A Study of $\tau^- \to \omega\pi^-\nu_\tau$ Using OPAL at LEP

Elizabeth A. Vokurka

THE UNIVERSITY of MANCHESTER

High Energy Group

Department of Physics and Astronomy

October 25, 1998

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering.

Contents

Abstract **8**

Declaration and Author **9**

10

1 Introduction **11**

1.1 Why the ω? 12
1.2 Contents of the Following Text 13

2 Theory **14**

2.1 The τ 14
2.1.1 V-A Interactions 15
2.1.2 Semi-leptonic Decay 15
2.1.3 $\langle \pi^-\pi^+\pi^-\pi^0 \mid V_\mu - A_\mu \mid 0 \rangle$ 17
2.2 G-parity 19
2.3 Vital Statistics of the ω 21
2.3.1 Quark Mixing and Composition 22
2.3.2 ω Decay Products 23
2.4 Second Class Currents 23
2.4.1 Impact on the Standard Model 26

3 Previous ω Studies **27**

3.1 ARGUS 28
3.2 CLEO 28
3.3 ALEPH 29
3.4 Justification for Another Measurement 29

4 The Experiment **31**

4.1 LEP 31
4.2 OPAL 33

4.3 OPAL Coordinate System . . . 34
4.4 Central Tracking System . . . 34
4.4.1 Silicon Microvertex Detector (SI) . . . 34
4.4.2 Vertex Detector (CV) . . . 35
4.4.3 Jet Chamber (CJ) . . . 35
4.4.4 Z-Chambers (CZ) . . . 35
4.5 Time-of-Flight Counters (TOF) . . . 36
4.6 Electromagnetic Calorimeter (ECAL) . . . 36
4.6.1 Electromagnetic Presampler (PB and PE) . . . 36
4.6.2 Lead Glass Calorimeter (EB and EE) . . . 37
4.7 Hadron Calorimeter (HB, HE, and HP) . . . 38
4.8 Muon Detector (MB and ME) . . . 39
4.9 Forward Detector (FD) . . . 39
4.10 Silicon Tungsten Detector (SW) . . . 39

5 Event Simulation **41**
5.1 KORALZ . . . 41
5.2 TAUOLA . . . 42
5.3 GOPAL . . . 43
5.4 Simulation Specifics . . . 43
5.5 Where ARE the ω's? . . . 43

6 Event Selection **46**
6.1 Online Selection . . . 46
6.1.1 Low Multiplicity . . . 47
6.2 Tau Pair Selection . . . 47
6.2.1 The Tau Platform . . . 48
6.2.2 Selection in Analysis . . . 48
6.3 Primed for Fitting . . . 50
6.3.1 Useful Variables . . . 50
6.3.2 π^0 Identification . . . 52
6.4 The Problem . . . 53

7 Statistical Methods **65**
7.1 The Principle of Least Squares . . . 65
7.1.1 Fitting the Data . . . 66
7.2 Maximum Likelihood . . . 68
7.2.1 Determining the Likelihood . . . 68
7.2.2 Maximising the Likelihood . . . 69
7.2.3 Results of the Basic Maximum Likelihood Method . . . 70

7.2.4 Maximum Likelihood Limitations 72
7.3 Improved Maximum Likelihood 74
7.3.1 Results of Improved Maximum Likelihood 76
7.3.2 Possible Improvements on the Improvement 76
7.4 Neural Networks 77
7.4.1 Brain Power 78
7.4.2 Structure of the Net 78
7.4.3 Optimising Weights 79
7.4.4 Neural Nets in Practice 80
7.4.5 Effect on Improved Maximum Likelihood 81
7.5 Fisher Discriminants 82

8 Evaluation of the Branching Ratio 90
8.1 Determination of the Branching Ratio 90
8.2 Best Fitting Method 92
8.2.1 Linear Discrimination 92
8.2.2 Minimised χ^2 93
8.2.3 Uncorrected Maximum Likelihood 93
8.2.4 Corrected Maximum Likelihood 93
8.2.5 Neural Net Inputs 93
8.2.6 The Verdict 94

9 Results and Conclusions 95
9.1 The Branching Ratio 95
9.1.1 Systematic Errors 95
9.2 Potential for Second-Class Current Limit 97
9.3 Impact of Result 97
9.4 Further Study 98
9.4.1 Possible Hope for the Future 98

A τ Pair Selection 99

B Maximum Entropy 104
B.1 Bayesian Probability Calculus 104
B.2 Application to Image Reconstruction 105
B.2.1 The Conditional Probability of the Data 105
B.2.2 Prior Probability 106
B.3 Reconstruction of Clusters in the ECAL 107
B.4 An Alternative Method 110

C Linear Discrimination **111**
C.1 Discrimination 111
C.1.1 Simple 2 Distribution Case 111
C.1.2 The Linear Case 114
C.2 Multiple Distributions 115
C.3 Application of Linear Discriminants 116
C.3.1 Counting Techniques 116
C.3.2 Fitting Techniques 117
C.4 An Inappropriate Choice 117

References **119**

Acknowledgments **125**

List of Tables

2.1 Principal ω decay products 23

3.1 Previous $\tau \to \omega\pi\nu$ measurements 27

6.1 Cumulative effects of cuts on Monte Carlo 54
6.2 Effect of cuts on data 55
6.3 Constraints on poorly measured events 59
6.4 Background composition 60

7.1 Basic maximum likelihood fit 72
7.2 Effect of variable correlation 73
7.3 Corrected maximum likelihood fit results 77
7.4 The tagging efficiency of the net 81
7.5 Neural net basic maximum likelihood fit 82

8.1 Estimates of the proportion of ω decays. 91
8.2 Branching ratio estimates 92

C.1 Counting method 117
C.2 Fitting method 118

List of Figures

2.1 $\tau \rightarrow \nu_\tau + X$ through the weak current. 16

4.1 CERN's accelerator array. 32
4.2 A breakdown of OPAL's subdetectors. 33

5.1 $\pi^+\pi^-\pi^0$ mass combination in closest agreement with M_ω 44
5.2 The two mass combinations for the $3\pi\pi^0$ sample 45

6.1 One unassigned cluster removes a significant amount of background. . 56
6.2 The highest energy unassigned neutral cluster 57
6.3 A 3-prong $e^+e^- \rightarrow \tau^+\tau^-$ event. 58
6.4 Data compared with MC. 61
6.5 The overall energy deposited in the ECAL 62
6.6 MC $3\pi\pi^0$ *best* guess for the $\pi^+\pi^-$ combination. 63
6.7 Little, if any, distinguishable peak from the ω resonance. 64

7.1 % of decays from the least squares fit for each variable. 83
7.2 χ^2 from the least squares fit for each variable. 84
7.3 Standard distributions and their Gaussian mappings. 85
7.4 Neural net schematic. 86
7.5 Neural net feature extraction of the ω resonance 87
7.6 Outputs for the trained neural net 88
7.7 The 1994 data agrees with the MC net output. 89

B.1 Maximum entropy clustering for an isolated π^0. 108

C.1 The linear discriminant 112

Abstract

A study of $\tau^- \rightarrow \omega\pi^-\nu_\tau$ was performed using the 1994 data from the OPAL detector at CERN. Four independent statistical methods were examined to determine the best estimate of $Br(\tau^- \rightarrow \omega\pi^-\nu_\tau)$. A corrected maximum likelihood with neural net enhanced kinematic features as the probability density functions has been shown to be the most statistically robust fitting algorithm with an estimate of $Br(\tau^- \rightarrow \omega\pi^-\nu_\tau) = 2.02 \pm 0.32$ %. Due to the resolution of the detector's electromagnetic calorimeter, no sample of pure $\tau^- \rightarrow \omega\pi^-\nu_\tau$ was accessible. This removes the possibility of identifying second class currents at OPAL.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

The Author

The author was educated at the University of Illinois in Urbana-Champaign and the University of Manchester before joining the high energy physics group at the University of Manchester in 1994. The work presented here was conducted at the University of Manchester and CERN, Geneva.

Experimental particle physicists' ultimate goal is not to produce any marketable good or service. If they were interested in promoting their own welfare, they would become merchant bankers. Physicists have taken over the mantle of cloistered monks. They cultivate and preserve a highly technical, distressingly misunderstood, and (possibly) somewhat worthless collection of theories. Each hopes that some day, someone, somewhere will appreciate the beauty of these theories and the experimental ingenuity that has gone into confirming/debunking them. And eventually, these bits of information may even advance humanity beyond our wildest dreams. Gregor Mendel, we take our hats off to you and your peas.

But in this day and age, peas are small potatoes...

Chapter 1

Introduction

For years, particle physics has endeavoured to explain physical phenomena in terms of properties and interactions of a small number of particles. In 1930, Pauli began the journey toward understanding the fundamental constituents of matter by introducing the possibility of an electron neutrino [1]. The idea of electrons orbiting a nucleus of protons and neutrons was no longer sufficient to account for the nature of microscopic processes like β-decay. By the 1970's, things were getting a bit out of hand as scores of "fundamental" particles were being discovered willy-nilly.

Fortunately, years before this threat to simplicity began, three men formulated a theory to explain interactions via vector fields. This was the basis of a "standard model" that has yet to be significantly shaken by the physics community. Glashow [2] proposed an $SU(2)_L \times U(1)_Y$ gauge symmetry [3] to describe electroweak interactions while maintaining partial symmetry. Weinberg and Salam [4] then enhanced this theory to accommodate the present-day massive vector bosons, the $W^\pm$ and Z^0. In order to extend this theory beyond leptonic electroweak interactions, the GIM mechanism [5], Cabbibo mixing [6], and colour [7] have been added to explain hadronic interactions.

Although this theory has held firm for well over 20 years, there are still unanswered questions and untested hypotheses. There is still debate as to whether neutrinos are massive, and therefore able to oscillate from one mass eigenstate to another. The mass of the Higgs boson is unmeasured, and its existence remains unproven, leaving the questions of the precise model of the Higgs mechanism, the origin of mass, and the renormalisability of the electroweak gauge theory all open ended. A "grand unifying group" G is still speculative, and a decaying proton has not been observed to give the simplest model, $SU(5)$ GUT, more credibility. The composition and existence of dark matter is as much a mystery as the apparent lack of antimatter in the universe. The Standard Model, notorious for its sturdiness, has a long way to go to completely and elegantly describe the macro and microscopic. With the increased sensitivity of experiments and the improvement of experimental techniques, attempts to solidify the mathematical basis of the universe continue to gain speed. We have reached a point in experimentation that calls for high precision measurements to determine the reliability of the Standard Model.

1.1 Why the ω?

At the higher energies, searches are being performed to look for the Higgs boson and supersymmetric particles. Their discovery would probably help to strengthen existing theory. Yet most present studies attempt to break the Standard Model. They try to find discrepancies that the accepted theory fails to explain. Searches for exotic decays like R-parity violating SUSY particles are conducted alongside the more conventional searches at higher energies. Present experiments obtain high precision measurements such as $(g-2)$ of the muon for high order electroweak corrections. Precision R_b measurements also study high order corrections while being sensitive to the top mass. The search for forbidden decays is the final method of probing the Standard Model. If a theory does not allow certain decays to occur, and they are proven to exist, then the theory must be altered or even abandoned.

The goal of the study of $\tau \rightarrow \omega\pi\nu_\tau$ is ultimately a study of the forbidden decay, $\tau \rightarrow b_1\nu_\tau$. The b_1 resonance, with a mass of 1231$\pm$10 MeV [14], while allowed in strong decays, is not allowed in this semileptonic decay due to G-parity violation described in the next chapter. New methods of analysis and improvements in experimental modelling have made it possible for the OPAL collaboration to determine whether a competitive analysis of the $\tau^- \rightarrow \omega\pi^-\nu_\tau$ decay is possible. Heightened knowledge of the properties of this interaction can be used as a test of the Standard Model as shown in Chapter 2.

1.2 Contents of the Following Text

The following pages contain an exposé of the attempts to obtain as much information as possible from the $\tau \rightarrow \omega\pi\nu$ resonance using the OPAL detector at CERN. Ultimately, a branching ratio is the only accessible information, since an extraction of a pure ω sample is rendered impractical due to the limitations of detector design. This study is more concerned with the power of statistical techniques than the precise measurement of a branching ratio. With the next generation of detector highly dependent on complex pattern recognition for physics analyses, the limits of statistical manipulations must be pushed at the same rate as instrumental precision.

Chapter 2

Theory

The reader is spared a detailed description of the Standard Model. If you are looking for one, please consult [2], [4], [5], [6], [7], and [8] (or someone else's thesis). Instead, the emphasis of this chapter will be on the τ lepton's role in fundamental particle physics and the ω's impact on τ physics.

2.1 The τ

The τ is a 3rd generation sequential lepton. While almost 3500 times more massive, it interacts similarly to an electron due to lepton universality.

The τ is the unsung hero of meson resonance studies. While τ's are heavy enough to decay semi-leptonically, they do not suffer from the large non-resonant effects present in hadron collisions due to limited phase space. Over 99% of τ decays contain 1 or 3 charged particles because of the τ's relatively small mass (1777 MeV) [9].

2.1.1 V-A Interactions

The weak charged current is carried by the W^+ and W^- which couple to fermions by V-A, left-handed, interactions only. The current consists of a proper vector, V, and an axial-vector, A. There are other combinations of the vector and axial couplings to choose from [10], but V-A triumphed in a "classic and beautiful" experiment by Goldhaber *et al.* that determined the ν_e helicity from γ-rays emitted by resonant excited β-decaying ^{152}Sm nuclei produced by β-decaying ^{152}Eu [11]. For the tau, V-A coupling can be shown experimentally through the measurement of the Michel parameter, ρ [12] (or by assuming lepton universality).

$$\rho = \frac{3}{8}\frac{(V-A)^2}{V^2+A^2} \tag{2.1}$$

Measure this Michel parameter accurately (e.g. by using the shape of the electron spectrum from τ decays [9]), and you know the composition of the weak charged current...pure V-A.

2.1.2 Semi-leptonic Decay

A detailed account of the Feynman rules [13] for electroweak interactions shall be dispensed with. Feynman rules are the basis for a simple, symbolic method of calculating the perturbative expansion of scattering amplitudes in relativistically covariant perturbation theory. These rules can be implemented to construct the amplitude that accounts for decays such as $\tau^- \rightarrow \omega\pi^-\nu_\tau$. Using the V-A form of the weak current mentioned above, the weak interaction amplitude [3] comes in the form

$$\mathcal{M} = \frac{4G}{\sqrt{2}} J^{(\tau)\mu} J^\dagger_{(X)\mu}. \tag{2.2}$$

G is the weak coupling or Fermi constant. For leptons, $G = (1.16639 \pm 0.00002) \times 10^{-5} GeV^{-2}$ taking $\hbar c$ as unity [14]. J^{μ} is a charge-raising weak current, while its Hermitian conjugate, $J_{\mu}^{\dagger}$, lowers the charge. Figure 2.1 shows the Feynman diagram of a τ decaying semi-leptonically into a ν_{τ} and a final state, X.

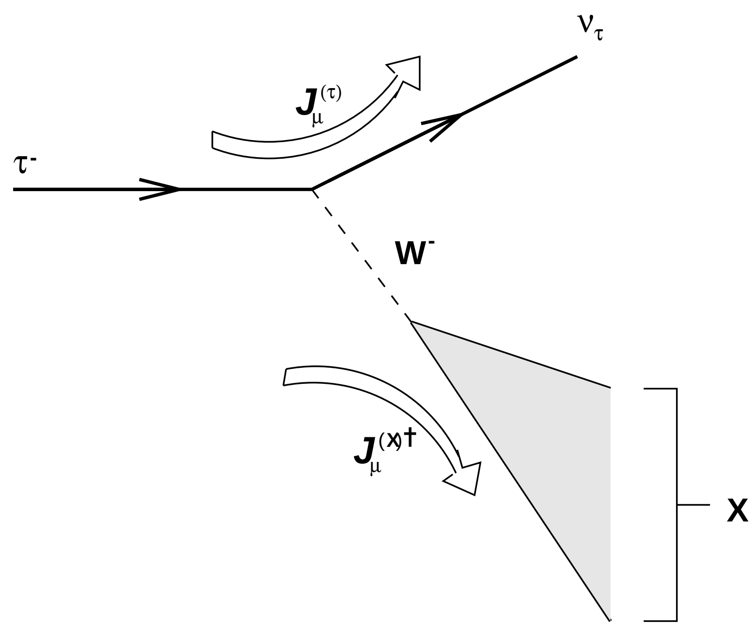

Figure 2.1: A Feynman diagram of the semi-leptonic decay $\tau \rightarrow \nu_{\tau} + X$ through the weak current.

$J^{(\tau)\mu}$ takes the form

$$J^{(\tau)\mu} = \psi_{\nu_{\tau}} \gamma_{\mu} \frac{1}{2}(1 - \gamma^{5}) \psi_{\tau}. \tag{2.3}$$

ψ is the wavefunction for a specific incoming or outgoing particle. $\frac{1}{2}(1 - \gamma^{5})$ is the normalised projection operator summing over spins with $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, a product of the four 4×4 Dirac γ-matrices, to condense notation. This operator ensures that parity is violated maximally, unlike the electromagnetic current.

The charge lowering weak current, $J_\mu^{(X)\dagger}$, is the matrix element of the V-A current for the specific final state X, $\langle X \mid V_\mu - A_\mu \mid 0\rangle$ [15]. $J_\mu^{(X)\dagger}$ is the only component that changes when calculating the matrix element for the range of semileptonic tau decays. By replacing X in the expectation value with the desired τ decay channel, its decay rate can be calculated keeping the rest of the matrix element constant.

2.1.3 $\langle \pi^-\pi^+\pi^-\pi^0 \mid V_\mu - A_\mu \mid 0\rangle$

$J_\mu^\dagger$ for the $3\pi\pi^0$ state in τ decay can be split into a sum of two parts.

$$J_\mu^\dagger = J_\mu^{I\dagger} + J_\mu^{\omega\dagger} \tag{2.4}$$

The production of the $\omega\pi$ final state is hidden in this current. $J_\mu^{I\dagger}$ describes two pion ρ resonances that the pions pass through, while $J_\mu^{\omega\dagger}$ deals with $\rho\omega\pi$ coupling and ω production. $J_\mu^{I\dagger}$ originates from the chiral Lagrangian assuming massless quarks, thus splitting up the left and right handed components of quark spins. The anomalous part, $J_\mu^{\omega\dagger}$ is not invariant under isospin [16]. Isospin is a quantum number that comes from assuming the u and d quarks are essentially interchangeable due to their reasonably identical mass. u and d can therefore be considered upper and lower states of a spinor in isotopic spin space [17]. It is analogous to conservation of angular momentum.

For the pions, π^+, π^0, π^-, and π^-, having the 4-momenta, p_1, p_2, p_3, and p_4, respectively, and the incoming W 4-momentum, $q = p_1 + p_2 + p_3 + p_4$, the currents can be written [16]

$$J_\mu^I = \frac{2\sqrt{3}}{f_\pi^2} \cos\theta_c B_\rho(q^2) \sum_{k=2}^{4} R_k A_{\mu\nu}^k B_\rho(s_k)(p_k - p_1)^\nu, \tag{2.5}$$

where

$$R_k = 1 \; for \; k = 3,4, R_k = -2 \; for \; k = 2 \tag{2.6}$$

$$A^k_{\mu\nu} = g_{\mu\nu} - \sum_{l=2, l\neq k}^{4} \frac{(q-2p_l)_\mu (q-p_l)_\nu}{(q-p_l)^2} \tag{2.7}$$

$$s_k = (p_k + p_1)^2 \tag{2.8}$$

$$B_i(Q^2) = \frac{m_i^2}{m_i^2 - Q^2 - i m_i \Gamma_i(Q^2)} \tag{2.9}$$

and

$$\begin{aligned} J^\omega_\mu \;=\; & \left(\frac{1}{m_\rho^2 - i m_\rho \Gamma_\rho} \widehat{BW}_\rho(q^2, m_\rho, \Gamma_\rho) + \sigma \frac{1}{m_{\rho'}^2 - i m_{\rho'} \Gamma_{\rho'}} \widehat{BW}_{\rho'}(q^2, m_{\rho'}, \Gamma_{\rho'}) \right) \\ & \cos\theta_c G_{\omega 3\pi} g_{\rho\omega\pi} F_\rho \left\{ \frac{1}{m_\omega^2 - i m_\omega \Gamma_\omega} \widehat{BW}_\omega((q-p_4)^2, m_\omega, \Gamma_\omega) \right. \\ & [p_{3\mu}((1-p_4) \cdot p_2 p_1 \cdot p_4 - (q-p_4) \cdot p_1 p_2 \cdot p_4) \\ & + p_{2\mu}((q-p_4) \cdot p_1 p_3 \cdot p_4 - (q-p_4) \cdot p_3 p_1 \cdot p_4) \\ & p_1 \mu((q-p_4) \cdot p_3 p_2 \cdot p_4 - (q-p_4) \cdot p_2 p_3 \cdot p_4)] \\ & \left. + (3 \leftrightarrow 4) \right\}, \end{aligned} \tag{2.10}$$

where

$$\widehat{BW}_i(Q^2, M_i, \Gamma_i) = \left(\frac{m_i^2 - i m_i \Gamma_i}{m_i^2 - Q^2 - i m_i \Gamma_i} \right). \tag{2.11}$$

For equation 2.5, f_π is the form factor for the pions describing their decay [3]. $cos\theta_c$ is the Cabbibo or mixing angle which removes strangeness from the decay. B_ρ is the Breit-Wigner enhancement of the ρ, both before and after the pions are produced. A Breit-Wigner resonance curve describes the shape of the amplitude of a decay as a function of width (lifetime), Γ_i, mass, m_i, and total energy, Q, of the state. $A^k_{\mu\nu}$ projects out the transverse components of the pion pairs.

Several factors in equation 2.10 can be combined to form F_2 a hadronic form factor in the $V-A$ matrix element,

$$F_2 = F_\rho g_{\rho\omega\pi}\left(\frac{1}{m_\rho^2 - im_\rho\Gamma_\rho} + \sigma\frac{1}{m'_\rho - q^2 - im_{\rho'}\Gamma_{\rho'}}\right). \qquad (2.12)$$

This is the only non-zero form factor of the hadronic matrix element for 1^- (J^P) states like the $\omega\pi$ system [18]. The full hadronic matrix element will be discussed in more detail in section 2.4. This form factor includes the ρ form factor, F_ρ, the coupling for a $\rho\omega\pi$ system, $g_{\rho\omega\pi}$, and the relative rates of ρ/ρ' coupling with the resultant $\omega\pi$ system, σ. $G_{\omega 3\pi}$ is the coupling constant for the ω going to 3π. The ω Breit-Wigner dictates the level of ω production across the allowed energy range. The rest of the anomalous current cycles over the allowed 4-vector combinations making up the $\omega\pi$ system.

Once the matrix element is squared to produce a probability for the $\pi^-\pi^+\pi^-\pi^0$ decay, the chiral and anomalous parts of the hadronic current will interfere, i.e. the final level ρ's and ρ''s in the chiral current will intermingle with the ω in the anomalous current during production. This interference is not significant across most of the energy range, since the width of the ω, and therefore its Breit-Wigner, is very narrow.

2.2 G-parity

Charge alone is not enough to create selection rules for charged particles, since the charge conjugation operator, C, can only have eigenvalues for neutral systems such as the π^0. For strong interactions of charged systems, the combination of an isospin rotation, R, with charge conjugation is called G-parity[19].

$$G = CR = Cexp(i\pi I_2) \tag{2.13}$$

G acting on a wave function rotates the system by 180^o about the y-axis in isospin space, followed by a sign change of the charge. Charged states may be eigenfunctions, but only if they interact strongly, thus maintaining isospin invariance.

To obtain the isospin eigenvalues, an isospin state $\chi(I, I_3 = 0)$ is defined which is analogous to the angular momentum wavefunction $Y_l^{m=0}(\theta\phi)$.

$$\chi(I,0) \xrightarrow{R} (-1)^I \chi(I,0) \tag{2.14}$$

In a nucleon-antinucleon state for which G-parity was first explored [20] the C-operator gives a factor $(-1)^{l+S}$ with S as total spin and l orbital angular momentum. Therefore the effect of G on this neutral ($I_3 = 0$) system is

$$G \mid \psi\rangle = (-1)^{l+S+I} \mid \psi\rangle \tag{2.15}$$

This can be extended to all strong interactions, since they are invariant under isospin rotations. The G-operator becomes a pion counter. R flips isospin, converting π^+ to π^- ($u\bar{d} \to \bar{u}d$). Charge conjugation flips the sign back.

$$\begin{aligned} G \mid \pi^+\rangle &= \pm \mid \pi^+\rangle \\ G \mid \pi^-\rangle &= \pm \mid \pi^-\rangle \\ G \mid \pi^0\rangle &= \pm \mid \pi^0\rangle \end{aligned} \tag{2.16}$$

This alone does not constrain the sign of the parity. Consider the neutral pion. It is an eigenstate of C with an eigenvalue of $^+1$, since it decays into an even number of photons, $\pi^0 \to \gamma\gamma$. With the isospin rotation eigenvalue being $(-1)^I = -1$, this means that the π^0 has an overall negative G-parity.

$$G \mid \pi^0\rangle = - \mid \pi^0\rangle \tag{2.17}$$

G-parity for the charged pions is ambiguous, since charged particles do not have eigenvalues for the C-operator. To maintain the symmetry of the isospin triplet, all pions have the same G-parity.

Isospin is additive, C is multiplicative, and therefore, so is G. The simple way of finding out a system's G-parity is to count the pions, since

$$G \mid \psi(n\pi)\rangle = (-1)^n \mid \psi(n\pi)\rangle. \tag{2.18}$$

2.3 Vital Statistics of the ω

The $\tau \rightarrow \omega\pi\nu_\tau$ decay was first observed at DORIS-II in 1987. According to the standard model, the production of the $\omega\pi$ is embedded in the vector current in equation 2.4.

The omega is a neutrally charged strongly interacting vector meson, or more esoterically, $I^G(J^{PC}) = 0^-(1^{--})$. J^P is a description of spin and symmetry in space. The ω, therefore, has a total angular momentum of 1 and odd parity. Odd parity means that the ω wave function flips sign when a spatial inversion is performed. It has no charge because of its orientation in isotopic spin space $(I = 0)$. Since $I_3 = 0$, the ω can be made up of $u\bar{u}$, $d\bar{d}$, or $s\bar{s}$ [1]. Charge parity, C, for the ω is odd since it can decay into $\pi^0\gamma$. G-parity is negative for the ω [21].

There are other important features of the ω that are worth mentioning at this point [14]. Its narrowness of width, $\Gamma = 8.43 \pm 0.10$ MeV, proves useful in distinguishing this resonance from the rest of the $3\pi\pi^0$ sample. The width is restricted due to G-parity conservation, forcing the omega to decay into $\pi^+\pi^-\pi^0$ most of the time. This decay has a low rate due to limited phase space. The mass of the ω is measured to be $M_\omega = 781.94 \pm 0.12$ MeV.

2.3.1 Quark Mixing and Composition

Since the ω is a member of the vector meson nonet with hypercharge $Y = 0$ (strangeness, S, and baryon number, B, both 0), it is a product of octet and singlet mixing [11]. The $I = S = 0$ octet, ϕ_8, and singlet, ϕ_0, wavefunctions are defined by

$$\begin{aligned} \phi_8 &= (d\bar{d} + u\bar{u} - 2s\bar{s})/\sqrt{6} \\ \phi_0 &= (d\bar{d} + u\bar{u} + s\bar{s})/\sqrt{3}. \end{aligned} \tag{2.19}$$

The ω and ϕ are linear combinations of the above wavefunctions.

$$\begin{aligned} \phi &= \frac{1}{\sqrt{3}}(\phi_0 - \sqrt{2}\phi_8) \\ \omega &= \frac{1}{\sqrt{3}}(\phi_8 + \sqrt{2}\phi_0) \end{aligned} \tag{2.20}$$

The square root corresponds to the theoretically derived value of the large mixing angle $\theta \approx 35^o$ [22], thus allowing the ω to be free of strange quarks. Combining the above equations, the ϕ and ω wavefunctions are achieved.

$$\begin{aligned} \phi &= s\bar{s} \\ \omega &= (u\bar{u} + d\bar{d})/\sqrt{2} \end{aligned} \tag{2.21}$$

Experiment is in close agreement with this "ideal" or "magic" case. The masses of the ω and ρ are close, while the ϕ has a much higher mass due to its strangeness. This also provides some understanding about the ω decay modes, since the ω decay products are composed of u and d quarks exclusively.

2.3.2 ω Decay Products

Table 2.1 shows the omega decay channels [14]. Note that there is a significant proportion of two pion decays quoted. While this decay would be G-parity violating if it were to occur through a strong interaction, the decay is electromagnetic. The electromagnetic nature of this decay is shown by its much smaller branching fraction due to a weaker coupling constant α as opposed to α_s. The electroweak current does not obey G-parity, and therefore, there is no surprise that this decay is present. It is not the ω itself that might show G-parity violation. The ω becomes interesting if it is the product of a G-parity violating decay, $\tau \rightarrow b_1\nu_\tau$.

Decay Product	*Fraction* (Γ_i/Γ)
$\pi^+\pi^-\pi^0$	$88.8 \pm 0.7\%$
$\pi^0\gamma$	$8.5 \pm 0.5\%$
$\pi^+\pi^-$	$2.21 \pm 0.3\%$

Table 2.1: Principle ω decay products. The other decay channels are statistically insignificant and principally electromagnetic.

2.4 Second Class Currents

In 1936, Gamow and Teller realised that there is a way of integrating strong interactions into the scalar weak lagrangian that is invariant under proper Lorentz and parity transformations [23]. The information available in the interactions is the spin and momenta of the particle states in the system. There are three linearly independent vectors and three linearly independent axial-vectors that can be used to construct the matrix element, M_α, a linear combination of vector and axial components.

$$M_\alpha = V_\alpha + A_\alpha \tag{2.22}$$

The matrix element is expressed in terms of initial and final state wave functions, ψ_i and their 4-momenta. The 4-fermion Lorentz invariants take the form [24]

$$\begin{array}{rc} \psi_f\psi_i & scalar \\ \psi_f\gamma_\alpha\psi_i & vector \\ \psi_f\sigma_{\alpha\beta}\psi_i & tensor \\ \psi_f\gamma_5\gamma_\alpha\psi_i & axial-vector \\ \psi_f\gamma_5\psi_i & pseudoscalar \end{array} \tag{2.23}$$

Each term comes into the matrix element with a different strength defined by its form factor f_i or g_i

$$\begin{aligned} V_\alpha &= \psi_f(f_1\gamma_\alpha + f_2\sigma_{\alpha\beta}q^\beta + f_3q_\alpha)\psi_i \\ A_\alpha &= \psi_f(g_1\gamma_\alpha + g_2\sigma_{\alpha\beta}q^\beta + g_3q_\alpha)\gamma_5\psi_i \end{aligned} \tag{2.24}$$

The form factors are functions of q^2 where $q = p_i - p_f$, the overall 4-momentum. At low q^2, these form factors can be considered constant.

While the form factors are restricted by the conserved vector current hypothesis, invariance under time reversal, and the Goldberger-Treiman relation [23], the important limitations here are those imposed by G-parity conservation. All terms in equation 2.24 behave similarly in isospin space. They have negative parity. The G-parity of each term therefore relies on its charge parity. The scalar, pseudoscalar, and axial terms possess even C-parity, while the vector and tensor factors are odd under C transformation. This is due to antisymmetrising the matrix element with

Heisenberg's "rule" for proper symmetrisation [25]. This leads to the f_1, f_2, and g_2 terms having even G-parity and terms f_3, g_1, and g_3 having odd G-parity. $f_3 q_\alpha$, the effective scalar, is forbidden by the conserved vector current (CVC) hypothesis [23].

A way to explore the behaviour of the different terms is to turn off the strong force for a moment and pretend that only the weak force is contributing to the matrix elements, $(M_V)_0$ and $(M_A)_0$ [23]. This is equivalent to calculating the currents for free quarks.

$$
\begin{aligned}
(M_V)_0 &= \psi_f \gamma_\alpha \psi_i \\
(M_A)_0 &= \psi_f \gamma_\alpha \gamma_5 \psi_i
\end{aligned}
\tag{2.25}
$$

$(M_V)_0$ is even while $(M_A)_0$ is odd under G transformation. Strong interactions are G-invariant, but when combined with the weak interaction two terms, f_3 and g_2, transform with opposite G-parity than expected. These elements of the current have been dubbed "second class" [20]. f_3 is already ruled out due to the conserved vector current hypothesis, but g_2, "weak electrism", is only forbidden due to G-parity conservation.

While G-parity forbids the g_2 current, G-parity may not be universally conserved in semileptonic decays [26]. G-parity relies on isospin invariance and the assumption that the u and d quarks are different states of the same particle. If there is a difference between the u and d masses, there is a possibility of violating G-parity, and creating a particle with even G-parity from the axial part of the matrix element.

While second-class currents were originally searched for in β-decaying nuclei [20], it was suggested in 1978 that an experiment looking at τ decays could produce a more accurate measure of G-parity violation [27]. The dominant b_1 decay is $b_1^- \to \omega\pi^-$ [14]. This system eventually decays into 4 pions, and therefore the b_1 has even G-parity. It is an axial-vector, ($J^P = 1^+$), and therefore, is produced by the axial part of the matrix element. With its even G-parity, this should be forbidden, while there is no problem with the standard 1^- $\omega\pi$ system that is produced by the vector component of the matrix element.

The helicity of the $\omega\pi$ system can be identified by studying the relative orientation, χ, of the normal to the ω decay plane and the bachelor π direction [29][32]. The overall spin and parity of the system, J^P, has a characteristic distribution in the variable $\cos\chi$. The second-class axial-vector current, potentially producing a b_1, has $J^P = 1^+$, should have a flat $\cos\chi$ distribution. The standard $\tau \to \omega\pi\nu_\tau$ decay travelling through the vector current as a first-class decay has $J^P = 1^-$ and a $\cos\chi$ distribution proportional to $\sin^2\chi$. A fit to the resultant χ measurements of a pure $\omega\pi$ sample determines the overall spin and parity of the system. Contamination from other $\tau \to 3\pi\pi^0$ decays would swamp χ, since there is no inherent helicity available from the plane made by any 3 pions.

2.4.1 Impact on the Standard Model

The existence of second-class currents such as the semileptonic decay $\tau \to b_1\nu_\tau$ will not shatter the Standard model. The search for these forbidden decays will provide a better understanding of assumptions made about isospace. While the success of the theory is compelling, there is still no hard evidence that the u and d quarks are different isospin projections of the same particle. Their masses may not be exactly equal. Their structure may not be infinitely small. Evidence of second-class currents would require physics to slightly revise its view of strong interactions.

In order to look for $\tau \to b_1\nu_\tau$, $\tau \to \omega\pi\nu_\tau$ must be studied. The first step is to determine a branching ratio to see how many $\omega\pi$ decays are expected from all τ's. Once $Br(\tau \to \omega\pi\nu_\tau)$ is measured, it is necessary to isolate a reasonably pure sample of $\tau \to \omega\pi\nu_\tau$ to provide a competitive measure of $\mathrm{Br}(\tau \to b_1\nu_\tau)$.

Chapter 3

Previous ω Studies

Experiment	*Branching Ratio (%)*
ARGUS (1986)	$1.5 \pm 0.3 \pm 0.3$
CLEO (1987)	$1.60 \pm 0.27 \pm 0.41$
CLEO (1995)	$1.95 \pm 0.07 \pm 0.11$
ALEPH (1996)	$1.91 \pm 0.07 \pm 0.06$
World Average [14]	1.91 ± 0.09
CVC Prediction	1.79 ± 0.14

Table 3.1: Previous measurements of the $\tau \to \omega\pi\nu$ decay. The World Average is a fit to the two CLEO measurements only. The measurements are all within error of prediction, although the most recent measurements are systematically higher than the CVC calculation.

Table 3.1 contains the previous measurements of $Br(\tau^- \to \omega\pi^-\nu_\tau)$ to date along with the most recent experimentally independent conserved vector current prediction of the branching ratio [28]. The measurements are few and far between. This is due to the complexity in identifying, with any kind of accuracy, the π^0 resultant in an ω decay. Each experiment attempting the measurement must have the ability to resolve photons with a great deal of precision. The charged particle tracking system and the electromagnetic calorimeter must have high resolving power.

3.1 ARGUS

The exclusive decay channel $\tau^- \to \omega\pi^-\nu_\tau$ was first observed by the ARGUS collaboration at DORIS-II [29]. Through measurements of $e^+e^- \to \tau^+\tau^-$ events with a 3-1 topology, a reasonably clean sample of 556 $\tau^- \to \pi^-\pi^-\pi^+\pi^0\nu_\tau$ decays was collected. The ultimate concern was proper kinematic fitting of the photon candidate(s) to form one or more π^0. The branching ratio of the ω decay channel was determined by fitting a Gaussian with a simple Breit-Wigner and a smooth background. A rough check on the $\pi^+\pi^-\pi^0$ Dalitz plot confirmed the ballpark measurement found in the table above.

Due to limited statistics, the binning for the second class current search was huge. No evidence of these events were found, and only a 90% confidence level for second class contribution of less than 50% was placed on their observation.

3.2 CLEO

With only a minimal search for G-parity violating events already performed and a branching ratio that could be improved upon for $\tau \to \omega\pi\nu_\tau$, a group at the CLEO detector at the Cornell Electron Storage Ring made a stab at measuring the branching ratio [30].

A 3-1 event topology sample with 3 times the statistics of ARGUS was analysed. Photons were selected as possible π^0 candidates and kinematically fit to the π^0 mass. A Dalitz plot was used to make a selection of events, and went so far as to pick the best $\pi^+\pi^-\pi^0$ combination by cutting on the interior half of the 3π mass plot. This best combination sample was then fitted to an ω peak using a Gaussian signal function with a polynomial background.

After the installation of a new detector (CLEO II), a similar measurement was once again performed [31]. These two measurements (table 3.1) are combined to form the present global average. No significant contribution from second class currents was found.

3.3 ALEPH

ALEPH [32] [33] was no longer restricted to a 3-1 topology of tau decays, since 3-3 τ pairs at the higher *cms* are unambiguous with respect to hadronic decays. They abandoned fitting methods implementing either Gaussian or Breit-Wigner functions, since the ω resonance is neither of these curves. Instead they relied on Monte Carlo simulations that had progressed significantly providing reasonable modeling of the data. Any discrepancy would manifest itself as a difference in branching ratio which would suggest non-Standard Model physics (or a bug in the Monte Carlo).

For the π^0 fit, a cut is made on the invariant mass of the two (or one) photon system assigned to the π^0 candidate, which is then constrained to M_{π^0}. Once the π^0 candidate is determined, a fit to the Monte Carlo expectation is made to determine the proportion of ω's in the data sample. A knowledge of the width of the ω resonance and $Br(\tau \to 3\pi\pi^0)$ allows for a measurement of the $\tau \to \omega\pi\nu$ branching ratio.

The contribution from second-class currents to the quoted ALEPH branching ratio has been measured to be less than 9% with a 95% confidence level.

3.4 Justification for Another Measurement

While the measurement of $Br(\tau^- \to \omega\pi^-\nu_\tau)$ has been gaining precision over the last decade, there is still room for improvement. The earlier measurements were plagued with limited statistics and inadequate fitting methods due to their reliance on standard Gaussian and Breit-Wigner functions to model the ω resonance. Monte

Carlo modeling has evolved enough in order to search for tiny discrepancies with the generally accepted standard model.

There is potential, with the statistics and improved photon resolving techniques, to produce a competitive measurement at OPAL. In contrast to the brute force of ALEPH's MC fitting technique, more sophisticated statistical methods have been implemented.

Chapter 4

The Experiment

The following analysis was performed using OPAL, a detector at CERN, the European Centre for Particle Physics. In both human resources and land mass, CERN is one of the world's largest scientific facilities.

4.1 LEP

A reason for its physical size is LEP [34], the Large Electron-Positron storage ring. The ring houses bending and focusing magnets, an RF acceleration system, and experimental detector sites at the four collision points. A detailed description of LEP's design can be found in [35]. LEP was piggybacked onto an already existing system of synchrotrons in order to parasitically use it as an injector (figure 4.1). Until the end of the 1995 run, LEP was running at a 91.2 GeV centre of mass energy , the Z^0 production resonance, for e^+e^- collisions. In the past two years, the energy has been systematically increased to study physics at and around the W pair production threshold of 161 GeV centre of mass energy. The work presented here deals with the data collected in 1994 on the Z^0 peak.

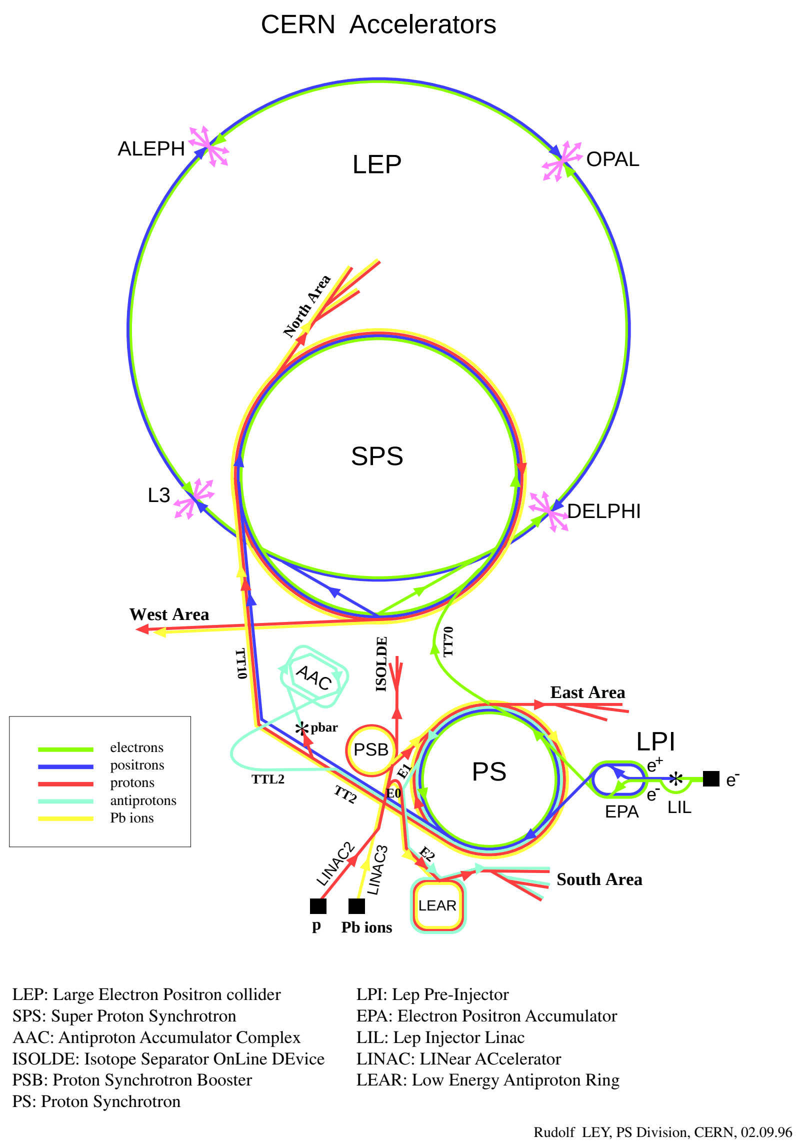

Figure 4.1: CERN's accelerator array.

4.2 OPAL

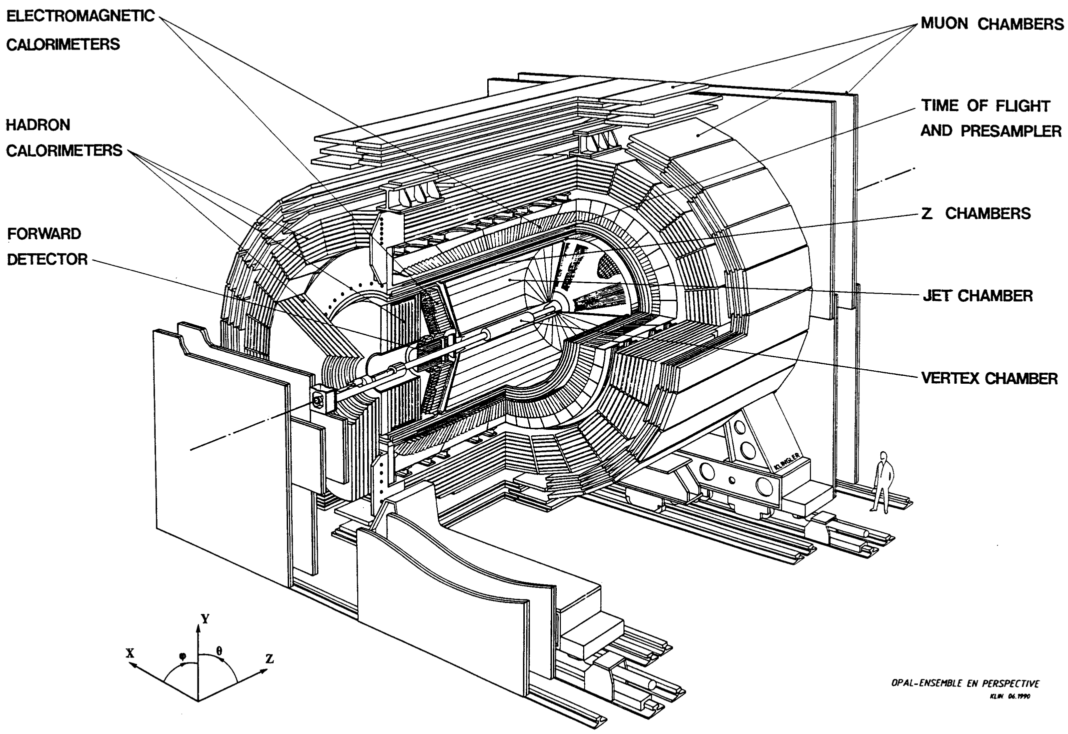

Figure 4.2: A breakdown of OPAL's subdetectors.

OPAL (figure 4.2) is considered the *safest* of the detector facilities spaced evenly around LEP. This simply means that OPAL was designed with established detector technology of the time [36]. The analysis discussed in this paper utilises data collected in 1994, after the 1991 upgrade and before the 1996 increase in energy. There have consistently been 8 electron bunches colliding with 8 positron bunches in the ring at one time, 16 subdetectors, and a beryllium beam pipe [37]. The centre of mass energy was consistently within ± 3.3 MeV of the Z^0 mass [38].

Several of the subdetectors have been used for reliable tau pair event identification, as well as identification of τ decay products. Only a cursory mention of the subdetectors will be found here. The endcap subdetectors will only be mentioned in passing, because they have not been used. A more detailed description of the detector and its finer points can be found in [39], [40], and [41].

4.3 OPAL Coordinate System

In the case of OPAL, position is defined by the OPAL Master Reference System, MRS, with the origin at the anticipated interaction point. The z axis lies along the beam pipe with the incident e^- direction as positive. The x axis points toward the centre of the LEP ring and the y axis completes the right-handed coordinate system by pointing upwards. θ is the polar angle measured from z and ϕ is the azimuthal angle from x. The MRS has already compensated for the 1.39% tilt to the horizontal on which LEP has been built due to the area's geography.

4.4 Central Tracking System

The central detector is housed in a pressure vessel under a uniform axial magnetic field of 0.435T and a pressure of 4 bar. It consists of a silicon microvertex detector, a vertex detector, a pictorial drift chamber, and z-chambers.

4.4.1 Silicon Microvertex Detector (SI)

The silicon microvertex detector is composed of two barrel shaped, single sided silicon microstrip detectors. The microvertex detector was installed in 1991 with r-ϕ readout only. In 1993, this was upgraded by gluing r-z wafers to the backs of the existing r-ϕ plates.

4.4.2 Vertex Detector (CV)

The vertex detector is a two layer cylindrical drift chamber. The inner layer's sense wires run parallel to the beam pipe, while the outer layer sense wires are skewed by an $\sim 4^0$ stereo angle. From the axial layer, a high precision measurement of the drift time enables an accurate measurement of the r-ϕ coordinate. A coarse z measurement is determined from the time difference between signals at the wires' ends. More precision is achieved offline for the z measurement by utilising the information from both the axial and stereo wire outputs.

4.4.3 Jet Chamber (CJ)

The jet chamber is another drift chamber split into 24 ϕ sectors of 159 axially positioned sense wires each. At least 8 points are obtained for a track over a solid angle of 90% of 4π. A drift time measurement, charge division, and the position of the wire enables the calculation of the track position in three dimensions (r, θ, ϕ). An average resolution in r-ϕ of 135 μm at the mean drift distance of 7 cm and an average resolution in z of 6 cm is achieved. This information can also be useful in calculating dE/dx, the energy loss per unit length.

4.4.4 Z-Chambers (CZ)

The z-chambers provide a precise z measurement that the jet chamber is unable to supply. These drift chambers allow for a z coordinate resolution of 100-350 μm depending on the position of the track from the wire. The z-chambers cover 94% of the azimuthal angle, ϕ, and accept tracks within $|\cos\theta| < 0.72$.

4.5 Time-of-Flight Counters (TOF)

Positioned just outside the magnetic coil, the time-of-flight counter is a system of scintillation counters. Charged particles can be identified within the energy range of 0.6 to 2.5 GeV and a polar angle range of $|\cos\theta| < 0.82$. The information gathered from the TOF is the main cosmic ray veto.

4.6 Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter is the detector closest to the interaction point that extends into the endcap. The combination of the barrel and endcap regions of the EM calorimeter covers almost 99% of the solid angle and is used to identify electrons and photons. The ECAL provides the information necessary to resolve closely positioned photon pairs produced in the $\tau \rightarrow \omega\pi\nu$ decay. The decay products must pass through approximately 2 radiation lengths, $\sim 2X_0$, of material before reaching this detector array. Therefore, most of the electromagnetic showers will begin before the lead glass calorimeters of the ECAL. In order to minimise the loss of information, presampling devices have been installed to measure the energy and position of showers. This improves the ability to discriminate between γ and π^0.

4.6.1 Electromagnetic Presampler (PB and PE)

The barrel electromagnetic presampler, PB, consists of an array of 16 axially positioned chambers containing two layers of limited streamer tubes. While the anode wires run along the z-direction, the top and bottom cathode strips are offset by $\pm 45^o$. A z component of the track's position is determined by charge division from the ends of the wire. The r-ϕ components are calculated by comparing the readouts of the cathode strips to the anode wires. The fact that the energy deposited in the streamer tubes should be approximately proportional to the hit multiplicity allows for a correction to the measured energy deposited in the EM calorimeter and therefore improves energy resolution.

Instead of a limited streamer tube design like the PB, the endcap electromagnetic presampler, PE, is an array of multiwire proportional counters.

4.6.2 Lead Glass Calorimeter (EB and EE)

In the barrel, 9440 lead glass blocks, each angled toward the interaction region, define the barrel lead glass calorimeter. The projective geometry reduces the number of neutral particles escaping through gaps between crystals. These blocks have an acceptance of $|\cos\theta| < 0.81$ (identical to PB). To maximise the containment of an electromagnetic shower in a limited space, the blocks are made of a specially engineered heavy glass and are only 24.6 radiation lengths long. Each has a cross section of $\sim 10\times \sim 10$ cm^2. The Čerenkov light produced in the lead glass is collected by magnetic field tolerant phototubes at the outside of each crystal.

The energy resolution in the lead glass blocks, without any material in between them and the incident beam is $\sigma_E/E =\sim 5\%/\sqrt{E}$ at low energies [39]. Degradation of the resolution occurs with the introduction of the 2 X_o of detector before it, decreasing the resolution by a factor of 2. ~50% of this loss in precision can be recovered, though by using information from the PB.

A quick, back-of-the-envelope calculation shows that the lateral separation, D, between two photons from a π^0 (mass M_{π^0} and energy E_{π^0}) over a distance L is

$$D = \frac{2LM_{\pi^0}}{\sqrt{E_{\pi^0}^2 - M_{\pi^0}^2}} \tag{4.1}$$

Therefore, over the 2.5 metres of closest approach between the interaction point and the EB, a neutral pion of average energy, around 10 GeV, would have 7 cm photon separation. With the ECAL block faces being significantly larger, most of the photon showers from π^0 decays will overlap in a block, only making one neutral ECAL cluster identifiable.

The endcap electromagnetic calorimeter, EE, provides only 22 radiation lengths on average and the lead glass blocks are positioned axially. Because the EE is in the full OPAL magnetic field, special Vacuum Photo Triodes (VPTs) are used to read out the light from the EM showers in the blocks.

4.7 Hadron Calorimeter (HB, HE, and HP)

The magnetic return yoke is exploited for its 4 interaction lengths of iron by placing active detectors between the yoke layers to create the hadron calorimeter. The three sections of the hadron calorimeter cover 97% of the solid angle and absorb virtually all the hadronic activity, leaving only the muons to pass through to the muon detectors. Because of the probability that a hadronic shower will begin in the ECAL, the information obtained from the hadron calorimeter, HCAL, must be used in conjunction with that from the ECAL.

Both the hadron barrel, HB, and endcap, HE, are of limited streamer tube design with the tubes sandwiched in between leaves of iron. The signals off the streamer tube wires are only used for monitoring detector performance. The track signals are read off the cathode pads and strips on the inner and outer surfaces of the chambers respectively. The energy of a hadronic shower is estimated from the sum of the analogue signals off the pads. A rough measurement of a particle's track can be taken from the strips with a resolution of the wire pitch.

The hadron pole-tip calorimeter, HP, increases the angular acceptance of hadron calorimeter from $|\cos\theta| = 0.91$ to $|\cos\theta| = 0.99$. The HP contains multiwire proportional counters.

4.8 Muon Detector (MB and ME)

The barrel muon detector, MB, [36] fits around the HCAL and its supports. ϕ and coarse z coordinates are determined from drift time on the anode wires that run parallel to the beam axis. Finer z measurements are made by reading the induced signals off two different sized diamond shaped cathode pads on the top and bottom of the chamber. The polar angular acceptance range is $|\cos\theta| < 0.68$ for four layers of drift chamber and $|\cos\theta| < 0.72$ for one or more layers.

The endcap muon detector consists of four layers of streamer tube detectors, two layers of quadrant chambers and two layers of two patch chambers.

4.9 Forward Detector (FD)

The forward detector has been used to determine the LEP luminosity by detecting low angle Bhabha scattering events. This measurement is necessary when normalising the reaction rates from Z^0 decays. Only the beam pipe and a small section of the aluminium pressure vessel impairs its acceptance. The forward detector is composed of a calorimeter (including a presampler and layered proportional tube chambers), a gamma catcher to plug up the hole between EE and the forward calorimeter, and a far forward monitor to measure the number of electrons deflected outwards by the LEP quadrupole magnets. Each section is made of lead-scintillator.

4.10 Silicon Tungsten Detector (SW)

The silicon tungsten detector [41], SW, is also designed to measure luminosity through the detection of Bhabha scattering events. Alternate layers of silicon detectors and tungsten make up the calorimeters at ± 238.94 cm from the interaction point. An extra layer of silicon detects preshowering. After the first several weeks of the 1993 data taking period, the SW replaced the FD for luminosity measurements. Both detectors are functional, but general data analysis preferentially considers the SW measurements.

Chapter 5

Event Simulation

The data cannot stand alone and provide adequate information in an analysis of this kind (unless a multidimensional integration is performed to determine acceptance). Knowledge of how particle interactions behave in the environment of the detector has been collected over the years from both theorists and experimentalists. This information is used to produce a reasonably accurate model of random particle interactions...a Monte Carlo, MC. Random event generators are used to test existent theories as well as determine the limitations of a detector through systematic studies.

5.1 KORALZ

The KORALZ Monte Carlo program [42] is primarily designed to generate τ pairs and their decay products for e^+e^- collision experiments with a centre of mass energy around the Z^0 resonance. In addition to τ pairs, it is capable of producing other fermion-antifermion final states (under some restrictions). KORALZ is used across LEP to produce $e^+e^- \rightarrow \tau^+\tau^-$ and $e^-e^+ \rightarrow \mu^+\mu^-$ MC samples.

By hardwiring in educated guesses and experimentally measured values for some parameters such as the masses of the top quark, Higgs boson, and Z^0, and the vector and axial coupling strengths, KORALZ is capable of generating the initial and final 4-vectors of most $e^+e^- \rightarrow \tau^+\tau^-$ interactions. The generator takes into account physical processes that affect the production of the tau pair and their decays. Among these are multiple initial and final state hard radiation, $O(\alpha)$ radiative corrections and bremsstrahlung, and some spin polarisation effects.

5.2 TAUOLA

TAUOLA [15][16] is the τ decay library embedded in KORALZ. This is where particle theory comes into play. More than 20 decay channels are modelled. The matrix element is calculated for the V-A current corresponding to a specific final state X, with X being everything except the neutrino as in $\tau \rightarrow \nu_\tau + X$. Once the matrix element is determined, a decay rate is calculated, and the program is off and running. The latest version of TAUOLA, TAUOLA 2.4, improves on the matrix element calculation and generation algorithm for $\tau^- \rightarrow \nu_\tau 2\pi^-\pi^+\pi^0$. It now features the intermediate-state resonances of the ω and ρ plus an additional current to take into account $\rho\omega\pi$-coupling. The relevant theory and currents have been described in Chapter 2.

Due to the method of calculating $\tau^- \rightarrow \omega\pi^-\nu_\tau$, TAUOLA does not hardwire an estimate of $Br(\tau \rightarrow \omega\pi\nu_\tau)$. Instead the program generates decays using $Br(\tau \rightarrow 3\pi\pi^0\nu_\tau) = 0.0598$ and $\Gamma_\omega = 0.0085$. $\sim 1.43\%$ of the initial tau pair Monte Carlo low multiplicity sample is identified as a $3\pi\pi^0$ decay lying within the ω mass window. This is an estimate of the branching ratio which underestimates the proportion of $\omega\pi$ decays (see figure 5.1).

5.3 GOPAL

An event generator simply produces the initial and final 4-vectors for interactions. It disregards the material through which the particles travel. The inner workings of OPAL are simulated using the GOPAL program [43] which utilises GEANT [44], a detector modelling package. The substance and geometry is specified, and then the particles are sent through a virtual detector. Scattering and absorption are accounted for to produce as precise a copy of the real experiment as possible. Once GOPAL has processed the KORALZ events, they are digitised, and from here on out, treated like experimental data.

5.4 Simulation Specifics

For the purpose of this analysis, GOPAL 1.33, using GEANT 2.2130, was used to model the the workings of the 1994 OPAL detector. In addition to the standard OPAL low multiplicity 600K $e^+e^- \to \tau^+\tau^-$ MC sample, 200K $3\pi\pi^0$(not ω), and 200K $\omega\pi$ events were generated.

5.5 Where ARE the ω's?

The generation does not automatically separate events going through specific resonances such as the ω, since interference terms must be taken into account. This means that the Monte Carlo program identifies $\tau^- \to \pi^-\pi^+\pi^-\pi^0$ without being able to specify whether the decay went through a resonance. Therefore, $\tau^- \to \omega\pi^-\nu_\tau \to \pi^-\pi^+\pi^0\pi^-\nu_\tau$ is lumped together with other $\tau^- \to \pi^-\pi^+\pi^0\pi^-\nu_\tau$ decays for the sake of Monte Carlo decay tree identification. The matrix element must be altered to get a pure sample of omegas. This is simplified by the supremely narrow nature of the resonance, so interference can essentially be ignored. Figure 5.1 shows the extent to which the resonance is isolated. For $\sim$ 92% of the ω sample, at least one $\pi^+\pi^-\pi^0$ invariant mass combination lies within 1% of the world average M_ω of 782 MeV, while only $\sim$ 3% of the non-ω sample falls in this mass window. The spread of the ω resonance is due to the shape of the Breit-Wigner in the amplitude.

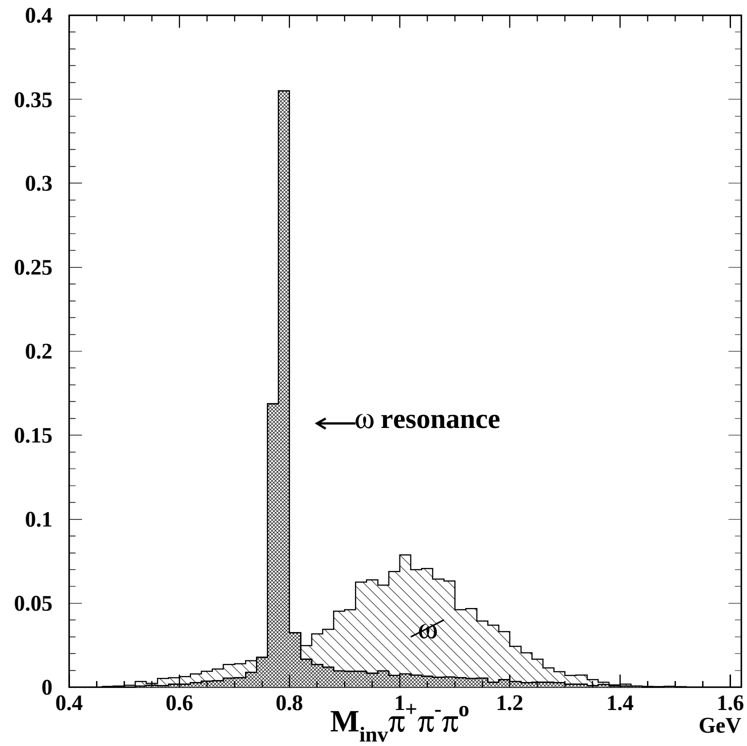

Figure 5.1: The reconstructed Monte Carlo *truth* invariant masses of the $\pi^+\pi^-\pi^0$ combination in closest agreement with the ω mass is shown in the same plot for $3\pi\pi^0(\omega)$, shaded, and $3\pi\pi^0(not\ \omega)$, hatched, events. Most of the ω resonance is sharply defined. The asymmetric tail is due to an increase in phase space at higher energies.

The selection of an ω sample is performed by turning on and off the chiral and anomalous parts of the $3\pi\pi^0$ current. These two independent currents govern non-ω and ω production respectively in the matrix element for $\tau^- \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$ decays discussed in Chapter 2. The two samples obtained, 200K ω and 200K non-ω $\tau \rightarrow$

$3\pi\pi^0\nu_\tau$ decays, are used for $3\pi\pi^0$ separation in lieu of any positive ID from the total MC. Figure 5.2 shows the two possible combinations for ω candidates, highlighting the separation between the non-resonant $\tau \to 3\pi\pi^0\nu$ decays and those passing through an ω resonance.

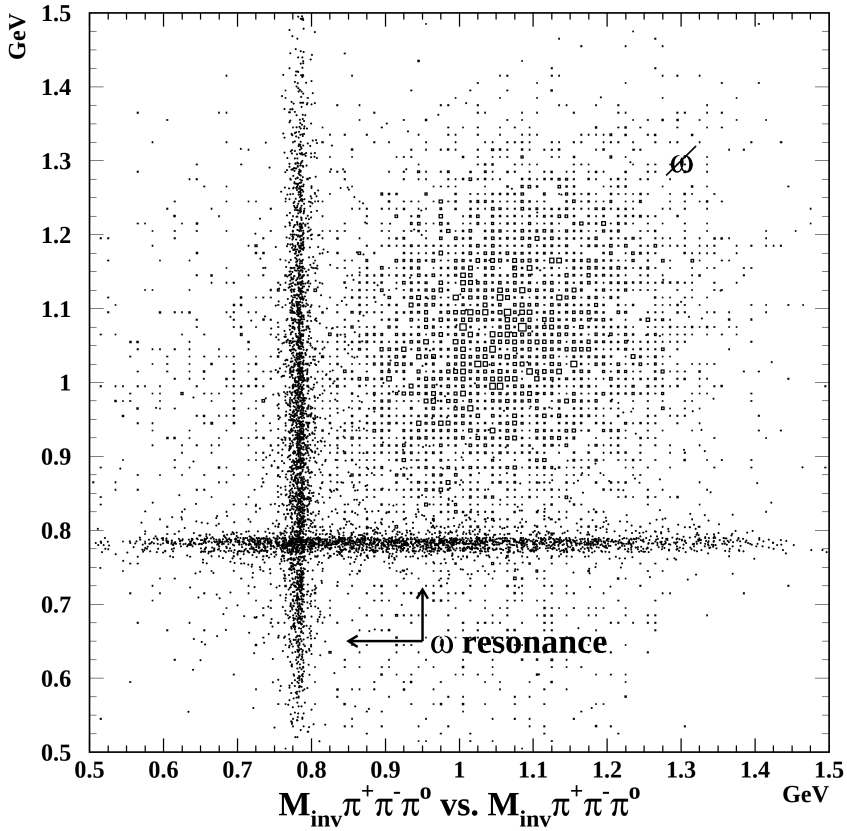

Figure 5.2: The two mass combinations for the $3\pi\pi^0$ sample are shown versus each other. The dots represent the ω sample that tightly adheres to the resonance energy. The boxes are the non-ω background.

Chapter 6

Event Selection

The previous chapter discussed only the generation of phantom events and their interaction with a make-believe detector. Nothing constructive has yet been mentioned with reference to the analysis of these or real data. In order to get any useful information from the electronic output from the detector, it is necessary to reduce the size and pack the data samples into convenient, and preferably clean, subsets relevant to this study.

6.1 Online Selection

OPAL is triggered as interesting physics passes through the detector, allowing for storage of vital information from the 14 subdetectors while reducing background contamination. The "central trigger logic" processor [45] is split into two parts. "direct" trigger signals such as total energy deposited and track multiplicity come from single subdetectors with generally high thresholds. "θ-ϕ matrix" triggers provide information about spatial coincidences between individual components throughout the detector. The efficiency for tagging $e^+e^- \rightarrow \tau^+\tau^-$ events with "visible energy" (the sum of energy in the ECAL and central detector) is $99.9 \pm 0.1\%$. With this analysis focusing on events within the barrel region containing 3 pions which deposit significant energy in the detector, the efficiency is effectively 100%.

Events are reconstructed by the ROPE [46], Reconstruction of OPAL Events, program. ROPE is designed to take any analysis level of data and form data structures "containing all our knowledge about reconstructed events" [37]. The basic levels of analysis are *digits* (raw number from the triggers), DSTs (Data Summary Tapes of highly processed and packaged data for specific events), and DDSTs (both *digits* and DSTs). The DDSTs, produced by the OD processor [47] in ROPE provide the easily accessible format required here. From the 52.5 pb^{-1} luminosity of 1994 OPAL data [48] and the τ pair production cross section on the Z peak, $\sigma(Z \rightarrow \tau\tau) = 1.98$ nb [49], almost 104000 τ pair events are available for analysis.

6.1.1 Low Multiplicity

OPAL low multiplicity DDSTs for the 1994 data have been analysed to reduce hadronic background early on without cutting down the τ pair sample. A low multiplicity event satisfies *any* of the following criteria [37]:

- A track with large enough transverse momentum, $p_{xy} > 0.7$ GeV, a small impact parameter, $|d_0| \leq 1$ cm, a small decay length, $|z_0| \leq 50$ cm, and at least 20 hits in the central detector system.
- Two or more electromagnetic clusters with total corrected energy > 6 GeV.
- Two electromagnetic clusters back to back to within 25^0, one having a total corrected energy > 2 GeV.

6.2 Tau Pair Selection

This study relies heavily on statistical analysis, rather than hard cuts on measured parameters. Yet before the mathematical tricks can be put to use, there must be a reasonable sample of τ events.

6.2.1 The Tau Platform

The Tau Platform [50] [51] is a set of OPAL standard subroutines designed to create and manipulate *mini*dsts of tau pair decays, tau platform files. Although the software is capable of identifying specific tau decays, no prior knowledge of the identity of tau decay products is assumed. The statistical analysis is designed to accomplish the necessary separation.

All of the DDST event samples, 1994 data and 600K $e^+e^- \rightarrow \tau^+\tau^-$, 200K $e^+e^- \rightarrow \tau^+\tau^- \rightarrow 3\pi\pi^0\nu_\tau\ 3\pi\pi^0\nu_\tau$ (no ω), and 200K $e^+e^- \rightarrow \tau^+\tau^- \rightarrow 3\pi\pi^0\nu_\tau\ 3\pi\pi^0\nu_\tau$ (ω) Monte Carlo, are dragged through the tau platform in order to loosely select events (Appendix A) to create compact, τ rich files and provide easily accessible information about the events. The tau platform is then used throughout the analysis as a convenient environment in which to work. The background for all τ-pair events selected by the tau platform from the low multiplicity sample of data is $(1.09 \pm 0.29)\%$ [52]. The overall selection efficiency for tau pairs in this sample is estimated to be 57.1 ± 0.2 from Monte Carlo, which corresponds to an efficiency of $\sim 92.0\%$ within the barrel, $|\cos\theta| < 0.7$.

6.2.2 Selection in Analysis

Minimal cuts are imposed on the low multiplicity tau platform files before the data are fit using a fiesta of statistical methods. Events must have:

- 2 and only 2 cones.
- Been positively identified as tau pairs (Appendix A).
- Passed a *tight* selection (also Appendix A).
- Been contained in the barrel, the almost cylindrically symmetric central region of the OPAL. (This is primarily due to the fact that the algorithm used to identify neutral clusters has only been implemented in the barrel region where

the geometry of the ECAL is reasonably uniform. Attempts to extend π^0 identification into the endcaps have been unsuccessful.)

- Occurred while the subdetectors and triggers are in sufficient working order, i.e. have "good" status. (This only works with the data, since the simulation of the detector does not go so far as to simulate fuses being blown, the beam being dumped into OPAL...)
- At least one τ decaying into three charged particles (candidates for the charged pions in the decay $\tau^- \to \pi^- \omega \nu_\tau \to \pi^- \pi^+ \pi^- \pi^0 \nu_\tau$).
- No photon conversions in the 3-prong cone.
- In the 3-prong decay, at least one neutral ECAL cluster that has not been assigned to any of the three charged tracks (Appendix B).
- The first two unassigned neutral clusters (or one, if there is only one) with a significant energy, $E_{unass}(\gamma_1, \gamma_2) > 1.0$ GeV.

The effect of these cuts on the MC sample can be shown in table 6.1. Assuming $Br(\tau \to 3\pi\pi^0)$, Γ_ω, and an accurate matrix element calculation, the selection has an efficiency $25.02 \pm 0.07\%$ for ω's and a background of $61.6 \pm 0.3\%$ from τ decays other than $\tau \to 3\pi\pi^0$. The majority of the ω sample is lost in the first four cuts due to detector geometry and misidentification of the tau. The cuts imposed in order to clean away background affect the signal minimally. Table 6.2 shows the same cuts for data. The percentage lost in the data agrees with Monte Carlo loss after the "good" status cut is performed (an overall loss of $\sim$8.8% for 1994 data and $\sim$8.6% for MC in statistics between the status cut and the end of the hard cuts). The background from non-tau decays is minimal in the analysis sample.

The last two cuts are shown in figures 6.1 and 6.2. No significant loss is made in the $3\pi\pi^0$ samples while dramatically cutting the other τ pair backgrounds. The final sample after all the hard cuts should contain clean, 3-prong τ decays with at least one unassigned neutral cluster of some significance. An example of an event with both τ decays passing the cuts is figure 6.3.

From the standard τ pair Monte Carlo, there is no way of determining the efficiency of the $\omega\pi$ and non-ω decays in the $3\pi\pi^0$ sample. The individual decay channel MC's show that both the ω and non-ω decays have efficiencies of approximately 25%, while the $3\pi\pi^0$ sample in the general MC has an efficiency of $33.02 \pm 0.20\%$. This is due to some of the initial cuts made on the sample. The specialised MC's consist of back-to-back $3\pi\pi^0$ decays, thus forcing the initial cut of $N_{cone} = 2$ to remove a larger sample than a standard MC that would preferentially generate $3\pi\pi^0$ decays in a 3-1 topology. What can be inferred is that the two styles of $3\pi\pi^0$ decay have the same efficiency, and therefore, in the standard sample, should have an efficiency close to 33%. The proportions of each decay channel must be determined through statistical fitting. An estimate can be made by cutting on $M_\omega = 781.94 \pm 0.12$ MeV from the true 4-vectors of the general $3\pi\pi^0$ Monte Carlo sample. This gives 12.24% for $\omega\pi$'s and 26.43% for non-ω $3\pi\pi^0$ decays in the final analysis sample. This is an *underestimate* of the proportion of ω's in the sample due to the broad spread of the $M_{inv}(\omega)$ outside of the sharp resonance, even before taking into account the resolution of the detector.

6.3 Primed for Fitting

There now exists a subset of the initial sample that should provide enough information to extract branching ratios for $\tau^- \to \omega\pi^-\nu_\tau$ as well as some of the background decays.

6.3.1 Useful Variables

An attempt was made to obtain all discriminating variables in order to facilitate signal and background separation. In total, 34 variables were investigated ranging from invariant mass calculations to the reconstructed tau direction using the Kühn method [54]. This sample was pruned to 16 variables with maximum separation of the Monte Carlo decays, concentrating on the $\omega\pi$ and similar decays. The list below describes the variables used. The limits (in parentheses) assigned to several of them remove rogue events that pass the initial cuts, but still stray far from the smooth distributions.

1. Number of unassigned neutral clusters: N_{unass}.

2. Total energy in the cone divided by the total charged track momentum: $\sum E_{total}/\sum p_{track}$ (< 3.0 GeV).

3. Invariant mass of the three charged pions: $M_{inv}(3\pi)$.

4. Energy of the unassigned neutral clusters divided by the total charged track momentum: $\sum E_{unass}/\sum p_{track}$.

5. Invariant mass of the $2\pi^+\pi^-\pi^0$ system: $M_{inv}(3\pi\pi^0)$ (< 4.0 GeV). The determination of π^0 candidates will be discussed below.

6. *Best* combination invariant mass of $\pi^+\pi^-\pi^0$: $M_{inv}(\omega)$, i.e. closest to the accepted ω mass.

7. Other combination invariant mass of $\pi^+\pi^-\pi^0$: $M_{inv}(X\omega)$.

8. Energy of the most energetic unassigned neutral cluster: E_{γ_1}.

9. Highest energy $\pi^+\pi^-$ invariant mass combination: $M_{inv}(2\pi_{high})$ (< 2.0 GeV).

10. Other combination of $\pi^+\pi^-$: $M_{inv}(2\pi_{low})$.

11. Invariant mass of ω candidate and the lone pion: $M_{inv}(\omega\pi)$. The *best* omega is used and its mass is constrained to the world average.

12. Invariant mass of the other $\omega\pi$ combination: $M_{inv}(X\omega\pi)$.

13. Energy of the two highest energy unassigned neutral clusters: $E_{\gamma_1\gamma_2}$.

14. Energy of the lone pion: $E_{lone\ \pi}$ (< 50.0 GeV).

15. Total momentum in a cone: $\sum |p_{all}|$ (< 92.0 GeV).

16. Charged momentum in a cone: $\sum |p_{charged}|$.

Table 6.3 shows the effect of these additional constraints on the number of events available for analysis. The Monte Carlo sample is split into 4 classes of particle types that pass through the above selection: $\omega\pi$, $3\pi\pi^0$(*no* ω), 3π, and *everything else*. The first three names sufficiently describe the composition of the subsamples. The *everything else* subset contains the decays mentioned in table 6.4. The differentiation between the types of background closely modelling the $\omega\pi$ decay assists in calculating the branching ratios as well as provides a cross check of its accuracy.

Data and Monte Carlo agree(figure 6.4). The proportions of ω and non-ω decays in these plots is roughly determined for illustration by taking the true MC mass from the standard $3\pi\pi^0$ sample and cutting on the world average mass of the ω. The true $\omega\pi$ and $3\pi\pi^0$(*no* ω) samples are then used in the fits with the estimates and the measured $\tau \rightarrow 3\pi\pi^0$ efficiency.

6.3.2 π^0 Identification

Almost 99% of the time, a π^0 will decay into 2 photons. While the maximum entropy clustering algorithm (Appendix B) used to enhance the electromagnetic calorimeter's image resolution is capable of distinguishing between one and two photons to some extent, a large proportion of the time, the two photons are resolved into a single cluster. This is due to the fact that the distance between the photons in an average pion decay is of the order of the dimensions of an ECAL block (see Chapter 4). Therefore, single unassigned neutral clusters are considered to be an unresolved 2γ system for the purpose of this analysis. The clustering algorithm does not generally err in the direction of too many clusters (figure 6.5). It is this 2γ system that is taken as a decay product of the ω candidate. The neutral cluster system's 4-vector is constrained to the π^0 mass for invariant mass calculations.

6.4 The Problem

Up to this point, no distinction has been made between ω and non-ω $3\pi\pi^0$ samples. This is no surprise. Once that beautifully sharp ω resonance in figure 5.1 is pushed through the detector, it becomes less than impressive (figure 6.6). In fact, when all the remaining signal and background are combined, the peak at the ω resonance is barely noticeable (figure 6.7). The invariant mass of a reconstructed omega is the variable with largest signal to background separation available, yet no cut can be made on it to produce a reasonable sample of omega candidates. The background would swamp the signal. Something more powerful than a simple series of one dimensional cuts needs to be implemented. All the available information about an event must be taken into account to obtain any semblance of a $\tau \to \omega\pi\nu$ branching ratio. And this is where statistics steps in...

Cut Imposed		*% Accepted*	*% of Sample*
$N_{cone} = 2$:	$\omega\pi$	73.69 ± 0.07	4.69 ± 0.02
	$3\pi\pi^0$ (no ω)	73.54 ± 0.07	$\uparrow$ combined
	3π ($a1$)	82.3 ± 0.1	7.78 ± 0.03
	everything else	81.17 ± 0.04	87.53 ± 0.03
τ ID :	$\omega\pi$	64.25 ± 0.08	4.62 ± 0.02
	$3\pi\pi^0$ (no ω)	61.22 ± 0.08	$\uparrow$ combined
	3π ($a1$)	74.9 ± 0.1	7.70 ± 0.03
	everything else	74.72 ± 0.04	87.68 ± 0.03
"Tight" Selection :	$\omega\pi$	47.12 ± 0.08	4.52 ± 0.02
	$3\pi\pi^0$ (no ω)	45.24 ± 0.084	$\uparrow$ combined
	3π ($a1$)	70.4 ± 0.1	7.68 ± 0.03
	everything else	70.55 ± 0.04	87.80 ± 0.03
In the Barrel :	$\omega\pi$	36.29 ± 0.08	4.73 ± 0.03
	$3\pi\pi^0$ (no ω)	35.92 ± 0.08	$\uparrow$ combined
	3π ($a1$)	49.5 ± 0.2	8.04 ± 0.04
	everything else	47.12 ± 0.05	87.23 ± 0.04
Status "Good" :			NA
3 Charged Tracks :	$\omega\pi$	32.37 ± 0.07	24.3 ± 0.1
	$3\pi\pi^0$ (no ω)	31.02 ± 0.07	$\uparrow$ combined
	3π ($a1$)	43.4 ± 0.2	45.3 ± 0.2
	everything else	2.56 ± 0.02	30.5 ± 0.2
No Conversions :	$\omega\pi$	26.44 ± 0.07	29.7 ± 0.2
	$3\pi\pi^0$ (no ω)	26.29 ± 0.07	$\uparrow$ combined
	3π ($a1$)	40.1 ± 0.2	57.6 ± 0.2
	everything else	0.796 ± 0.008	12.8 ± 0.1
≥ 1 Neutral Cluster :	$\omega\pi$	26.07 ± 0.07	32.8 ± 0.2
	$3\pi\pi^0$ (no ω)	26.08 ± 0.07	$\uparrow$ combined
	3π ($a1$)	34.4 ± 0.2	54.0 ± 0.2
	everything else	0.741 ± 0.004	13.3 ± 0.1
$E_{unass_{12}} > 1.0$ GeV :	$\omega\pi$	25.02 ± 0.07	38.3 ± 0.2
	$3\pi\pi^0$ (no ω)	25.42 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.9 ± 0.2	47.1 ± 0.2
	everything else	0.673 ± 0.008	14.5 ± 0.2

Table 6.1: The cumulative effects of the cuts imposed on 600K MC τ-pair events are shown here. The effects of the cuts for ω and non-ω $3\pi\pi^0$ events are taken from their respective 200K MC samples. Only the selection efficiency of the *combined* $3\pi\pi^0$ sample can be determined due to decay generation and identification in the Monte Carlo.

Cut Imposed	*% Accepted*
Initial Sample :	2166740
$N_{cone} = 2$:	48.90 ± 0.03
τ ID :	26.12 ± 0.03
"Tight" Selection :	6.06 ± 0.02
In the Barrel :	3.82 ± 0.01
Status "Good" :	3.46 ± 0.01
3 Charged Tracks :	0.561 ± 0.005
No Conversions :	0.409 ± 0.004
≥ 1 Neutral Cluster :	0.370 ± 0.004
$E_{unass_{12}} > 1.0$ GeV :	0.305 ± 0.004

Table 6.2: Here, the same cuts are applied to the 1994 data. Assuming that the resultant sample contains only tau decays, 3.2% of the τ decays are kept for the sample. This is in good agreement with 3.4% of the MC τ's left after the hard cuts.

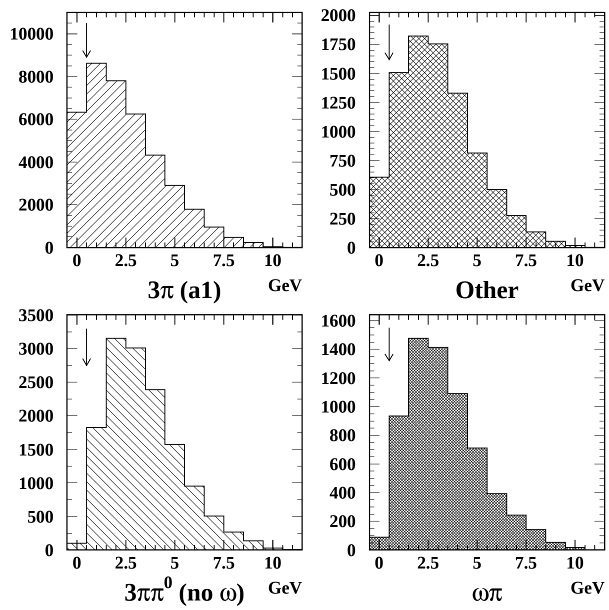

Figure 6.1: Requiring at least one unassigned neutral cluster to be present in a cone (cut shown as an arrow) removes a significant amount of background without greatly affecting the ω signal. The detector generally finds at least one of the photons from a π^0 decay.

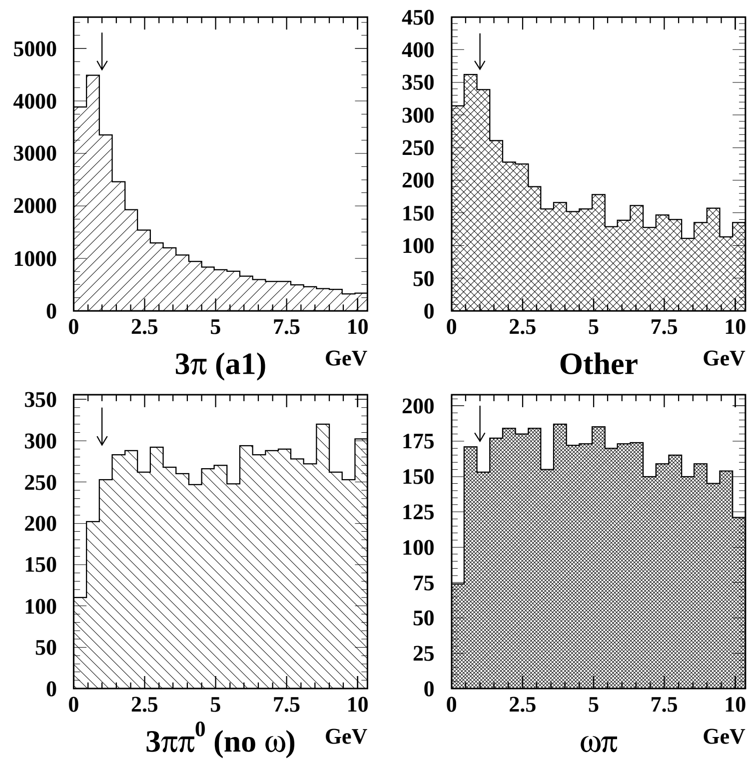

Figure 6.2: The electromagnetic energy measured in the highest energy unassigned neutral cluster in 3π decays and other backgrounds is peaked near 0.0 GeV. This is due to soft initial and final state radiation being identified as neutral clusters. The cut of 1.0 GeV is shown by the arrow for each sample.

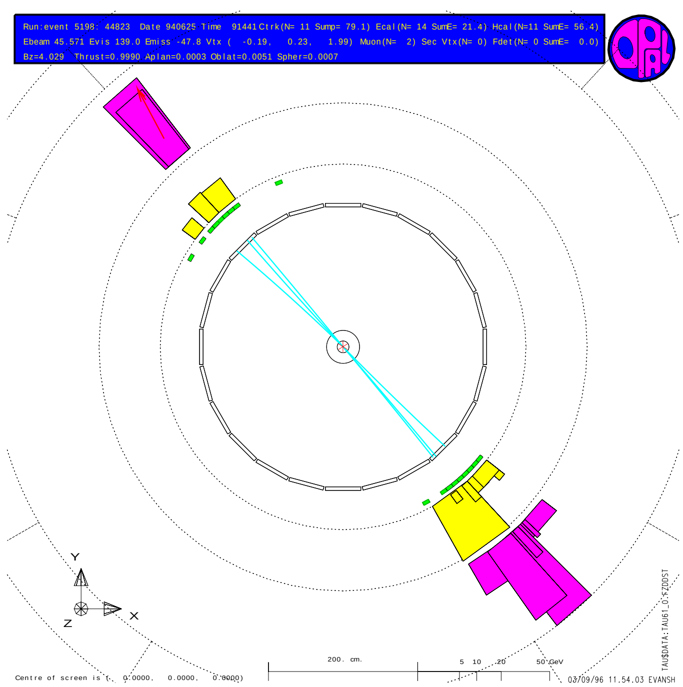

Figure 6.3: An $e^+e^- \to \tau^+\tau^-$ event where both taus decay into three charged (and at least one neutral) particles.

Cut Imposed		*% Accepted*	*% Sample*
$E_{total}/p_{track} < 3.0$ GeV :	$\omega\pi$	24.95 ± 0.07	38.5 ± 0.2
	$3\pi\pi^0$ (no ω)	25.13 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.9 ± 0.1	47.6 ± 0.2
	everything else	0.638 ± 0.008	13.9 ± 0.2
	1994 data	0.302 ± 0.004	
$M_{inv}(3\pi\pi^0) < 4.0$ GeV :	$\omega\pi$	24.61 ± 0.07	38.6 ± 0.2
	$3\pi\pi^0$ (no ω)	24.82 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.4 ± 0.1	47.4 ± 0.2
	everything else	0.629 ± 0.008	13.9 ± 0.2
	1994 data	0.296 ± 0.004	
$M_{inv}(2\pi_{high}) < 2.0$ GeV :	$\omega\pi$	24.50 ± 0.07	38.6 ± 0.2
	$3\pi\pi^0$ (no ω)	24.71 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.3 ± 0.1	47.4 ± 0.2
	everything else	0.625 ± 0.007	13.9 ± 0.2
	1994 data	0.295 ± 0.004	
$E_{lone} < 50.0$ GeV :	$\omega\pi$	24.43 ± 0.07	38.7 ± 0.2
	$3\pi\pi^0$ (no ω)	24.654 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.2 ± 0.1	47.5 ± 0.2
	everything else	0.621 ± 0.007	13.9 ± 0.2
	1994 data	0.294 ± 0.004	
$\sum \lvert p_{all} \rvert < 92.0$ GeV :	$\omega\pi$	24.41 ± 0.07	38.7 ± 0.2
	$3\pi\pi^0$ (no ω)	24.63 ± 0.07	$\uparrow$ combined
	3π ($a1$)	24.2 ± 0.1	47.5 ± 0.2
	everything else	0.620 ± 0.007	13.9 ± 0.2
	1994 data	0.294 ± 0.004	

Table 6.3: Imposing constraints throughout the analysis to remove poorly measured events has little effect on the efficiency of the selection. The $\omega\pi$ sample only loses an additional 0.5%.

Decay Channel	*% of Sample*
$\pi^-\pi^0$ (ρ)	23.41
$\pi^-\pi^+\pi^-2\pi^0$	20.23
$\pi^-2\pi^0$	15.64
$K^-\pi^+\pi^-$	9.04
$K^-K^+\pi^-$	7.45
π^-K^0 (K^*)	7.37
$\pi^-3\pi^0$	4.59
K^0K^-	2.35
$K^0\pi^-\bar{K}^0$	2.07
$\pi^-\bar{K}^0\pi^0$	2.00
$\pi^-\pi^+\pi^-3\pi^0$	1.93
$K^-\pi^0K^0$	1.55
π^-	0.59
$K^-\pi^0$ (K^*)	0.41
e^-	0.39
$\pi^-\pi^+\pi^-\pi^+\pi^-$	0.27
$\pi^-\pi^+\pi^-\pi^+\pi^-\pi^0$	0.18
$2\pi^0K^-$	0.12
$\pi^-\pi^02\gamma$	0.10
$\pi^-\pi^0\gamma$	0.09
$\pi^-\pi^03\pi^0$ ($\pi^-\pi^0\eta$)	0.09
μ^-	0.06
K^-	0.04
Unassigned	0.04

Table 6.4: A breakdown of the *everything else* background category. The majority of the background is from other decays with one or more π^0's. The *Unassigned* category compensates for inadequacies in the Monte Carlo modelling.

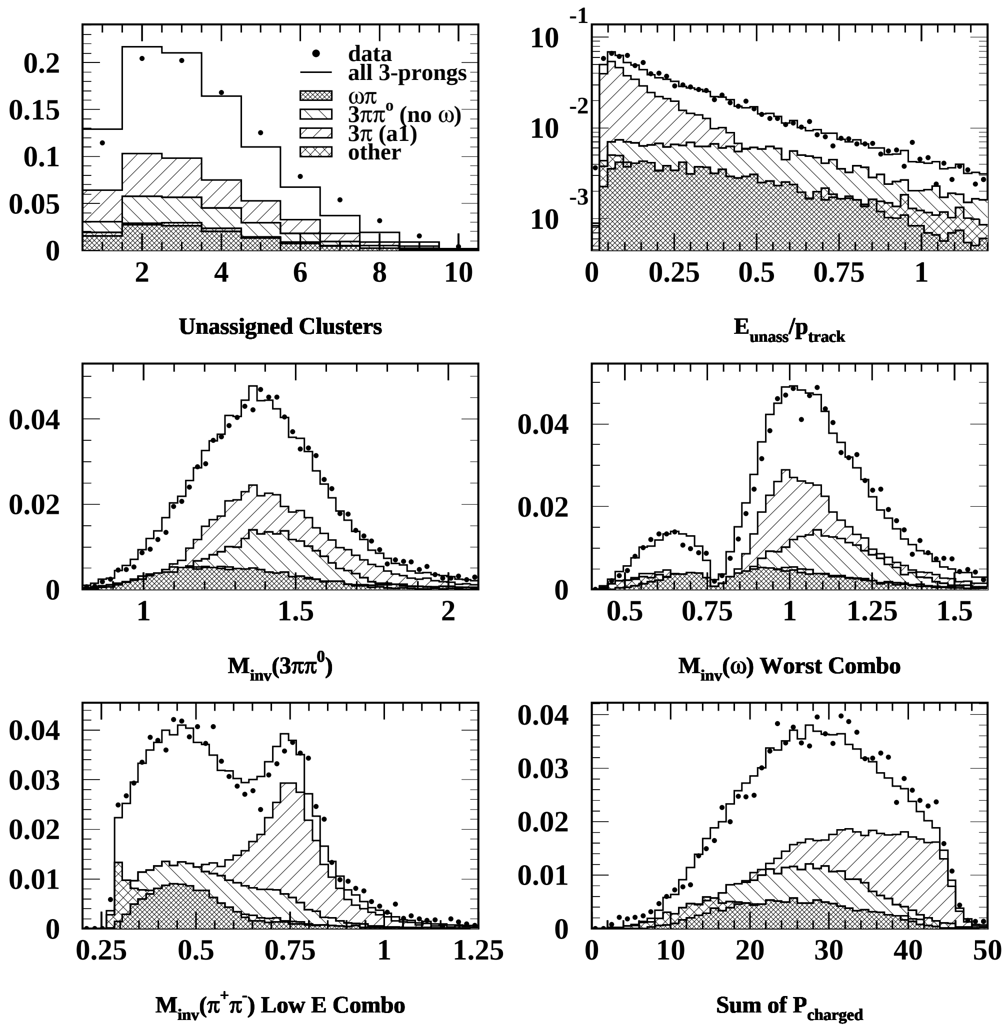

Figure 6.4: Data and Monte Carlo are compared for several of the discriminating variables. Data are the points and MC, the histograms. The shading corresponds to the possible decay categories. The agreement between data and MC appears reasonable.

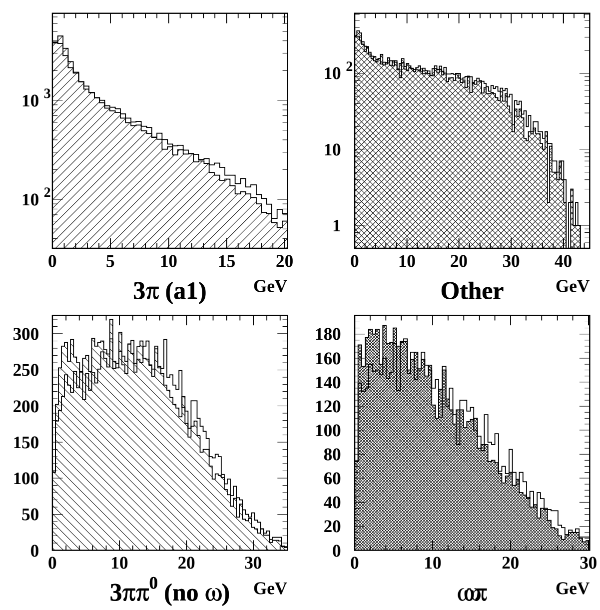

Figure 6.5: For each of the four decay channel categories, the overall energy deposited in the ECAL is determined (unshaded). Little energy is lost if the energy deposited is calculated from only the two most energetic clusters (shaded).

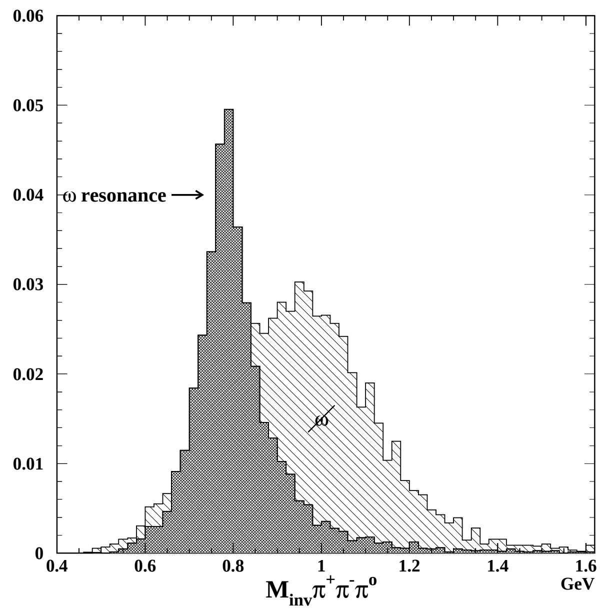

Figure 6.6: For 200K samples of $\tau \rightarrow 3\pi\pi^0\nu_\tau(no\ \omega)$ and $\omega\pi$ MC decays, the $\pi^+\pi^-\pi^0$ combination that comes closest to an ω reconstructed mass is determined. The invariant mass of this $3\pi\pi^0$ combination is much more smeared than the invariant mass for the the Monte Carlo truth 4-vectors.

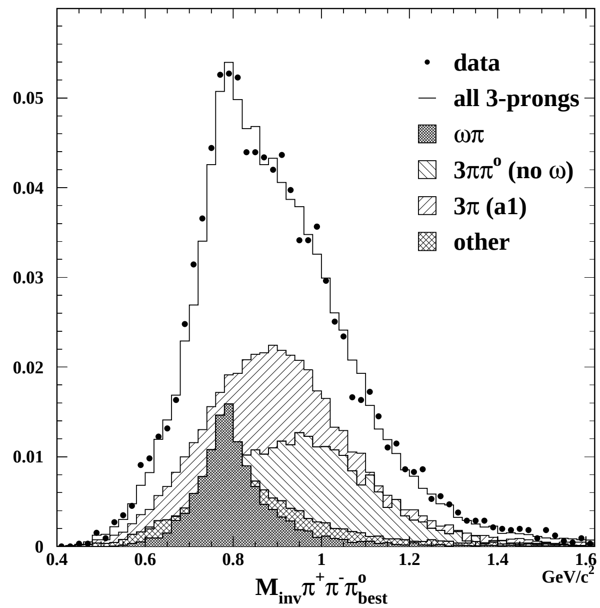

Figure 6.7: With the inclusion of the rest of the background decay channels, the $\pi^+\pi^-\pi^0$ invariant mass distribution has only a minimal peak from the ω resonance (shaded) at M_ω of 782 MeV.

Chapter 7

Statistical Methods

While hard cuts on individual variables may not be sufficient to separate decay channels in a sample, there is no need for despair. The very shape of the distributions can be used to achieve branching ratios.

During the course of the analysis, several methods of statistical manipulation have been implemented. Throughout this chapter, a specific method or twist on a method is introduced. Once a mathematical construct is in place, its effect on the analysis is discussed.

7.1 The Principle of Least Squares

One of the most basic ways to determine how well a mathematical model matches the available data is by determining χ^2, a measure of the model closeness to data [55].

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i; a)}{\sigma_i} \right]^2 \tag{7.1}$$

χ^2 is a sum over the difference between the measurements y_i and the prediction $f(x_i; a)$ from precise values x_i. Scaling by the error, σ_i, normalises the difference to the expected accuracy. If the derivative of $f(x_i;a)$ is known, a least squares fit can be performed by calculating the approximation of the parameter a solving

$$\frac{d\chi^2}{da} = \sum_i 2\frac{1}{\sigma_i^2}\frac{df(x_i;a)}{da}[y_i - f(x_i;a)] = 0. \tag{7.2}$$

This estimate of a, $\hat{a}$, should come close to the true value of a.

The *principle of least squares* can be derived from the principle of maximum likelihood soon to be described in section 7.2, but it should be apparent that the difference between data and prediction squared would be minimum where the two are reasonably equal. $\frac{\chi^2}{N}$ should be close to one, making the difference of the observed values and the prediction of the order of the known accuracy.

7.1.1 Fitting the Data

If there are several parameters $\vec{a}$ being adjusted to minimise the χ^2, then a series of equations are solved simultaneously. In the case of the 1994 data, the prediction of the data's overall distribution for a variable is a weighted sum from normalised MC distributions. The function for bin k and the 4 subsets of the Monte Carlo: $\omega\pi$, $3\pi\pi^0$ (no ω), 3π, and *everything else* tau decays is

$$f = \sum_{i=1}^{4} w_i m_{ik}. \tag{7.3}$$

The elements of $\vec{w}$ are the relative proportions of the decays in the sample. $\sum_i \hat{w}_i = 1$ for a sensible fit. m_{ik} is the number of entries in bin k from normalised distribution i. The χ^2 becomes

$$\chi^2 = \sum_{k=1}^{bins} \frac{(n_k - \sum_{i=1}^{4} w_i m_{ik})^2}{n_k}. \tag{7.4}$$

The errors $\sqrt{n_k}$ are Poisson. n_k is the number of normalised data in bin k, and is used for the error to ease derivation, since it should be a reasonable estimate of the number of normalised MC entries in a given bin. Any bins with $n_k \leq 5$ are disregarded.

$$\frac{d\chi^2}{dw_i} = \sum_{k=1}^{bins} 2m_{ik} \frac{(n_k - \sum_{j=1}^{4} w_j m_{jk})}{n_k} \tag{7.5}$$

When this derivative is 0, χ^2 is minimised.

$$\sum_{k=1}^{bins} \frac{m_{ik} \sum_{j=1}^{4} \hat{w}_j m_{jk}}{n_k} = \sum_{k=1}^{bins} m_{ik} \tag{7.6}$$

The histograms are all normalised, and so the right-hand side is equal to 1. The simultaneous equations for all i can be described by

$$M\vec{\hat{w}} = \vec{u}. \tag{7.7}$$

$\vec{u}$ is a vector of units, and matrix M has elements

$$M_{ij} = \sum_{k=1}^{bins} \frac{m_{ik} m_{jk}}{n_k}. \tag{7.8}$$

Inverting the matrix, an estimate of $\vec{w}$ can be determined.

$$\vec{\hat{w}} = M^{-1}\vec{u} \tag{7.9}$$

Figure 7.1 shows the result for the 16 variables described in section 6.3.1. The errors are determined with the assumption that the n_k in equation 7.4 is an estimate of the Monte Carlo error. Figure 7.2 contains the χ^2 per degree of freedom calculations. The proportions should stay between 0.0 and 1.0 to be physical, but there have been no restrictions made on $\vec{\hat{w}}$'s boundary values. The number of neutral clusters is poor at determining the appropriate proportions of decays in the data sample, due to its limited range. There are only a maximum of 10 unassigned neutral clusters in a decay, and therefore, only 10 bins over which the fit can be taken, with little distinction between the two $3\pi\pi^0$ samples. The range of $\vec{w}$ produced over the 16 variables is indicative of the difficulty of distinguishing between the decays. A combination of the fits cannot be made, because the correlations between the variables have not been taken into consideration, and the fit performed using the number of unassigned neutral clusters in a decay will drive $\vec{w}$ to provide unphysical fractions of $\omega\pi$'s in the data.

The information from all the variables needs to be used simultaneously to determine more accurate decay fractions. Taking one step back statistically, and one step forward in complexity brings us to a standard log-likelihood fit.

7.2 Maximum Likelihood

One of the best methods of estimating a vector parameters, $\vec{w}$, with the variable set $\{x_1, x_2, x_3, ..., x_N\}$ is through the maximisation of a likelihood function, $\mathcal{L}$ [55][56]. Likelihood is a measure of the probability that the estimation $\vec{w}$ will give you $\vec{x}$. The object is to find the $\vec{w}$ the makes a given set of $\vec{x}$'s most likely.

7.2.1 Determining the Likelihood

Probability density functions, *pdf*'s, are used to determine the probability that the measurement of variable i, x_i, will lead to an outcome, j. These *pdf*'s are the same normalised histograms of MC distributions that were used for the least squares fit. The Monte Carlo contains the best prediction of how different decay channels behave, with the binning providing the probability of measuring each value of parameter. While previous analyses used various combinations of Breit-Wingers and Gaussians to fit their data, this is not as accurate as using the shapes of MC distributions which are the best models of interferences, etcetera, available.

Using Monte Carlo decay identification trees and *pdf*'s, the probability $p_{ij}(x_i)$ that outcome/decay j will produce value x for variable i is determined. The combined probability, P_j, is then the chance that (assuming uncorrelated p_{ij}'s) a certain decay would produce all the measured values of a specific event, an event being a $\tau \to \omega\pi\nu$ decay candidate that has passed selection cuts.

$$P_j(x_1, x_2, x_3, ..., x_N) = \prod_{i=1}^{N} p_{ij}(x_i) \tag{7.10}$$

An event likelihood is produced from the combined decay probabilities.

$$\mathcal{L}_{event}(x_1, x_2, x_3, ..., x_N; \vec{\hat{w}}) = \sum_{j=1}^{ndecay} \hat{w}_j P_j \tag{7.11}$$

$\sum_{j=1}^{ndecay} \hat{w}_j$ is constrained to 1. In this case, $\vec{\hat{w}}$ is an array of weights assigned to each class of decay, i.e. $\omega\pi$, $3\pi\pi^0$ (no ω), 3π, and *everything else*. This will evolve into a series of branching ratios, but now, it is simply a set of proportions in the event samples accepted by the final selection.

Once the likelihood for a given event is calculated, the likelihood that these weights, $\vec{\hat{w}}$, will give the correct event distribution for a sample is straightforward. It is the product of the event likelihoods.

$$\mathcal{L}_{overall}(\vec{\hat{w}}) = \prod_{event=1}^{nevent} \mathcal{L}_{event} \tag{7.12}$$

7.2.2 Maximising the Likelihood

In order to get the most probable combination of weights, the above equation is "maximised". Due to the fact that critical points are found by differentiation, it would be advantageous to utilise an easily differentiable function. Maximising a function, f, is the same as maximising f^2, f-2, $\log f$, etc. The log of the likelihood function is chosen so that gradients will be trivial to calculate.

$$\log \mathcal{L}_{overall} = \sum_{event=1}^{nevent} \log \mathcal{L}_{event} \tag{7.13}$$

A function *minimising* program is used to achieve the maximal combination of probability weights. MINUIT [57] is capable of taking a multiparameter function specified by the user and finding its minimum through gradient descent. In this case, $-\log \mathcal{L}_{overall}$ is minimised, thus maximising $\log \mathcal{L}_{overall}$.

7.2.3 Results of the Basic Maximum Likelihood Method

The kinematic variables mentioned in Section 6.3.1 are the *pdf*'s for the initial maximum likelihood fits. These 16 variables have been determined to be the best at separation between decay channels. Below, a maximum likelihood fit is also performed on neural net outputs (Section 7.4) generated from these initial variables. The outcome j represents $\omega\pi$, $3\pi\pi^0$ (no ω), 3π, or *everything else* decay channels. The array of estimated weights, $\vec{\hat{w}}$, is the approximate proportion of each decay channel in a sample, the only constraint to $\vec{\hat{w}}$ being

$$\sum_i \hat{w}_i = 1. \tag{7.14}$$

Negative weights are still allowed. This constraint makes it possible to reduce the weight array to three elements, $\vec{\hat{w}}'$. The hats will now be dropped, since the true decay proportions are unmeasurable. The relationship between $\vec{w}$ and $\vec{w}'$ is

$$\begin{aligned}
w_\omega &= w'_\omega \qquad (7.15)\\
w_{not\omega} &= (1 - w'_\omega) w'_{not\omega}\\
w_{3\pi} &= (1 - w'_\omega - w'_{not\omega}) w'_{3\pi}\\
w_{other} &= (1 - w'_\omega - w'_{not\omega} - w'_{3\pi}).
\end{aligned}$$

Only three parameters are now in the likelihood equation with the minimisation function as

$$-\log \mathcal{L}_{overall} = -\sum_{k=1}^{nevents} \log \left(\sum_{j=1}^{ndecay} w_j \prod_{i=1}^{N} p_{ijk}(x_{ik}) \right) . \tag{7.16}$$

Where w_j is a function of the elements of $\vec{w}'$.

The combined probabilities, P_{jk}, are calculated once for each tau decay candidate in the sample to be fitted. To assist in the minimisation, $\vec{w}$ is started out at a reasonable estimate of the percentages of each decay in the data, $\{12.7, 25.8, 33.1, 28.4\}$. This is taken from the middle of the road χ^2 fit for variable $M_{inv}(\omega)$, since this distribution shows the most separation for the different decay channels visually. The gradients are calculated with respect to each ${w_j}'$. To get a rough estimate of the position of the minimum, the simplex method of Nelder and Mead [58] is implemented. The simplex method is a genuine multidimensional minimisation routine, but it provides no information about the errors on $\vec{w}$ and cannot be expected to converge to the minimum. Simplex simply gets close to the minimum, and a gradient descent method is then utilised to actually reach it. The gradient descent method is most affective when getting close to the minimum itself. Once the minimum is determined, a search for additional local minima is performed to enhance the chances that the minimum found is absolute. The accuracy of the search is fixed to provide "one-standard deviation" errors on each of the parameters for the log-likelihood fit. The errors are determined in two ways. A cumulative error matrix from the gradient descent method provides parabolic errors. These are compared to asymmetric errors obtained by following the function up until 1σ per parameter is achieved.

The results achieved by this minimalist maximum likelihood method are shown in table 7.1. While the numbers are close, the low quality of the fit can be seen with the comparison between the MC fit and the true values of decay proportions. Even

Variable Combination		*% Proportion (w_j)*
MC True values :	ω	12.2 (38.7 ± 0.2)
	$3\pi\pi^0$ (not ω)	26.5
	3π $(a1)$	47.5 ± 0.2
	everything else	13.9 ± 0.2
16 variable fit (1994 Data) :	ω	13.1 ± 1.2
	$3\pi\pi^0$ (not ω)	17.9 ± 1.2
	3π $(a1)$	55.8 ± 1.7
	everything else	13.1 ± 2.4
16 variables fit (MC) :	ω	14.22 ± 0.44
	$3\pi\pi^0$ (not ω)	18.69 ± 0.43
	3π $(a1)$	52.69 ± 0.61
	everything else	14.48 ± 0.86

Table 7.1: The table above shows a basic maximum likelihood fit using the τ pair, $\omega\pi$, and $3\pi\pi^0$ (no ω) Monte Carlo samples for the *pdf*'s. The 1994 data sample and corresponding MC generation have been fitted for $\vec{w}$. The true Monte Carlo values for $\tau \to 3\pi\pi^0\nu$ decays is only available for the entire $3\pi\pi^0$ sample (in parentheses).

in the ideal case of the Monte Carlo, the overall percentage of $3\pi\pi^0$ decays for $\vec{w}$ is over 10σ off $\vec{w}_{MC}$. Estimates for the ω and non-ω subsamples are made by making a hard cut on $M_{inv}(\omega)$'s published accuracy. This is an *underestimate* of the proportion of ω's in the sample due to the ignorance of the ω sample's tail (see figure 5.1).

With a standard maximum likelihood fit the poor performance has to be expected.

7.2.4 Maximum Likelihood Limitations

The limitations of a basic maximum likelihood method become clear in the case where all variables in $\vec{x}$ are perfectly correlated: $x_1 = x_2 = x_3 ... = x_N$. P_j becomes

$$P_j = p_j(x)^N. \tag{7.17}$$

Variable Combination		*Proportion (%)*	Correlation
$N_{neutrals}$ and M_ω (MC):	ω	12.2 ± 1.6	0.054
	$3\pi\pi^0$ (not ω)	26.8 ± 5.1	-0.021
	3π ($a1$)	47.0 ± 6.0	0.018
	everything else	14.0 ± 8.1	0.031
$\sum P_{total}$ and $\sum P_{charged}$ (MC):	ω	41.6 ± 2.3	0.730
	$3\pi\pi^0$ (not ω)	0.0 ± 2.3	0.600
	3π ($a1$)	48.3 ± 2.8	0.891
	everything else	10.1 ± 4.2	0.710

Table 7.2: Note that even with the Monte Carlo used for the reference histograms and input sample, the weights are in good agreement with the *truth* for small correlations and completely wrong for large ones.

While the combined probability that an event has decayed via a specific channel should remain constant irregardless of the number of times a *pdf* is consulted, the likelihood is significantly reduced with the introduction of highly correlated variables. The standard maximum likelihood method assumes that the variables used to determine the combined probability are *uncorrelated.*

An example of this phenomenon is found in table 7.2. Two variable combinations are shown. The number of unassigned neutral clusters is slightly correlated with the best estimate of the ω invariant mass ($\hat{\rho} \approx 0.05$). The total momentum of the event is highly correlated with the sum of the momentum of only the charged tracks ($\hat{\rho} \approx 0.73$). The difference between the two fits using these pairs of variables is marked. Less correlation leads to a better fit.

The above inadequacy of the maximum likelihood method can be side-stepped by introducing a strategy to take correlations into account.

7.3 Improved Maximum Likelihood

Correcting the combined probabilities has been shown to improve a fit by 20% or more with large correlations between input variables [59]. The improvement is made by using an n dimensional unit Gaussian probability distribution centred at the origin to correct for the effects of variable correlations. This Gaussian, $G(\vec{y})$, is described by an $n \times n$ covariance matrix, V.

$$G(\vec{y}) = (2\pi)^{-n/2}|V|^{-1/2}e^{-\frac{1}{2}\vec{y}^T V^{-1}\vec{y}} \tag{7.18}$$

$|V|$ is the determinant of the covariance matrix in y-space. $\vec{y}$ is the vector that describes a point in this space.

By measuring the means, variances, and covariances of the variables, the multidimensional probability function can be written down directly. Gaussian distributions are used for convenience. With the measurement of the error, σ, and the mean, μ, the shape of the function is known. The job is made even simpler by choosing unit Gaussian distribution functions centred at 0, $\sigma^2 = 1$ and $\mu = 0$.

None of the kinematic variables are Gaussian, but it is possible to map variable x_i from a difficult distribution function onto variable y_i of a Gaussian distribution. A monotonic function, $y_i(x_i)$, capable of mapping from x-space to y-space is

$$y_i(x_i) = \sqrt{2}erf^{-1}(2F(x_i) - 1). \tag{7.19}$$

Each variable x_i has a probability factor F assigned to it using the equation

$$F(x_i) = \frac{\int_{x_{i_{min}}}^{x_i} pdf(x_i')dx_i'}{\int_{x_{i_{min}}}^{x_{i_{max}}} pdf(x_i')dx_i'}. \tag{7.20}$$

The denominator normalises the probability factor, so $F(x_{i_{min}}) = 0$ and $F(x_{i_{max}}) = 1$. $F(x_i)$ is determined from a standard pdf, and erf^{-1} is an inverse error function. An approximation of erf^{-1} is made using a piecewise continuous approximation with Chebyshev coefficients [60]. $y(x)$ performs a one-to-one mapping of each variable x onto y-space. The probability density functions for the elements of $\vec{y}$ are given by equation 7.18.

This method does not produce an n dimensional Gaussian. It only guarantees that the one dimensional projections of the function will be Gaussian (see figure 7.3). Although this is not ideal, it is a significant improvement on ignoring the correlations altogether. The number of unassigned neutral clusters will have to be disregarded for this statistical method due to its discrete nature.

Just as there is a one to one mapping of x_i to y_i values, there is a mapping of probabilities. With $G(\vec{y})$ as a combined probability density function in y-space that accounts for correlations, a corrected combined probability in x-space, P', can be determined.

$$P'_j(\vec{x})d^n\vec{x} = G_j(\vec{y})d^n\vec{y} \tag{7.21}$$

The Gaussians for each decay channel, j, are different because of the shapes of each variable distribution in x-space. If this was not the case, there would be no hope in differentiating one decay from another. Therefore, there are four different covariance matrices tying the variables together and leading to four independently corrected probabilities.

Manipulating the equation above results in a Jacobian that relates the improved likelihood with the basic method.

$$\frac{P'_j(\vec{x})}{G_j(\vec{y})} = \frac{d^n\vec{y}}{d^n\vec{x}} = \prod_{i=1}^{n} \frac{\partial y_i}{\partial x_i} = \prod_{i=1}^{n} \frac{p_{ij}(x_i)}{g_{ij}(y_i)} \tag{7.22}$$

where $\prod_{i=1}^{n} p_{ij}(x_i) = P_j(\vec{x})$ and $g_{ij}(y_i)$ is the analogue from a one dimensional Gaussian for variable i and decay channel j. Therefore, the corrected probability is related to the original combined probability by a factor, $c(\vec{x})$.

$$P_j'(\vec{x}) = c(\vec{x})P_j(\vec{x}) = \frac{G_j(\vec{y})}{\prod_{i=1}^{n} g_{ij}(y_i)}P_j(\vec{x}) \tag{7.23}$$

With $g_{ij}(y_i)$ a unit Gaussian, the correction factor becomes

$$c(\vec{x}) = \frac{\exp^{-\frac{1}{2}\vec{y}_j^T(V_j-I)\vec{y}_j}}{|V_j|^{\frac{1}{2}}}. \tag{7.24}$$

The new and improved version of the combined probability can then replace the uncorrected probability in the likelihood equation of the previous section.

7.3.1 Results of Improved Maximum Likelihood

The improvement is marked (table 7.3). The estimate of the overall percentage of $3\pi\pi^0$ decays for $\vec{w}$ now lies within error of the Monte Carlo *true* value. The errors are slightly higher, due to the proper treatment of correlations among the kinematic variables. The increase of the percentage of $\omega\pi$ in the data fit is comparable to the $\sim 20\%$ loss from a hard cut made on the ω resonance to get an estimate of ω's in the sample. Unfortunately, there is still a bias in the fit causing the fitting method to overestimate the combined $3\pi\pi^0$ decay proportion by over 5%. While the data is close to the expected proportion of $3\pi\pi^0$ decays $(38.7 \pm 0.2\%)$ in the sample at $39.3 \pm 1.8\%$, the MC fit is much higher at $43.9 \pm 0.7\%$.

7.3.2 Possible Improvements on the Improvement

The maximum likelihood method of fitting parameters is quite general, and little can be done to improve upon the algorithm in addition to correcting for correlations. Any improvements would have to be made in either the modelling of decay channel behaviour or the choice of variables to use. Variable choice is the concentration, since

Variable Combination		*% Proportion (w_j)*
MC True values :	ω	12.2 (38.7 ± 0.2)
	$3\pi\pi^0$ (not ω)	26.5
	3π ($a1$)	47.5 ± 0.2
	everything else	13.9 ± 0.2
Fit Excluding $N_{neutral}$ (1994 Data) :	ω	16.7 ± 1.3
	$3\pi\pi^0$ (not ω)	22.6 ± 1.3
	3π ($a1$)	51.0 ± 1.7
	everything else	9.7 ± 2.5
Fit Excluding $N_{neutral}$ (MC) :	ω	17.34 ± 0.46
	$3\pi\pi^0$ (not ω)	26.53 ± 0.50
	3π ($a1$)	47.96 ± 0.64
	everything else	8.16 ± 0.94

Table 7.3: The corrected maximum likelihood fit should deal with kinematic variable correlations with reasonable accuracy. The improvement in the fit can be shown by comparing the combined $3\pi\pi^0$ Monte Carlo *true* value (parentheses) with the estimates produced by the fits.

the alterations available in Monte Carlo theory would make little difference to the distribution shape, and therefore will make no impact on the maximum likelihood fit.

The further you can separate your channels in an event sample, the better chance you have of reducing biases and determining the correct makeup of that sample. Kinematic variables are not the only variables that can differentiate between decay types.

7.4 Neural Networks

The information in the kinematic variables is capable of being manipulated into a form that enhances features inherent in the individual decay channels that may not be directly apparent in the kinematic variables themselves. Due to the similarity in the two $\tau^- \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$ decay samples, finding variables with the greatest distribution separation *before* maximising the likelihood becomes a necessity. Neural nets are capable of generating such feature distributions.

7.4.1 Brain Power

You do not have to look far to see that neural networks have their advantages. In the 1940's the functionality of the vertebrate central nervous system, a biological neural net, was described in detail. In the early 80's, the computational culture started aggressively searching for the artificial equivalent. In contrast to the von Neumann, digital, structure of a conventional computer, which is exceptional at numerical and informational manipulation and storage, a neural network is designed to "adapt" to its environment and "recognise" broad characteristics of that environment.

In the case of this analysis, a set of optimal features is extracted from a variable set, $\vec{x}$. These features take into account correlations, and should optimise the separation of decay samples, thus making it easier on the maximum likelihood fit.

7.4.2 Structure of the Net

A neural net consists of layers of nodes: an input layer $\vec{x}$, at least one hidden layer $\vec{h}$, and an output layer $\vec{o}$ [61] (figure 7.4). In a feed forward net, each node in a layer affects the value of the nodes in the layers ahead of it in the net. Therefore, $\vec{o}$ is a function of $\vec{h}$ which is in turn a function of $\vec{x}$ The relationship is governed by a non-linear transfer function g. The transfer function for a node i depends not only on the value of each node j that interacts with it, but the relative strength of the interaction w_{ij}.

$$o_i(\vec{x}) = g\left(\sum_{j=0} w_{ij} g\left(\sum_{k=0} w_{jk} x_k\right)\right) \tag{7.25}$$

$\vec{o}$ is a measurement of how close the characteristics of an event resemble those of a specific decay. An additional node, $j = 0$, at each layer takes account of the threshold signal needed to trigger the node like triggering a nerve. This is an effort to adjust for rogue events that mimic "interesting" events. The range of values for each of the elements of the array is $[0, 1]$. A sigmoid-shaped function is used as g.

$$g(a_i) = \frac{1}{2}[1 + \tanh(a_i/T)] \tag{7.26}$$

a_i is the linear sum of weights and node values in equation 7.25. Temperature T sets the steepness of rise from 0 to 1 of the function, i.e. the inverse gain. For the sake of this analysis, it is set to 1.

7.4.3 Optimising Weights

The above algorithm provides an alternative method of presenting the information available in order to enhance variable separation of decays. To determine the optimal weight matrix relating nodes across layers, the net must be *trained.*

Training is accomplished by minimising the summed square error function, E

$$E = \frac{1}{2} \sum_{event} \sum_{i} (o_i^{(event)} - t_i^{(event)})^2 \tag{7.27}$$

with respect to the weight matrix elements. $\vec{t}^{(event)}$, the target pattern, is the value of $\vec{o}^{(event)}$ if the modelling was exact and the net was perfectly trained. The gradient descent method updates the weights proportionally to $\partial E/\partial w_{ij}$ where

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial w_{ij}} \tag{7.28}$$

for $event$. $\partial o_i/\partial w_{ij}$ is the partial derivative of the nonlinear transfer function $g(\sum w_{ij} x_j)$, $g' \times x_j$. $\partial E/\partial o_i = (o_i - t_i)$ for each variable per event. A fraction of this differential dictated by the learning rate is subtracted from weight w_{ij} until the correction becomes sufficiently small.

7.4.4 Neural Nets in Practice

JETNET [62], a standard neural network package, is used to train and implement a neural net for this analysis. 15 variables in Section 6.3.1, disregarding the number of unassigned neutral clusters, provide the input vectors. Before a final set of neural net optimal feature outputs are generated, the net had to be steered in the right direction.

Due to the fact that the $\omega\pi$ and non-ω $3\pi\pi^0$ decays are virtually indistinguishable except in $M_{inv}(\omega)$ measurements, it is difficult for the net to find the correct minimum in feature space. Training the sample from completely random initial weights, an efficiency of only $\sim 2.6\%$ for successfully classifying the event is achieved for the ω sample. In order to facilitate the minimisation and to increase the accuracy of the net, a smaller net is trained and then hand packed into the larger net. The gentle guidance provides some initial separation for the $3\pi\pi^0$ sample while being trained on the entire background. This *mini*net includes all input nodes, half of the hidden nodes and only the $3\pi\pi^0$ output nodes (see figure 7.4). It is trained and tested on two random, independent 100K normalised data sets from the 200K samples of $\omega\pi$ and $3\pi\pi^0$ (no ω) events. 77.4 ± 0.1 % of the time, the $\omega\pi$ events are successfully identified, while the efficiency is 66.2 ± 0.2 % for non-ω $3\pi\pi^0$ events. Figure 7.5 shows the separation. The sharp cutoffs seen are indicative of the fact that the net is incapable of positively identifying $3\pi\pi^0$ events with 100% accuracy due to the kinematic similarities of the decay channels [63].

The net is gently prodded into the desired region of feature space by hardwiring the $3\pi\pi^0$ net weights into their appropriate positions in the full matrix, W. The full

True Decay Class		*Identification Efficiency (%)*
MC True values :	ω	40.8 ± 0.6
	$3\pi\pi^0$ (not ω)	47.2 ± 0.4
	3π ($a1$)	63.2 ± 0.3
	everything else	49.2 ± 0.6

Table 7.4: The proportions of decays in the test half of the MC for which the correct net label is assigned is shown above. While there is still difficulty distinguishing between $3\pi\pi^0$ subclasses, the efficiency is significantly higher than the random success of identification, 25%.

net, with double the hidden nodes and 4 output nodes representing all decay types, is then trained with all nodes free to adjust. The results of the tiered training can be seen in figures 7.6 and 7.7. Table 7.4 shows the efficiency of successfully identifying the four classes of tau decays.

7.4.5 Effect on Improved Maximum Likelihood

The proper way of dealing with the neural net outputs is to construct four likelihoods, each with a single net output node as input. While neural nets are widely believed to deal with input correlations correctly [64], they produce outputs that are perfectly correlated. From each of these likelihoods, one element of $\vec{w}$ is calculated at a time. This provides the maximum likelihood algorithm with an input that enhances one channel identification at a time. This affectively fits for $\omega\pi/other$, $3\pi/other$, etc. The test is to see whether these elements obey $\sum \vec{w}_j = 1$.

The results can be found in table 7.5. They are comparable with the other methods of measurement, and correlations should have been properly dealt with. The individual likelihoods for each weight appears to discriminate between decay types quite well while maintaining the unity restriction within error. Note the discrepancy between the data and Monte Carlo fits for the separation of the $3\pi\pi^0$ samples. This

Variable Combination		*% Proportion (w_j)*
MC True values :	ω	12.2 (38.7 ± 0.2)
	$3\pi\pi^0$ (not ω)	26.5
	3π ($a1$)	47.5 ± 0.2
	everything else	13.9 ± 0.2
Individual Net Outputs (Data) :	ω	15.6 ± 2.5
	$3\pi\pi^0$ (not ω)	19.4 ± 2.1
	3π ($a1$)	52.9 ± 7.9
	everything else	12.2 ± 5.1
Individual Net Outputs (MC) :	ω	12.26 ± 0.91
	$3\pi\pi^0$ (not ω)	26.43 ± 0.83
	3π ($a1$)	47.46 ± 2.84
	everything else	13.86 ± 1.88

Table 7.5: The neural net as maximum likelihood inputs has success in determining the proportion of $\tau \to \omega\pi\nu$ in the sample as well as removing the bias in the Monte Carlo kinematic fit.

difference is within error, and the bias from the kinematic input variables has been removed. The combined $3\pi\pi^0$ percentages are now 35.0 ± 3.3 and 38.7 ± 0.1 for data and MC respectively.

7.5 Fisher Discriminants

In addition to the statistical methods already mentioned, an attempt was made at determining proportions of decays in the analysis sample using Fisher Discriminants. While linear discriminants were shown to be inappropriate for this study, a description of the results can be found in Appendix C.

Figure 7.1: The results of the χ^2 fractional fit for each of the 16 variables, assuming 4 classes of decay are shown. The number of neutral clusters is a poor discriminator, especially between the different classes of $3\pi\pi^0$ decays. While there appears to be consistency with the 3π and *everything else* decays, there is little proof that individual variables can identify $\omega\pi$ decays in a $3\pi\pi^0$ sample.

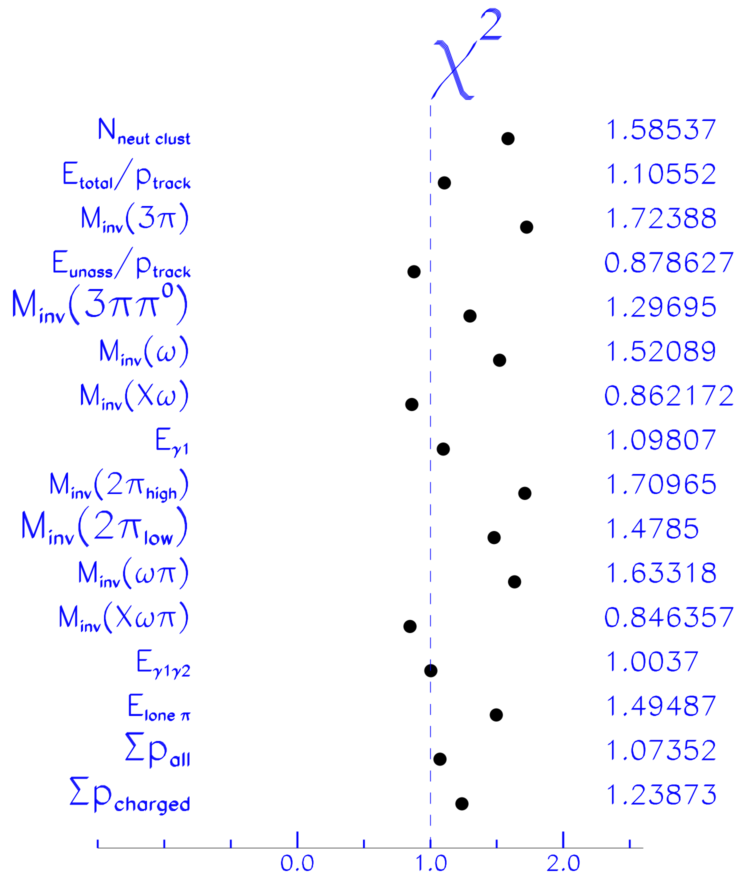

Figure 7.2: The $\chi^2/d.o.f$ does not provide much information about which variables have strong discriminating power. This is due to the fact that for most variables, there is little distinction between $\tau \to 3\pi\pi^0\nu_\tau$ decays that pass through an ω resonance and those that do not.

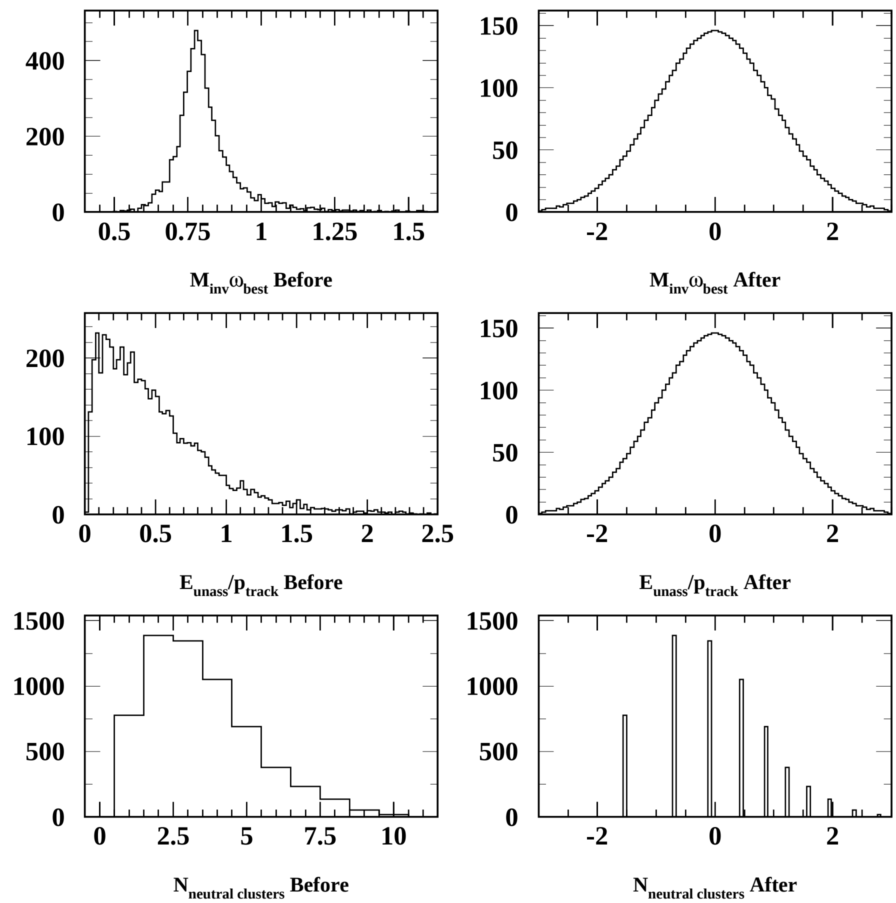

Figure 7.3: Above are 3 examples of distributions in x-space and their y-space "Gaussian" mappings. Since $N_{neutral\ clusters}$ is an integer, the mapping is not continuous. The program still does its best to move the integer values to a distribution that has $\sigma = 1$ and $\mu = 0$.

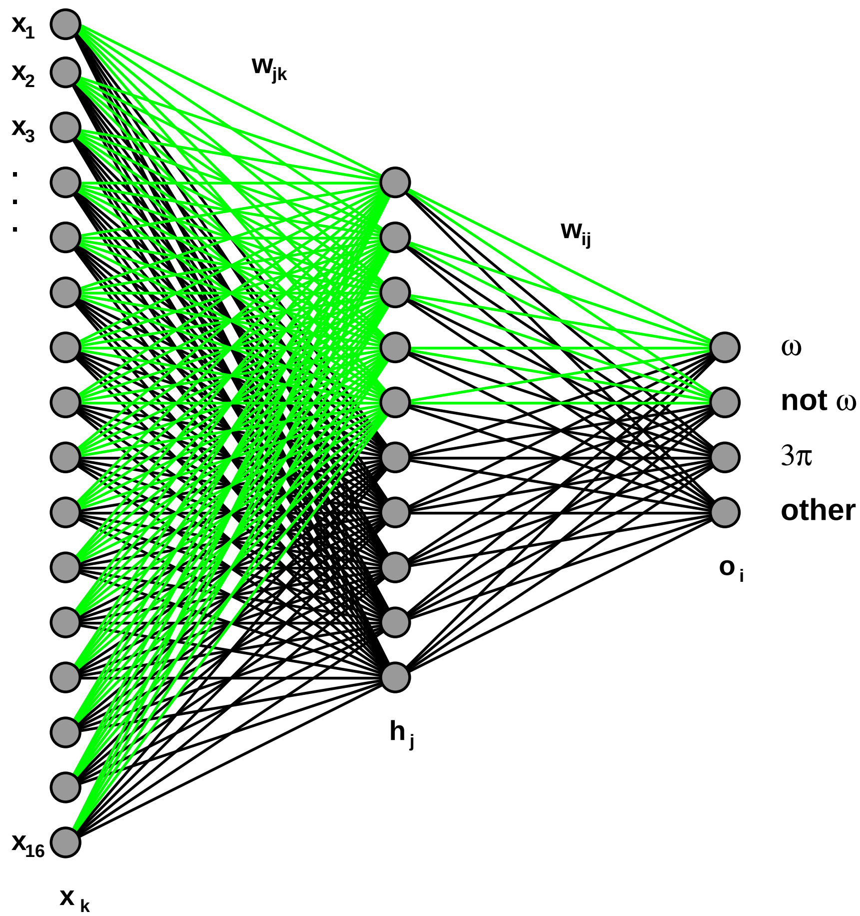

Figure 7.4: A standard neural net schematic. This is actually a superposition of two nets: $3\pi\pi^0$ subclassification (grey) and full $\tau \rightarrow \omega\pi\nu$ candidate subclassification (complete net). The initial net was optimised to identify ω decays in a $3\pi\pi^0$ sample. The weight matrix associated with that net was then hardwired into a complete net before training.

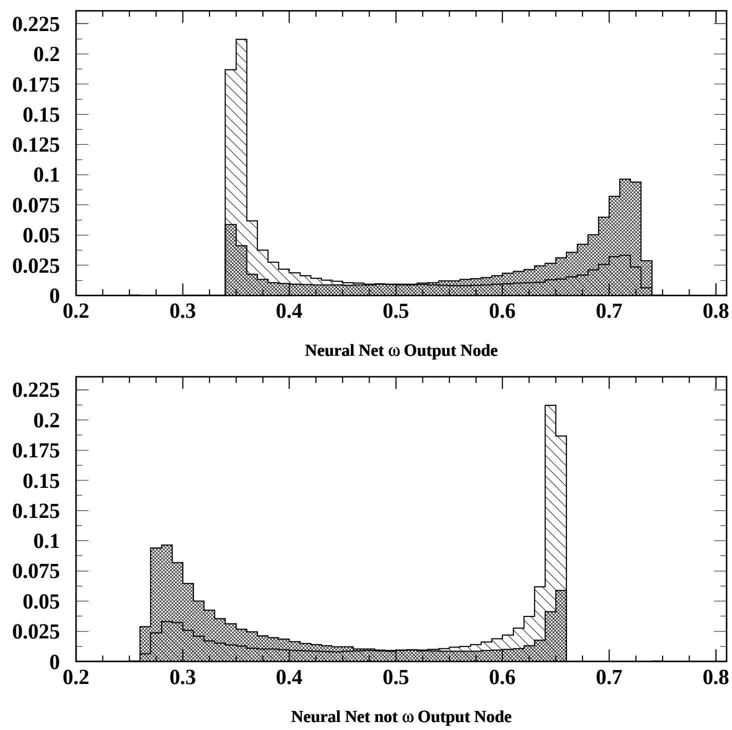

Figure 7.5: The separation between the ω resonance (shaded) and the rest of the $3\pi\pi^0$ sample (hatched) is markedly enhanced with neural net feature extraction.

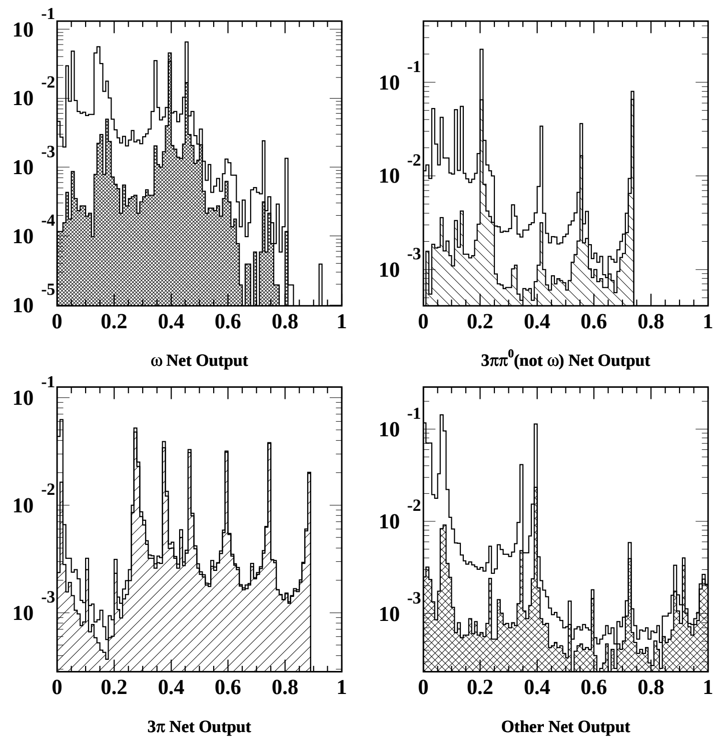

Figure 7.6: The outputs for the neural net trained and run on the entire Monte Carlo event sample is shown. The decay channels that correspond to the enhanced feature outputs are shown as shaded regions.

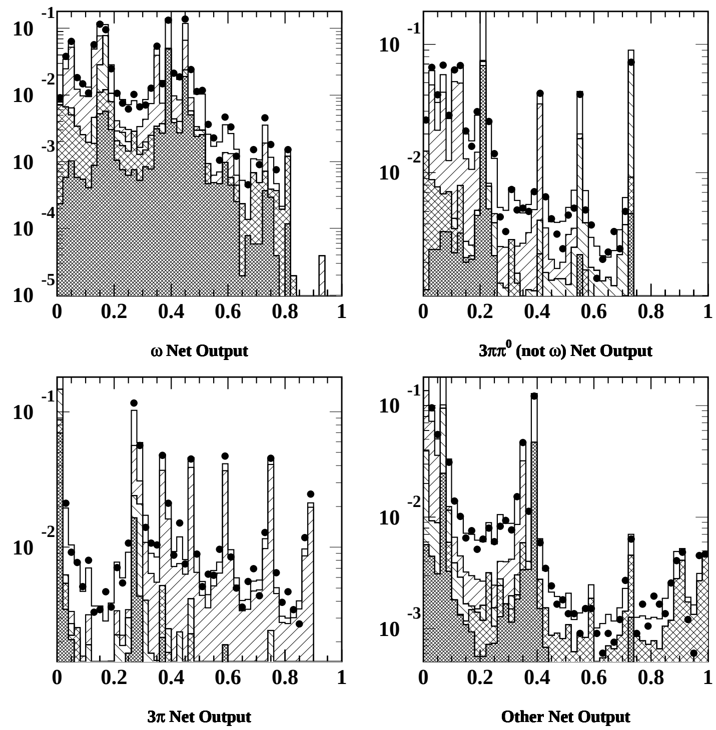

Figure 7.7: The 1994 data agrees with the Monte Carlo net output. The breakdown of individual decay channels is shown. The shading key is identical to figure 6.4

Chapter 8

Evaluation of the Branching Ratio

In the previous chapter, different methods for determining the proportion of different decay classes in a sample are discussed. Several estimates for the proportion of $\tau \rightarrow \omega\pi\nu_\tau$ in the data passing selection have been determined. These values need to be transformed into an analysis independent form, a branching ratio. A branching ratio is a measure of the percentage of all τ's produced in e^+e^- collisions that will decay via an ω, $Br(\tau \rightarrow \omega\pi\nu_\tau)$.

8.1 Determination of the Branching Ratio

The fractions gained from the fitting methods must be traced back to the original sample. At every step of selection, some events are lost, leaving the events that look most like ω decays. The number of lost events has been estimated by using the Monte Carlo in Chapter 6. The loss from tau platform and analysis selections are combined in table 8.1. The efficiency for the low multiplicity selection was not modelled by the MC, but it is assumed that the efficiency is high and that any inefficiency will affect both the ω and non-ω tau decays reasonably equally. This is backed up by the near perfect tagging of $e^+e^- \rightarrow \tau^+\tau^-$ decays with visible energy [45]. The final event

Selection Step	*ω Decays*	*All τ Decays*
Initial MC Size	400,000	1,200,000
% Passing Selection	33.02 ± 0.20	4.28 ± 0.02
% ω in 1994 Data From: $\chi^2_{M_{inv}(\omega)}$	12.6 ± 2.5	
Kinematic Max Likelihood (uncorrected)	13.4 ± 1.1	
Kinematic Max Likelihood (corrected)	16.7 ± 1.3	
Neural Net Max Likelihood	15.6 ± 2.5	
Linear Discriminant (optimal weights)	17.1 ± 3.2	

Table 8.1: The results from the four, reasonably independent methods of determining the proportions of $\tau \to \omega\pi\nu_\tau$ decays in the analysis sample are shown. All methods are in reasonable agreement. The error on the uncorrected maximum likelihood fit is artificially small due to its ignorance of correlations.

sample is considered to be a pure sample of taus as demonstrated by the comparison of selection effects on both samples and the $\sim 99\%$ purity of the tau sample after the tau platform "loose" selection (Chapter 6). There are 6067 in the 1994 data after selection, N_{sample}. The merits of each method of determining percentage of ω candidates in this sample will be discussed in the next section.

The estimate of the branching ratio is the number of $\tau \to \omega\pi\nu_\tau$ candidate decays over the number of all τ candidates.

$$Br(\tau \to \omega\pi\nu_\tau) = \frac{N_\omega}{N_\tau} \tag{8.1}$$

where

$$\begin{aligned} N_\omega &= \frac{\%\omega_{in\ sample} N_{sample}}{\%\omega_{passing\ selection}} \\ N_\tau &= \frac{N_{sample}}{\%\tau_{passing\ selection}}. \end{aligned} \tag{8.2}$$

The results for the five statistical methods are shown in table 8.2. The statistical

Fitting Method	*Branching Ratio*
$\chi^2_{M_{inv}(\omega)}$	1.63 ± 0.32
Kinematic Max Likelihood (uncorrected)	1.74 ± 0.14
Kinematic Max Likelihood (corrected)	2.16 ± 0.17
Neural Net Max Likelihood	2.02 ± 0.32
Linear Discriminant	2.22 ± 0.41

Table 8.2: The $Br(\tau \to \omega\pi\nu_\tau)$ estimates from the five fitting methods agree within error with each other and previous measurements of $Br(\tau \to \omega\pi\nu_\tau)$ (Chapter 3).

errors are significantly higher than other experiments (Chapter 3). This is due to the limited resolution of the electromagnetic calorimeter in identifying π^0's (Section ??). The spread of the measurements is due to the proper/improper treatment of features in the data.

8.2 Best Fitting Method

Now that there are 5 methods of determining the branching ratio, the most appropriate method needs to be determined. Each of the fitting techniques has its disadvantages. The skill is to choose the least flawed philosophy for this problem.

8.2.1 Linear Discrimination

In analyses such as this, there is a difficulty illustrated best by the invariant mass distribution of the ω candidates. Decay channels are delineated by the narrowness of resonances instead of large differences in distribution means. A Fisher discriminant is a measure of the distance between means. If a sharp peak is overlayed on the middle of a broad background, the discriminant is powerless. In this particular study, linear discrimination is a severely limited approach.

8.2.2 Minimised χ^2

The χ^2 measure has been calculated for each individual kinematic variable. This reduces the amount of information available for the fit. Combining the measures of the variables is impractical, since the behaviour of correlations after the fit is not available. Something must be done to make full use of the kinematic information and take multidimensional variance into account.

8.2.3 Uncorrected Maximum Likelihood

While a basic log-likelihood fit uses all the information available, it is designed to work on perfectly uncorrelated variable sets. The correlation biases the fit unpredictably, driving down highly correlated probabilities and driving up other parameters to compensate. The error is also artificially reduced due to effectively double counting highly correlated probabilities. Since all the variables generated are inherently correlated at some level, a correction must be made.

8.2.4 Corrected Maximum Likelihood

With the correction projecting single dimensions into Gaussian distributions, a sufficient covariance matrix is generated to account for correlations between variables. While this correction is a marked improvement, it is not a perfect fix. The 1-dimensional projections are Gaussian, but the distributions do not transform into an N-dimensional Gaussian. Irregardless, the fit and errors should be more accurate.

8.2.5 Neural Net Inputs

A way of sidestepping the issue of correlations is to provide the maximum likelihood fit with uncorrelated variables. Neural nets properly deal with correlations when training. While multidimensional outputs are perfectly correlated, each output can be used individually to provide information about one decay channel at a time. There is also the added bonus of enhancing and combining distinguishing features from the kinematic variables. This removes most of the uninformative information that might weigh down a log-likelihood fit during maximisation.

8.2.6 The Verdict

The chosen method of determining the branching ratio must take as much information into account and properly deal with the significant correlations between variables. Fisher discriminants, the χ^2 fit, and the uncorrected maximum likelihood fit are all immediately eliminated. While the corrected maximum likelihood fit on kinematic variables provides an improvement, the fit can still be bogged down with extraneous information. It is best to start with optimal input variables for the log likelihood fit, especially when decay distributions overlap significantly in variable space. From the available fitting methods, a maximum likelihood fit with neural net feature inputs that correctly deals with correlations is statistically the most appropriate method.

Chapter 9

Results and Conclusions

9.1 The Branching Ratio

With five distinct statistical fitting methods for estimating a branching ratio investigated, a correlation corrected maximum likelihood fit with neural net feature enhancing inputs provides the most appropriate algorithm for the 1994 $\tau \rightarrow \omega\pi\nu_\tau$ decay candidates at OPAL. The final estimate for $Br(\tau^- \rightarrow \omega\pi^-\nu_\tau)$ is 2.02 ± 0.32. The error is statistical.

9.1.1 Systematic Errors

Systematic errors provide an approximation of errors introduced by the experimental method that are biased/asymmetric. These errors remain constant, no matter how many times a measurement is taken, unlike random, statistical errors. In contrast to a cross section measurement, calculating a branching ratio is not greatly affected by most systematic errors. In this case, it is a ratio of $\omega\pi$ candidates divided by all other τ decays. The number of $\omega\pi$ candidates is determined by studying a subsection of τ decays that mimic the $\omega\pi$ channel. Detector acceptance and measurement biases should affect events in the subsample similarly. The majority of the systematics will disappear once a fraction is taken.

The most significant systematic concern that does not cancel in the branching ratio is the modelling of the ω in the Monte Carlo. This directly affects the fitting techniques which rely heavily on the accuracy of theory. If the width of the ω resonance was poorly measured or the inclusion of interference between the ρ and the ω significantly changed the $3\pi\pi^0$ distribution, systematic studies would become vital. Fortunately, the width is known to 1% and is narrow, limiting the effects of interference terms. An estimate of this modelling systematic can be achieved by making a hard cut on the Monte Carlo around 1% of the $\pi^+\pi^-\pi^0$ ω mass peak. 92% of the pure ω sample and 3% of the non-ω sample fall within this window. While there are twice as many non-ω events on average in the selected tau decays, the cut would provide $\sim$ 98% of the true ω sample for fitting. With the events from both classes being virtually identical in most kinematic distrubutions, this would drop the number of estimated ω's by $\sim$ 3%. Compared to the statistical error of $\pm$16%, this is minimal.

While this systematic error in itself does not require addition to the statistical error to maintain the number's accuracy, smaller systematics not mentioned may push the error from 0.32% to 0.33% or higher. A further systematic study would be necessary to determine their significance. The suspicion is that it would be minimal. This is compounded by the fact that the above measurement has the same statistical error as the 1986 ARGUS analysis (table 3.1) and is dramatically larger than any consecutive measurement. This is due to inadequate identification of the π^0 candidates. The result is statistically not competitive.

9.2 Potential for Second-Class Current Limit

No limit on the existence of the second-class current $\tau \rightarrow b_1\nu_\tau$ can be made in this study. While the linear discriminant and neural net variables provide some measure of likelihood that an event is an omega, the accuracy of that measure is limited. Statistical fitting to identify the proportion of ω's in the decays is more accurate, but does not locate individual events. The ω signal is so broad across distributions such as the reconstructed mass of the ω candidate that no hard cut can be made that allows for any kind of ω sample purity. Without a pure sample, no limit can be determined.

9.3 Impact of Result

The impact of this latest measurement is not significant. The branching ratio fits are all statistically limited compared to previous measurements. While increasing the event sample beyond 1994 will reduce the statistical error slightly, the resolution of the electromagnetic calorimeter will hamper any improvement on the error.

There is a notable feature of this study, though. The branching ratio, although within error of all previous values is significantly higher. In fact, there has been a chronological trend of increase across all measurements. The suspicion is that improvements in modelling of the ω resonance would increase the measurement. A primitive Breit-Wigner would not accurately model the asymmetric tail in figure 5.1. It has also been shown that a more sophisticated fitting technique provides a more accurate measurement than a χ^2 single variable fit similar to the one performed by ALEPH. Even with its large error, this $Br(\tau \rightarrow \omega\pi\nu_\tau)$ measurement may be showing that the previous calculations underestimate the branching ratio.

9.4 Further Study

If it were possible to obtain a reasonably pure sample of $\omega\pi$ tau decays, it might be worthwhile for OPAL to have a stab at a limit to second-class currents. The standard method of adjustable windowing [32] to provide such a sample will not work due to the broadening of the ω resonance. ECAL resolution leads to poor π^0 reconstruction, even with the improved maximum entropy clustering algorithm. Both detector and statistical manipulation of the data have been strained to the limits, and still there is little chance that OPAL will ever be competitive in this region of τ physics.

9.4.1 Possible Hope for the Future

There could be an improvement in the study of $\tau \to \omega\pi\nu_\tau$ in the next generation of particle detectors. BaBar, with its extensive production of τ's, has the potential for generating a precision measurement of $Br(\tau^- \to \omega\pi^-\nu_\tau)$ and isolating a reasonably pure sample of ω's [65]. While the ECAL resolution is not dramatically better than OPAL's, the average energy of π^0's is greatly reduced, enabling more accurate π^0 reconstruction.

Appendix A

τ Pair Selection

The τ pair preselection performed by the tau platform has changed little from the initial work done in tau physics by OPAL [66]. Most of the alterations [53] have occurred in the endcap region where this analysis does not venture. The selection below should provide an unbiased and clean sample of τ pairs to analyse. These are the τ pair requirements for the **barrel only**. The cuts define the *tight* preselection with the *loose* preselection in parentheses [50].

- Multihadronic events are rejected by specifying charged track and ECAL cluster multiplicity requirements.
  - A good track has enough hits in the central jet chamber,

    $N_{hit}(CJ) \geq 20$

    some momentum transverse to the beam axis,

    $P_t \geq 0.1$ GeV

    and lies reasonably close to the interaction point.

    $|d_0| \leq 2$ cm: the impact parameter

$|z_0| \leq 75$ cm: the decay length

$R_{min} \leq 75$ cm: the distance of the first CJ hit's distance to the beam axis

– A good cluster in the barrel region ($cos\theta \leq 0.7$) hitting at least one block

$N_{blk} \geq 1$

with a significant amount of energy deposited in the ECAL.

$E_{clust}(raw) \geq 0.1$ GeV

– Once tracks and clusters are identified as *good*, the multihadrons can be removed by requiring the number of good charged tracks and ECAL clusters to be small,

$1 \leq N_{track} \leq 6 \qquad (1 \leq N_{track} \leq 10)$

$N_{clust} \leq 10 \qquad (N_{clust} \leq 15)$

- It is convenient to regard each τ as a jet in order to remove multihadronic and two-photon events which are characterised by wide-spread or acolinear event topologies [67].

  – Events are selected with two cones of significant ECAL cluster and track energy

  $N_{cone}(charge) = 2 \qquad (N_{cone\ min} = 0,\ N_{cone\ max} = 9999)$

  $P_{cone} \geq 0.01 E_{beam}$

  with a small enough half-angle for the cone containing the charged tracks and ECAL clusters.

  $\theta_{half\ cone} = 35^0$

- Cosmic rays are rejected with tighter track cuts and TOF requirements.
  - At least one track needs to have a very small impact parameter and average decay length.

    $|d_0|_{min} \leq 0.5$ cm (0.5)

    $|z_0|_{ave} \leq 20$ cm (20)

  - There are the added requirements that at least one TOF signal fires around the nominal expected value,

    $|t_{meas} - t_{exp}| \leq 10$ ns (10.0)

    and no pair of TOF counters, i and j, separated azimuthally by more than 165^0 has triggered too far apart.

    $|t_i - t_j| \leq 10$ ns (10.0)

- The last significant background to lepton pairs is from two-photon events. This is removed by acolinearity and visible energy cuts.
  - The track and cluster momentum sum of the two cones must have small acolinearity

    $\theta_{acol} \leq 15^0$ (999^0)

    while the highest calculated value of visible energy must be notable.

    $E_{vis} \equiv \sum_{cone} Max(E_{cls}, E_{trk}) \geq 0.03 E_{CM}$ (0.01)

  - Also, if the visible energy is still quite small,

    $E_{vis} \leq 0.20 E_{CM}$ (0.20)

    events with minimal transverse momentum are removed.

$$P_t(cls) \geq 2 \text{ GeV} \qquad (2.0)$$
$$P_t(trk) \geq 2 \text{ GeV} \qquad (2.0)$$

- Interesting physics is determined through FYZ1 preselection bit requirements.
  - For $e^+e^- \rightarrow l^+l^-$, the low multiplicity bit is set to *TRUE*.
- Once a reasonably clean sample of lepton pairs is selected, e^+e^- and $\mu^+\mu^-$ can then be separated from the desired $\tau^+\tau^-$ pairs.
  - $e^+e^- \rightarrow e^+e^-$ events have negligible missing energy, and dump most of their energy in the ECAL subdetector. A cut is made on deposited energy for an event.

    $$\sum_i E_i^{CLUST} \leq 0.8E_{CM}$$
    $$or \sum_i E_i^{CLUST} + \sum_i E_i^{TRACK} \leq E_{CM}$$
  - Once the electron pairs are removed, muon pair events are identified. The first step is to identify muon candidates. This is done in one of three ways. The number of triggered layers of muon chambers must be significant,

    $$N_{layers}^{MU} \geq 2$$

    Or the HCAL has been triggered often, especially in the last layers and in single layers,

    $$N_{layers}^{HC} \geq 4$$
    $$N_{last3}^{HC} \geq 1$$
    $$N_{hits/layer}^{HC} > 2$$

    Or there is very little energy deposited in the ECAL.

    $$E_{cls} < 2 \text{ GEV}$$

– Not all $\mu^+\mu^-$ events, once identified, are rejected due to the fact that both τ's in an event may decide to decay into muons. To remove most of the $e^+e^- \to \mu^+\mu^-$ background a total energy cut is made:

$$\sum_{cone=1}^{2} E_{cone}^{CLUST} + E_{cone}^{TRACK} \geq 0.6 E_{CM}$$

What should be left after the above cuts is a workably pure sample of tau pairs with which to conduct the analysis. The efficiency, ϵ, over a geometrical acceptance of $|\cos\theta| < 0.90$ is $\sim 88.8\%$ [67]

Appendix B

Maximum Entropy

> *It rained for the last few days. The weather report calls for sun tomorrow. I'll take my umbrella with me in the morning.*
>
> **- an advocate of Maximum Entropy**

Maximum Entropy was developed in the late seventies as a way of reconstructing images from astronomical data [68]. In particle physics, Maximum Entropy can be used to translate energy deposited in an array of electromagnetic calorimeters into the most likely combination of clusters from particle showers in the subdetector [69].

B.1 Bayesian Probability Calculus

The image processing algorithm of Maximum Entropy is based on *Bayesian* probability calculus in which

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \tag{B.1}$$

$P(A)$ is the probability of A occurring without any regard to event B. $P(B|A)$ is the conditional probability that given A, B will occur.

Bayesian statistical analysis builds on logical induction by representing all previous knowledge about a process as a prior probability function, $P(A)$. The conclusions drawn from this function are modified in light of any new data $P(B|A)$, maximising the chances that the prior probability models the system.

B.2 Application to Image Reconstruction

For neutral cluster identification, Bayes' theorem becomes

$$P(image|data) = \frac{P(image)P(data|image)}{P(data)} \tag{B.2}$$

where the *image* (I), the most likely image to make up the previously acquired data, defines the prior probability, and *data* (D) is the new data in the form of shower shapes and energy deposited in the electromagnetic detector. The probability of measuring new data is considered a constant, therefore

$$P(I|D) \propto P(D|I)P(I). \tag{B.3}$$

The object is to maximise $P(I|D)$, the probability of obtaining a specific image given the data will be greatest for the most likely outcome. Suitable models for $P(D|I)$ and $P(I)$ must be determined.

B.2.1 The Conditional Probability of the Data

With n data points for a given D, a Gaussian limit can be written as:

$$\begin{aligned} P(D|I) &= \prod_{i=1}^{n} (2\pi\sigma_i^2)^{-\frac{1}{2}} exp(-\frac{1}{2}\chi_i^2) \\ &\propto exp(-\frac{1}{2}\chi^2) \end{aligned} \tag{B.4}$$

Each of the showers can be considered a Gaussian with constant error, taking into account how probable the particle was to be in any position in the detector grid. χ^2 is calculated from the image I's transformation into data space, using known experimental resolution. In simpler terms, the observed data are given a measure as to how likely they produced the resultant I by comparing them to the expected form of the data from I. The closer χ^2 is to 1, the more the image and data agree.

B.2.2 Prior Probability

There are an infinite number of images consistent with any given data set due to the continuous errors in resolution. Therefore, some extra assumption has to be made to obtain the "best" image from this set. A prior probability, $P(I)$, is chosen for this purpose. It is the probability one would assign to an image *before* the data is observed, i.e. the background noise in the system.

$P(I)$ is taken as the combinatorial probability of obtaining a random sampling of N counts in a grid with n_i hits in pixel i.

$$P(I) = Prob(n_1, n_2, n_3 ...) = 2^{-N} \frac{N!}{n_1!n_2!n_3!...} \tag{B.5}$$

which can be rewritten using Stirling's approximation, $\log n! \approx n(\log n - 1)$.

$$P(I) = Prob(n_1, n_2, n_3 ...) \propto \exp(S) \tag{B.6}$$

where

$$S = -\sum_i n_i \log \frac{n_i}{N}. \tag{B.7}$$

This "entropy term", $P(I)$, is maximum for a flat distribution of equal numbers hits in each pixel, i.e. constant noise.

Taking the signals from the electromagnetic calorimeter to be an intensity map with intensity x_i in each pixel, S becomes

$$S = -\sum_i x_i \log \frac{x_i}{b} \tag{B.8}$$

with b as an arbitrary constant. It defines what fraction of intensity is considered a unique count.

The best image choice for data D assuming a flat background is found by maximising $P(I|D)$ which combines the entropy and conditional probability terms.

$$P(I|D) \propto exp(\alpha S - \frac{1}{2}\chi^2) \tag{B.9}$$

With $\alpha = 1$, the probability is Bayesian. α is included to allow for adjustment of the noise in the system during minimisation and is chosen such that the maximised result will have χ^2 per degree of freedom equal to 1.0. The χ^2 term on its own would find a solution which looks identical to the data ($\chi^2 = 0.0$). The "entropy" term would completely flatten out the image. Combining the two provides the *flattest* solution consistent with the data.

B.3 Reconstruction of Clusters in the ECAL

So how does this apply to the ECAL blocks? Figure B.1 shows the maximum entropy process in a 3×3 array of ECAL blocks. ECAL information is shown as boxes in the grid proportional to the energy deposited in the blocks at coordinates $\alpha\beta$ ($E_{\alpha\beta}$). The maximum entropy image divides these blocks into a 10×10 array of pixels, each of intensity, x_{ij}, with ij being the position of the 1×1 pixels.

Appropriate methods of calculating the two terms in equation B.9 must be found. The entropy term, S, in the above proportionality takes into consideration the maximum entropy image's sub-pixel intensities, x_{ij}.

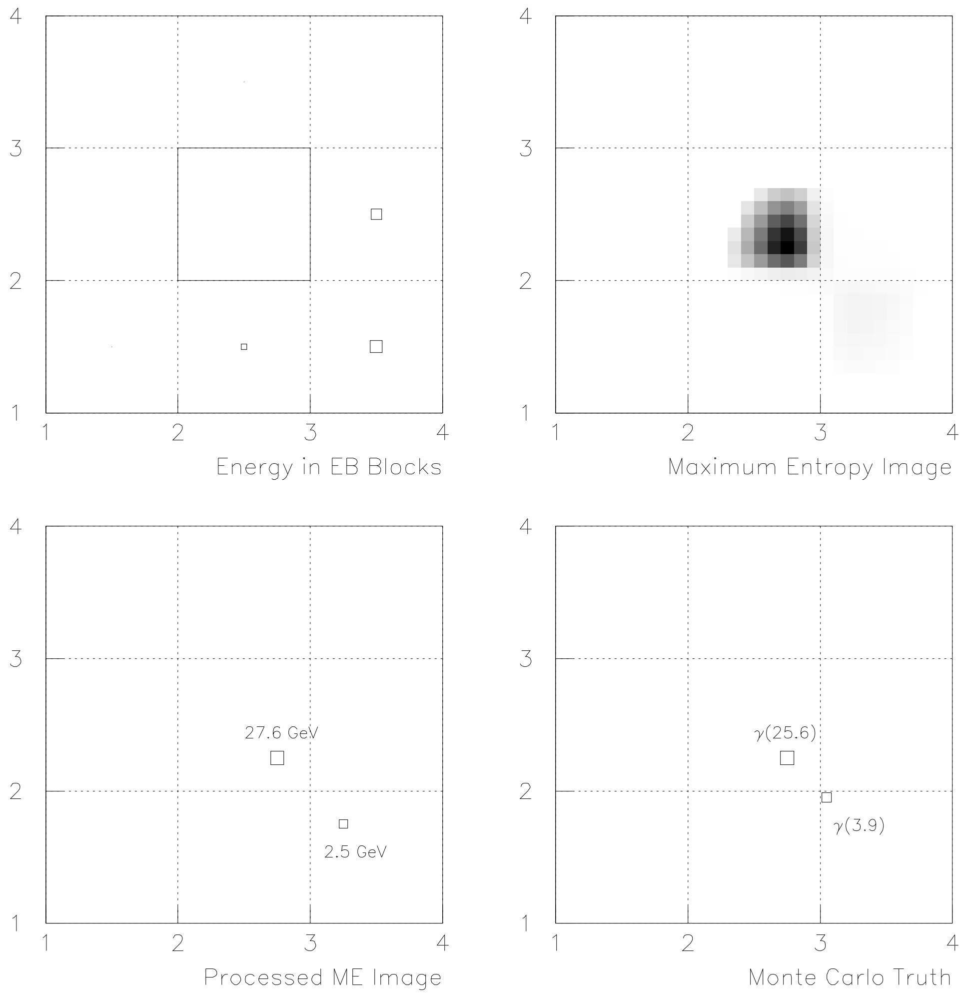

Figure B.1: An isolated π^0 that has decayed into two photons. Additional energy in the blocks surrounding the main deposit has forced the maximum entropy clustering algorithm to identify two clear maxima. A reasonable description of the truth is achieved by the image.

For χ^2 a transformation matrix, $T_{(\alpha\beta)(ij)}$, is needed to map image space, x_{ij}, onto data space, $D_{\alpha\beta}$. This describes how a specific image corresponds to the data. It is determined in groups of 3×3 blocks as a function of the central block, the block with the most ECAL energy of a cluster. Most of the information about the position of the shower is found in the surrounding blocks that contain the tails of the distribution. The shower profile is characterised by the Molière radius (2.94 cm in OPAL's ECAL), and the error on the energy is assumed to be

$$\sigma_E = 0.05\sqrt{E/GeV} + 0.01GeV \tag{B.10}$$

A detailed account of the energy sharing parameterisation that defines T is given in [70].

The equation to maximise becomes

$$\alpha S - \frac{1}{2}\chi^2 = -\alpha \sum_{ij} x_{ij} \log x_{ij} - \frac{1}{2} \frac{\sum_{\alpha\beta}(E_{\alpha\beta} - \sum_{ij} T_{(\alpha\beta)(ij)} x_{ij})^2}{\sigma^2_{\alpha\beta}} \tag{B.11}$$

with

$$\chi^2 = \frac{\sum_{\alpha\beta}(E_{\alpha\beta} - \sum_{ij} T_{(\alpha\beta)(ij)} x_{ij})^2}{\sigma^2_{\alpha\beta}}. \tag{B.12}$$

The maximisation algorithm to find the most appropriate x_{ij}'s can be found in [71]. The adjustment of α to accommodate $\chi^2 = 1.0$ for a set of x_{ij}'s produces an estimated event reconstruction, where particles are assigned to local maxima and given corresponding energies. This is shown to compare quite reasonably with the true Monte Carlo energies and position as in figure B.1.

B.4 An Alternative Method

A mention should be made of an alternative clustering algorithm [72]. The tau platform contains a second photon finder which does not take a discrete, statistical approach. Instead of subdividing the ECAL blocks into pixels and calculating a transformation matrix, it fits shower distributions directly to ECAL energies, minimising the χ^2. Unlike the above method, this one has been extended into the endcap. It is still in a developmental phase, although preliminary studies do not show a marked improvement to the maximum entropy method.

Appendix C

Linear Discrimination

An alternative method of separating decay distributions, Fisher or linear discrimination [73], has been studied. A boundary is determined to separate a population into two or more subgroups from a vector of individual event measurements.

C.1 Discrimination

Discrimination is the process of defining a rule built on previous knowledge to allocate new members into the correct subpopulation of a group. Each event is considered a point in N-dimensional space, S, with coordinates $x = (x_1, x_2, x_3, ..., x_N)$. Two or more distinct, but overlapping, continuous densities can be distinguished if individual members of the populations are available from a Monte Carlo, etc. It is unnecessary to know the exact shape of the distribution.

C.1.1 Simple 2 Distribution Case

Consider two equally large, equally probable, equally weighted distributions, A and B. There is a region R where there are as many A events inside as there are B events outside (figure C.1).

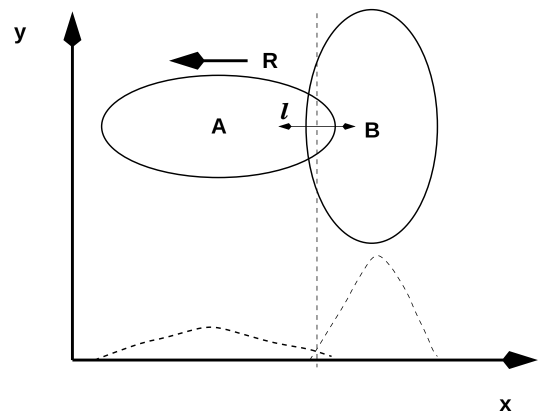

Figure C.1: For two distributions of equal size and importance, a vector l can be identified that connects the distribution centres. This is the discriminant, and orthogonal vectors can be assigned to provide subsamples of varying purity and selection efficiency. the most popular cut on the discriminant provides equal but opposite amounts of subsample on either side. In this two dimensional case, region R is defined by a straight line.

$$\int_R f_B dx = \int_{S-R} f_A dx = 1 - \int_R f_A dx \tag{C.1}$$

or, in other words,

$$\int_R (f_A + f_B) dx = 1. \tag{C.2}$$

f_i is the *relative frequency* that decay i is in that region of space. In laymen's terms, the chance of finding decay type A inside the region R is the same as finding decay type B outside the region, and vice versa. Because of this condition, the chance of misidentifying a member from either population is the same. As in most methods of distribution separation, the object is minimisation, in this case, the chance of misassigning an event to R. It can be done by minimising the error

$$1 + \int_R (f_B - f_A)dx \tag{C.3}$$

which is the sum of the fractions of the two distributions that have been misidentified. For the smallest error, $f_B - f_A$ should always be negative everywhere in R and positive outside R. The border is where the frequencies overlap.

$$f_A = f_B \tag{C.4}$$

As in the likelihood selection, the minimisation of the error on a function of f_A and f_B is equivalent to minimising logs, etc.

$$\log f_A = \log f_B \tag{C.5}$$

This becomes useful when dealing with unnormalised subpopulations.

Fractions of the total population for the clusters could be unequal ($\pi_B = 1 - \pi_A$ where $\pi_A \neq \pi_B$), as is the case for this analysis due to differences in branching ratios. The significance of misidentifying a member of the population (c_A and c_B) could be much different in the two populations. Misdiagnosing a person with a serious heart defect is much more serious than tagging a perfectly healthy person with the possibility of heart failure. With these added considerations, the error of misidentification becomes

$$c_B \int_R \pi_B f_B dx + c_A \int_{S-R} \pi_A f_A dx = c_A \pi_A + \int_R (c_B \pi_B f_B - c_A \pi_A f_A) dx. \tag{C.6}$$

This is minimised when

$$\frac{c_B \pi_B f_B}{c_A \pi_A f_A} = 1. \tag{C.7}$$

Calling once more on our old friend the logarithm, this simply introduces a constant term into the function.

$$\log \frac{f_B}{f_A} + \log \frac{c_B \pi_B}{c_A \pi_A} = \log \frac{f_B}{f_A} + k = 0 \tag{C.8}$$

$\log \frac{f_B}{f_A}$ is considered the discriminator, and equation C.5 is the discriminating function.

C.1.2 The Linear Case

For the simplest case, take a population consisting of two normal Gaussians with a common covariance matrix γ, the boundary is determined by

$$\frac{1}{2}(x - \mu_A)'\gamma^{-1}(x - \mu_A) = \frac{1}{2}(x - \mu_B)'\gamma^{-1}(x - \mu_B) \tag{C.9}$$

from equation C.5. When multiplied through, this simplifies to

$$x'\gamma^{-1}(\mu_A - \mu_B) - \frac{1}{2}(\mu_A + \mu_B)'\gamma^{-1}(\mu_A - \mu_B) = 0 \tag{C.10}$$

Just as with the maximum likelihood fitting method, the decay samples can be pushed into Gaussian space if they are not inherently Gaussian. A pooled unbiased estimate of a common covariance matrix can be a achieved for two samples with covariance matrices c_A and c_B and population sizes n_A and n_B.

$$c = \frac{n_A c_A + n_B c_B}{n_A + n_B - 2} \tag{C.11}$$

By determining the means as $\bar{x}_A$ and $\bar{x}_B$, the boundary is

$$x'c^{-1}(\bar{x}_A - \bar{x}_B) - \frac{1}{2}(\bar{x}_A + \bar{x}_B)'c^{-1}(\bar{x}_A - \bar{x}_B) = 0 \tag{C.12}$$

This can be proven by calculating the left side of the equation for $x = \bar{x}_A$ and $x = \bar{x}_B$. The boundary value of 0 lies halfway between the values for the two means.

Defining l as the vector of coefficients for the linear discriminant

$$l = c^{-1}(\bar{x}_A - \bar{x}_B) \tag{C.13}$$

equation C.12 becomes

$$l'x = \frac{1}{2}l'(\bar{x}_A + \bar{x}_B). \tag{C.14}$$

In equation C.8, the constant k is added to the right-hand side.

Normalisation of the distribution is not necessary. From the approach of minimising between-class and within-class variances, the proportionality

$$l \propto \gamma^{-1}(\mu_A - \mu_B) \tag{C.15}$$

can be derived without it [74].

C.2 Multiple Distributions

With many distributions, there is still only the possibility of comparing two sources at one time. Each pair has its own discriminant for distinction. Combining the information from the discriminants, identifies the most likely distribution for a given event.

Bayesian statistics steps in once again to define the probability that an event x belongs to distribution A given the percentage of A in the sample, p_A.

$$P(A|x) = \frac{p_A P(x|A)}{p_A P(x|A) + p_B P(x|B) + p_C P(x|C)...} \tag{C.16}$$

$P(x|A)$ is the Gaussian function describing distribution A. The most likely source will have the largest $p_j P(x|j)$. Once the log is taken to determine the discriminator, both sides of equation C.9 take the form

$$\mu_j' V^{-1} x - \frac{1}{2}\mu_j' V^{-1} \mu_j + \frac{1}{2}\log p_j \tag{C.17}$$

The discriminants for the multiple samples are

$$F_j = \mu_j' V^{-1} \mu_j + w_j \tag{C.18}$$

where w_j is adjustable and contains the constant terms.

C.3 Application of Linear Discriminants

Once a discriminant is chosen, it can be used on its own or in conjunction with a fitting method. The following caclulations were performed by Roger Barlow to complement the previous analysis.

C.3.1 Counting Techniques

The adjustable weights are all set to 0, initially for the four subsamples in the ω candidate distribution. This effectively ignores any information of the relative proportions in the sample and tests for the efficiency of the discriminator. The program will ultimately misassign decay channels. This can be corrected for by using the Monte Carlo to determine the estimated misassignment rate.

Adjusting the weights changes the statistical errors as well as the fractional estimate. To properly account for the proportions of decay channels in the sample, the weights are adjusted to minimise the errors. Table C.1 details the results of both methods.

Decay Channel		Highest F	MC Corrected	*% in Sample*
1994 Data	ω	1485 ± 39	980 ± 148	15.4 ± 2.3
($w_j = 0$)	$3\pi\pi^0$ (not ω)	1059 ± 33	1273 ± 108	20.0 ± 1.7
	3π ($a1$)	2894 ± 54	3281 ± 88	51.5 ± 1.4
	everything else	938 ± 31	843 ± 115	13.2 ± 1.8
1994 Data :	ω	2163 ± 47	1022 ± 114	16.0 ± 1.8
(w_j's adjusted)	$3\pi\pi^0$ (not ω)	832 ± 29	1223 ± 104	19.2 ± 1.6
	3π ($a1$)	3068 ± 55	3286 ± 87	51.5 ± 1.4
	everything else	313 ± 18	884 ± 95	13.2 ± 1.5

Table C.1: While the errors on the ω fraction is noticeably driven down by adjusting the probability weights, there is little alteration in the estimated makeup of the analysis sample.

C.3.2 Fitting Techniques

The counting techniques amount to a hard cut on the discriminant. Fitting techniques do not deal in absolute edges, and therefore have better statistical precision. The outputs for the four discriminants F_j are treated like the neural net outputs and passed through a corrected maximum likelihood fit. This is checked against an alternative fitting method, the method of optimal weights [75], which is not dependent on the Gaussian nature of the probability density functions. The method of optimal weights should be as efficient as a maximum likelihood method provided the a priori weights are treated properly. Results for both these methods are in table C.2.

C.4 An Inappropriate Choice

While a linear discriminant method provides comparable results with the other fitting methods, it is ultimately an inappropriate method of dealing with the ω candidate sample distributions. A Fisher discriminant identifies the path between source means in multidimensional space, disregarding standard deviations. From many projections

Decay Channel		*% in Sample*
1994 Data (ML fit):	ω	19.4 ± 2.0
	$3\pi\pi^0$ (not ω)	12.9 ± 1.7
	3π ($a1$)	56.6 ± 1.4
	everything else	11.1 ± 3.0
1994 Data (optimal weights):	ω	17.1 ± 3.2
	$3\pi\pi^0$ (not ω)	18.7 ± 2.8
	3π ($a1$)	50.7 ± 1.5
	everything else	13.6 ± 2.0

Table C.2: The fitting methods should provide a more accurate method of determining the fractional makeup of the analysis sample. The % of $\tau \rightarrow \omega\pi\nu_\tau$ in the sample is markedly higher than from the counting methods.

of the MC sample (see figure 6.7), the means of at least 2 distributions are close, while the spreads of the subsamples are completely different. A linear discriminant has no hope of identifying the distance between two means if there is no distance. The dramatic overlap of decay channel distributions renders linear discrimination impotent.

Bibliography

[1] B. R. Martin and G. Shaw, *Particle Physics*, John Wiley & Sons (1992) 24.

[2] S. Glashow, *Partial-Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579-588.

[3] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley and Sons (1984).

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264-1267. ; A. Salam, *Selected Papers of Abdus Salam*, ed. A. Ali *et al.*, World Scientific (1994) 244-254.

[5] S.L. Glashow, I. Iliopoulos, L. Maiani, *Weak Interactions With Lepton-Hadron Symmetry*, Phys. Rev. **D 2** (1970) 1285-1287.

[6] N. Cabbibo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531-533; M. Kobayashi, K. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interactions*, Prog. Theor. Phys. **49** (1973) 652-657.

[7] C. Bouchiat, I. Iliopoulos, Ph. Meyer, *An Anomaly-Free Version of Weinberg's Model*, Phys. Lett. **38 B** (1972) 516-523.

[8] S.L. Glashow and M. Gell-Mann, *Gauge Theories of Vector Particles*, Annals Phys. **15** (1961) 437-460.

[9] B. C. Barish and R. Stroynowski, *The Physics of the τ Lepton*, Physics Reports **157, No. 1** (1988) 1-62.

[10] Albrecht, *et al.*, (ARGUS Collaboration), *Determination of the Michel Parameter in τ Decay*, Phys. Lett. B **246** (1990) 278-284.

[11] D. H. Perkins, *Introduction to High Energy Physics*, 3rd edition, Addison-Wesley Publishing Co. (1987) 155, 206-258.

[12] M. Davier, *The Tau Lepton Current: Lifetime, Leptonic Branching Ratios and Michel Parameters*, Nucl. Phys. Proc. Suppl. **40** (1995) 395-411.

[13] R. P. Feynman, *An Operator Calculus Having Applications to Quantum Electrodynamics*, Phys. Rev. **84** (1951) 108-128.

[14] Particle Data Group, R.M. Barnett *et al.*, *Review of Particle Properties*, Phys. Rev. **D 54** (1996) 1233.

[15] S. Jadach, J. H. Kühn and Z. Wąs, *TAUOLA - A Library of Monte Carlo Programs to Simulate Decays of Polarized τ Leptons*, Comp. Phys. Comm. **64** (1991)) 275-295; S. Jadach, M Jeżabek, J. H. Kühn and Z. Wąs, *The τ Decay Library TAUOLA, Update With Exact $\mathcal{O}(\alpha)$ QED Corrections in $\tau \to \mu(e)\nu\bar{\nu}$ Decay Modes*, Comp. Phys. Comm. **70** (1992) 69-76.

[16] S. Jadach, J. H. Kühn and Z. Wąs, *The τ Decay Library TAUOLA, Version 2.4*, Comp. Phys. Comm. **76** (1993) 361.

[17] L. B. Okun, *Particle Physics: The Quest for the Substance of Substance*, Harwood Academic Publishers (1985) 163.

[18] R. Decker, *On the $\tau \to \nu_\tau \pi\omega$ Decay Mode*, Z. Phys. **C36** (1987) 487-500.

[19] D. H. Perkins, *Introduction to High Energy Physics*, 1st edition, Addison-Wesley Publishing Co. (1972) 113-116.

[20] S. Weinberg, *Charge Symmetry of Weak Interactions*, Phys. Rev. **112** (1958) 1375-1379.

[21] I. S. Hughes, *Elementary Particles*, Cambridge University Press (1988) 203-215.

[22] R. J. Oakes and J. J. Sakurai, *Spectral-Function Sum Rules, ω-phi Mixing, and Lepton-Pair Decays of Vector Mesons*, Phys. Rev. Lett. **19** (1967) 1266-1270.

[23] E. Commins, *Weak Interactions*, McGraw-Hill, Inc. (1973) 10-13.

[24] L. B. Okun, *Leptons and Quarks*, North-Holland Publishing Co. (1982) 30-35.

[25] J. J. Sakurai, *Invariance Principles and Elementary Particles*, Princeton University Press (1964) 111-135.

[26] A. Halprin, B. W. Lee, and P. Sorba, *Second-Class Currents and Very Light Quarks*, Phys. Rev. D **14** (1976) 2343-2348.

[27] C. Leroy and J. Pestieau, *τ-decay and Second-Class Currents*, Phys. Lett. **72B** (1978) 398-399.

[28] S.I. Eidelman and V.N. Ivanchenko, *τ Decays and CVC*, Nucl. Phys. Proc. Suppl. **40** (1995) 131-138.

[29] H. Albrecht *et al.* (ARGUS Collaboration), *Evidence for the Decay $\tau^- \to \nu_\tau \omega \pi^-$*, Phys. Lett. B **185**, (1987) 223.

[30] P. Baringer, *et al.* (the CLEO Collaboration), *Production of η and ω Mesons in τ Decay and a Search for Second-Class Currents*, Physical Review Letters, Vol. 59, No. 18 (1987) 1993-1995.

[31] R. Balest, *et al.* (the CLEO Collaboration), *Measurements of the Decays $\tau \to h^- h^+ h^- \nu_\tau$ and $\tau \to h^- h^+ h^- \pi^0 \nu_\tau$*, Phys. Rev. Lett. **75** (1995) 3809-3813.

[32] P. Bourbon, *Study of the $\tau^\pm \to 3\pi\pi^0\nu_\tau$ Decay in Aleph*, Nucl. Phys. B (Proc. Suppl.) **40** (1995) 203-212.

[33] D. Abbaneo, *et al.* (ALEPH Collaboration), *A Study of τ Decays Involving η and ω Mesons*, Z. Phys. **C74** (1997) 263-273.

[34] S. Meyers,*The LEP Collider, from design to approval and commissioning*, CERN 91-08. Available at **http://www.cern.ch/CERN/Divisions/SL/history/lep_doc.html#3.1**.

[35] W. Schnell(CERN), *Design Study of a Large Electron-Positron Colliding Beam Machine: LEP*, IEEE Trans.Nucl.Sci.26:3130, (1979) 3130-3134.

[36] R. Akers, et al., *The OPAL Muon Barrel Detector*, Nucl. Inst. and Meth. A 357 (1995) 253-273.

[37] S. L. Lloyd, *The OPAL Primer*, Version 97a, (1997) 30.

[38] R. Assmann, et al. (The working group on LEP energy), *Calibration of Centre-of-Mass Energies at LEP1 for Precise Measurements of Z Properties*, CERN-EP/98-040 (1998) 25.

[39] K. Ahmet et al., OPAL Collaboration, *The OPAL Detector Paper at LEP*, Nucl. Inst. and Meth. **A 305** (1991) 275.

[40] P. P. Allport, et al., *The OPAL Silicon Microvertex Detector*, Nucl. Inst. and Meth. A 324 (1992) 34-52.

[41] B. E. Anderson, et al., *The OPAL Silicon-Tungsten Calorimeter Front End Electronics*, IEEE Trans.Nucl.Sci., Vol. 41, No. 4, (1994) 845-852.

[42] S. Jadach, B. F. L. Ward and Z. Wąs, *The Monte Carlo Program KORALZ, Version 4.0, for the Lepton or Quark Pair Production at LEP/SLC Energies* Comp. Phys. Comm. **79** (1994) 503-522.

[43] D. R. Ward, *The OPAL Monte Carlo Program: GOPAL*, Proceedings, MC91: Detector and event simulation in high energy physics, Amsterdam (1991) 204-226; GOPAL133 Manual available at **http://www.cern.ch/opal/manuals/go/pro/go.html**

[44] R. Brun *et al.*, *GEANT Users Guide(Version 3)*, CERN EE EE 84/1, (1987). Available at **http://wwwinfo.cern.ch/asdoc/geant_html3/geantall.html**.

[45] M. Arignon, et al., *The Trigger System of the OPAL Experiment at LEP*, CERN-PPE/91-32 (1991).

[46] C. Hawkes et al., *ROPE408: Users Guide*, OPAL internal document (1994).

[47] A. Buijs, *The OPAL DST(OD) processor in ROPE*, Version 4.13, (1994).

[48] Nick Oldershaw, private communication.

[49] The OPAL Collaboration, *A Preliminary Update of the Z^0 Line Shape and Lepton Asymmetry Measurements with a Revised 1993-1994 LEP Energy and 1995 Lepton Asymmetry*, OPAL physics Note PN242, 24 July 1996.

[50] M. Sasaki et al., *The Tau Platform: Version 1.02*, OPAL internal document (Feb. 11 1993).

[51] A. Stahl and A. Posthaus, *The Tau Platform: Version 208*, OPAL internal document (May 12, 1997).

[52] D. A. Rigby and J. A. Wilson, *Measurement of the Muonic Branching Ratio of the τ*, OPAL Technical Note TN431, 18th October 1996, 21.

[53] P. Acton, et al., OPAL Collaboration, *Measurement of the Topological Branching Ratios at LEP* Physics Lett. **B 288** (1992) 6, 373-385.

[54] J. H. Kühn, *Tau Kinematics from Impact Parameters*, Physics Letters **B 313** (1993) 458-460.

[55] R. Barlow, *Statistics*, John Wiley and Sons (1989).

[56] Louis Lyons, *Statistics for Nuclear and Particle Physicists*, Cambridge University Press (1986) 85-102.

[57] F. James *et al.*, *MINUIT Minimization Package (Version 94.1)* CERN Program Library Long Writeup D506 (1994). Available at **http://wwwinfo.cern.ch/asdoc/WWW/minuit/minmain/minmain.html**

[58] J. A. Nelder and R. Mead, *A Simplex Method for Function Minimization*, Comput. J. **7**, (1965) 308.

[59] Dean Karlen, *An Improved Likelihood Selection Technique for Correlated Variables*, OPAL Technical Note TN459, 19 February 1997.

[60] W. H. Press, *et al.*, *Numerical Recipes*, Cambridge University Press (1988) 147-165.

[61] C. Peterson and T. Rögnvaldsson, *An Introduction to Artificial Neural Networks*, lectures given at the 1991 Cern School of Computing, Sweden, (August, 1991).

[62] C. Peterson, T.Rögnvalsson, and L. Lönblad, *JETNET 3.0: A Versitile Artificial Neural Network Package*, CERN-TH.7135/94 (1993).

[63] G. Bahan and R. Barlow, *Identification of b Jets Using Neural Networks*, Computer Physics Communications **74** (1993) 199-216.

[64] I. Aleksander and H. Morton, *Neural Computing*, Chapman and Hall (1990).

[65] R. Barlow, private communication.

[66] G. Alexander, et al., OPAL Collaboration, *Measurement of Branching Ratios and τ Polarisation from $\tau \to \pi(K)\nu$ Decays at LEP* Physics Lett. **B 266** (1991) 201-217.

[67] G. Alexander *et al.*, OPAL Collaboration, *Measurement of the Z^0 Line Shape Parameters and the Electroweak Couplings of Charged Leptons* CERN-PPE/91-67 (1991) 20-35.

[68] G.J. Daniell and S.F. Gull, *Image Reconstruction From Incomplete and Noisy Data*, NATURE, vol 272 (1978) 686-690.

[69] M. A. Thomson, *Maximum Entropy and Event Reconstruction*, OPAL Technical Note TN240, 6th July 1994.

[70] M.G. Vincter, *et al.*, *A study of Electromagnetic Shower Profiles in EB*, OPAL Technical Note 174, (1993).

[71] J. Skilling and R. K. Bryan *Maximum Entropy Image Reconstruction: General Algorithm* Monthly notices of the Royal Astronomical Society, **211**, (1984) 111-124.

[72] Sven Menke, private communication.

[73] M. G. Kendall, *The Advanced Theory of Statistics*, Vol. 3, Macmillan Publishing Co., (1983), 370-403.

[74] G. Cowan, *Statistical Data Analysis*, Clarendon Press (1998) 153-187.

[75] R. Barlow, *Event Classification Using Weighting Methods*, Journal of Computational Physics **72** (1987) 202-219.

Acknowledgments

I would like to acknowledge:

- The Marshall Aid Commemoration Commission (especially Catherine), PPARC, and the keepers of the Manchester High Energy Physics travel budget for their generosity in making this research financially viable.
- Roger for being the best darned supervisor a girl could ever wish for.
- Hal and everyone else who helped when Roger was too far away to bother.
- My parents for lovingly tolerating my decision to travel halfway across the world for no better reason than to write this tome.
- Potential Doctors of Philosophy Curtis and Jones for proving time and time again that my PhD experience could have been infinitely more horrifying.
- And finally, all those quaint little English people and everyone else over the past few years who, in the tradition of the Chinese curse, made my life "interesting".

thanks

- BETH