in a 2+1 Flavor Lattice QCD. Talk presented at Lattice2004, Available at hep-lat/0409116.


[81] S. Stone. Exclusive Semileptonic $b \to u\ell\nu$ Decays at CLEO. Presented at the 4th International Workshop on the CKM Unitarity Triangle (CKM2006). Available at http://agenda.hepl.phys.nagoya-u.ac.jp/getFile.py/access?contribId=13&amp;sessionId=4&amp;resId=0&amp;materialId=sldes &amp;confId=2.