

**A SEARCH FOR LEPTON FLAVOR VIOLATION IN
UPSILON DECAYS**

by

William Oliver Love

BA in Physics and Mathematics, Rhode Island College, 2000

MS in Physics, University of Pittsburgh, 2002

Submitted to the Graduate Faculty of
the College of Arts and Sciences in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH
COLLEGE OF ARTS AND SCIENCES

This dissertation was presented

by

William Oliver Love

It was defended on

March 5, 2010

and approved by

Vladimir Savinov, PhD, Associate Professor

Vittorio Paolone, PhD, Associate Professor

Adam Leibovich, PhD, Associate Professor

W. Vincent Liu, PhD, Associate Professor

Peter Brusilovsky, PhD, Associate Professor

Dissertation Director: Vladimir Savinov, PhD, Associate Professor

Copyright © by William Oliver Love
2010

A SEARCH FOR LEPTON FLAVOR VIOLATION IN UPSILON DECAYS

William Oliver Love, PhD

University of Pittsburgh, 2010

This document describes a search for lepton flavor violation (LFV) in the bottomonium system. Using the data collected with the CLEO III detector, we looked for the leptonic decays $\Upsilon(nS) \rightarrow \mu\tau$ ($n = 1, 2,$ and 3). The τ lepton was identified through its leptonic decay $\tau \rightarrow e\bar{\nu}_e\nu_\tau$, and signal yields were obtained through multidimensional likelihood fitting with probability density function shapes measured from independent data samples. No signal was observed. Therefore, we report our results as upper limits on the LFV branching fractions of Υ mesons at the 95% confidence level.

TABLE OF CONTENTS

PREFACE	ix
I. ELEMENTARY PARTICLES	1
II. SYMMETRY	6
III. LEPTON FLAVOR VIOLATION	10
IV. PRIOR WORK	15
V. THE CORNELL ELECTRON STORAGE RING	18
VI. THE CLEO DETECTOR	23
VII. PARTICLE DETECTION AND RECONSTRUCTION	29
VIII. THE SAMPLES	33
IX. SELECTION CRITERIA	36
X. THE METHOD OF MAXIMUM LIKELIHOOD	44
XI. SYSTEMATIC UNCERTAINTIES	50
XII. THE INCLUSIVE ANALYSIS	55
XIII. CONCLUSIONS	60
APPENDIX. PLOTS	61
BIBLIOGRAPHY	87

LIST OF TABLES

1	The Fundamental Forces	3
2	The Known Elementary Particles	4
3	Composite Particles	5
4	Symmetries of the Standard Model	9
5	Upsilon Mesons	14
6	Prior Work	17
7	CLEO's Upgrades	28
8	Trigger Criteria	30
9	Data Acquisition Performance	31
10	CLEO III Data	32
11	Samples	34
12	Background Processes	39
13	Fit Variables	40
14	Selection Efficiencies: $\Upsilon(1S)$	41
15	Selection Efficiencies: $\Upsilon(2S)$	42
16	Selection Efficiencies: $\Upsilon(3S)$	43
17	Fit Components	45
18	Fit Results	49
19	Systematic Uncertainties on Detection Efficiency	51
20	Inclusive Selection Efficiencies	56

LIST OF FIGURES

1	Particle Exchange	2
2	Beta Decay	11
3	Lepton Flavor Violation	13
4	Upsilon Meson Decay	16
5	Neutrinoless Tau Lepton Decay	16
6	Triple Muon Decay	16
7	The Cornell Electron Storage Ring	20
8	CESR Schematic	22
9	The CLEO III Detector	27
10	Bhabha Scattering	35
11	Annihilation	35
12	Backgrounds in the Exclusive Analysis	37
13	Background Correlation for Radiative Muon Pairs	47
14	Tau Pairs Cross Section	53
15	Results	54
16	Cut Optimization	58
17	Backgrounds in the Inclusive Analysis	59
18	Projections	64
19	Tau Pairs (x)	65
20	Tau Pairs (y)	66
21	Tau Pairs (E/p)	67
22	Tau Pairs (dE/dx)	68

23	Muon Pairs (x)	69
24	Muon Pairs (y')	70
25	Muon Pairs (u')	71
26	Muon Pairs (dE/dx)	72
27	Signal MC: $\Upsilon(1S)$	73
28	Signal MC: $\Upsilon(2S)$	74
29	Signal MC: $\Upsilon(3S)$	75
30	Momentum Resolution	76
31	Calibration Fit	77
32	Control Fit	78
33	Signal Fit: $\Upsilon(1S)$	79
34	Signal Fit: $\Upsilon(2S)$	80
35	Signal Fit: $\Upsilon(3S)$	81
36	Calibration Results	82
37	Control Results	83
38	Results: $\Upsilon(1S)$	84
39	Results: $\Upsilon(2S)$	85
40	Results: $\Upsilon(3S)$	86

PREFACE

The author thanks the staff of the Cornell Electron Storage Ring and the members of the CLEO collaboration for the excellent data used in this analysis, the Publications Office at Cornell's Laboratory for Elementary-Particle Physics for preparing Figures 7, 8, 9, 12, and 15, and the following individuals, without whose help he never would have been able to do this: Morgen Bell, Martha Bell, Robert Brown, Peter Brusilovsky, Louis Coban, Timothy Cryan, Steven Dytman, Richard Galik, Brian Heltsley, David Kreinick, Adam Leibovich, Vincent Liu, Mary Love, Thomas Meyer, James Mueller, Donna Naples, Vittorio Paolone, Vladimir Savinov, Peter Siska, John Spade, Brian Staszal, Werner Sun, Julia Thompson, and Alan Watts.

I. ELEMENTARY PARTICLES

Particle physics is the branch of physics that studies the ultimate constituents of matter and the basic interactions of these constituents. Both are now described in terms of elementary particles, which seem to be structureless geometric points endowed with physical properties. Some of the elementary particles have mass and comprise ordinary matter. Others give rise to the forces encountered by material objects. In this picture fundamental forces are described in terms of particle exchange: two elementary particles interact by exchanging a third, which is said to mediate the interaction. For example, the repulsive electrostatic force between two electrons is mediated by the exchange of photons. The photons carry energy and momentum from one electron to the other, with the result that there is a net exchange of energy and momentum between the two electrons. The electrons, in turn, move apart, and we say that one has exerted a force on the other and *vice versa*.

Particle physics is often called high-energy physics because experiments in the field usually involve accelerating subatomic particles to speeds close to that of light, which causes the particles to acquire large kinetic energies. High energy is required for two reasons: first, the phenomena under investigation take place over incredibly small distances, comparable in size to atomic nuclei. Indeed, the elementary particles are believed to be point-like, and the size of the smallest composite objects is around 10^{-15} m. In order to study any object, radiation with a wavelength smaller than the object is required. The Planck relation¹ immediately tells us that the study of the very small requires radiation of high energy. If our probe is a beam of particles, the de Broglie relation² implies the same thing: namely that

¹The energy of a photon is inversely proportional to the wavelength of the corresponding electromagnetic radiation: $E = hc/\lambda$.

²The momentum of a material particle is inversely proportional to its associated quantum-mechanical wavelength: $p = h/\lambda$. A short wavelength requires a large momentum and hence a high kinetic energy.

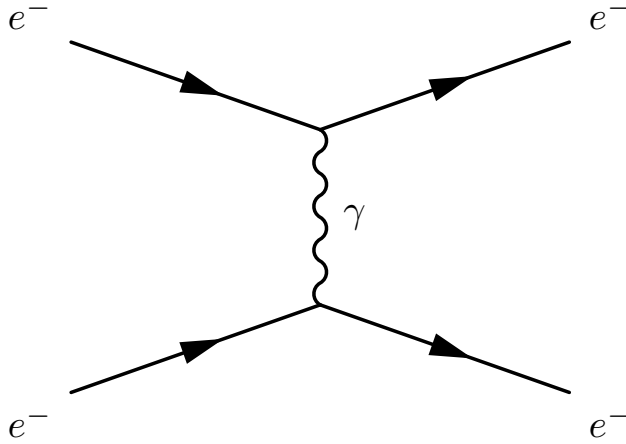


Figure 1: The repulsion of two electrons, described in terms of photon exchange.

high energy is required. The second reason for this is that many of the particles studied by particle physicists are relatively massive. Therefore, high energy is required to create them³.

The results of particle physics are described by the Standard Model. The model is a relativistic quantum field theory that describes seventeen elementary particles⁴ participating in three fundamental interactions. There are actually four known fundamental interactions: gravitation, electromagnetism, and the strong and weak nuclear forces, but the Standard Model does not describe gravitation⁵. This is because gravitation is so much weaker than the other three that its effects are not yet observable at the subatomic level [16].

The seventeen elementary particles can be subdivided into groups according to their intrinsic angular momentum, described by the particles' spin quantum numbers. Particles with $s = 1/2$ are fermions⁶ and comprise what we think of as matter in the macroscopic world. The elementary fermions are further subdivided into two groups: quarks, which participate in strong interactions, and leptons, which do not. The $s = 1$ particles are called gauge bosons⁷, and they correspond to forces. The gauge bosons are the particles which

³This is Einstein's mass-energy equivalence: $E = mc^2$.

⁴Additionally, the particles have their corresponding antiparticles, which have the algebraic signs of their internal quantum numbers flipped, *e.g.* electric charge: $Q \rightarrow -Q$.

⁵Gravitation is described by General Relativity, which is a classical field theory.

⁶They are so called because they obey Fermi-Dirac statistics. Therefore, they are subject to the Pauli exclusion principle.

⁷Bosons, including the gauge bosons, obey Bose-Einstein statistics.

Table 1: The four fundamental forces. Note: the graviton is still hypothetical.

INTERACTION	STRENGTH	RANGE	AFFECTS	MEDIATOR
strong	1	10^{-15} m	quarks	gluon
electromagnetic	10^{-2}	∞	quarks and charged leptons	photon
weak	10^{-5}	10^{-18} m	all elementary fermions	W and Z
gravitational	10^{-38}	∞	mass-energy	graviton

mediate the fundamental interactions. Finally the Standard Model contains one elementary particle that has $s = 0$, a boson called the Higgs boson, which has not yet been observed. However, its existence is predicted by the theory that unifies the electromagnetic and weak interactions.

From this basic picture much of the world can be described. Quarks are bound into composite particles by the strong interaction. Two types are possible: baryons, which are three-quark composites, and mesons, which are made of a quark bound to an antiquark. The quark composites have radii of the order of 10^{-15} m. Note that because of quantized angular momentum all baryons are fermions and all mesons are bosons⁸. Of the many possible composite particles only one is absolutely stable and does not spontaneously decay. This particle is the proton, a baryon composed of two u quarks and a d quark. A second baryon, the neutron, is stable when it binds with protons to form atomic nuclei⁹. Protons and neutrons, collectively called nucleons, are bound into nuclei by the residual effects of the strong interaction, whereby the quarks in adjacent nucleons exert attractive forces on one another. The largest nuclei are of the order of 10^{-14} m in radius.

At the next level of composite structure are atoms, which are formed when electrons are bound to nuclei via the electromagnetic interaction. Atoms are electrically neutral: the number of electrons bound to the nucleus equals the number of protons in the nucleus. When there is a surplus or deficiency of electrons, an ion results. Typical atomic radii are of the

⁸Baryons can have $s = 1/2$ or $s = 3/2$, and mesons can have $s = 0$ or $s = 1$.

⁹Free neutrons decay via the weak interaction with a mean lifetime of 885.7 s.

Table 2: The known elementary particles.

FERMIONS $s = 1/2$ (matter)					GAUGE BOSONS $s = 1$ (forces)			
TYPE	CHARGE	I	II	III		CHARGE	INTERACTION	
quarks	(<i>u</i> -type)	$+2/3$	<i>u</i>	<i>c</i>	<i>t</i>	γ	0	electromagnetic
	(<i>d</i> -type)	$-1/3$	<i>d</i>	<i>s</i>	<i>b</i>	<i>g</i>	0	strong
leptons	(neutral)	0	ν_e	ν_μ	ν_τ	<i>Z</i>	0	weak
	(charged)	-1	e^-	μ^-	τ^-	W^-	-1	

order of 10^{-10} m. This enormous size increase — atoms are four to five orders of magnitude larger than their nuclei — is attributable to the different binding forces.

Finally, atoms can bind electromagnetically to form molecules. Molecules are also electrically neutral, and here the result of an electron imbalance is called a molecular ion. Molecular radii range from the order of 10^{-10} m for simple inorganic compounds to the order of 10^{-9} m for the compounds commonly used in organic synthesis, and the macromolecules encountered in biochemistry are larger still. Huge numbers of atoms and molecules in aggregate comprise the solids, liquids, and gases of our macroscopic world. At this level of structure and above, matter is electrically neutral, and this allows the effects of gravity to become apparent¹⁰.

¹⁰The weakness of gravity means that huge numbers of particles are required: it takes around 10^{24} atoms to produce a measurable gravitational attraction.

Table 3: Some composite particles.

BARYONS (three-quark composites)				MESONS (quark-antiquark pairs)			
NAME	QUARKS	CHARGE	s	NAME	QUARKS	CHARGE	s
proton p	uud	+1	$1/2$	pion π^+	$u\bar{d}$	+1	0
neutron n	udd	0	$1/2$	pion π^0	$\frac{u\bar{u}-d\bar{d}}{\sqrt{2}}$	0	0
lambda Λ^0	uds	0	$1/2$	rho ρ^+	$u\bar{d}$	+1	1
omega Ω^-	sss	-1	$3/2$	upsilon Υ	$b\bar{b}$	0	1

II. SYMMETRY

A symmetry is an operation that leaves a system in a configuration which is indistinguishable from the system's original configuration. Examples include rotating an equilateral triangle by 120° about its center, reflecting the triangle through a line joining one of its vertices to the midpoint of the opposite side, or, in fact, doing nothing at all to the triangle¹. Symmetries do not need to be visual, geometric symmetries, although the examples just described are; any abstract operation that meets the criterion given above is a symmetry. For example, in Newtonian mechanics the equations of motion of a system are unchanged by the operation of relocating the origin of coordinates, and we therefore say that systems described by Newtonian mechanics possess translational symmetry.

Noether's theorem, proved in 1915, is a cornerstone of modern theoretical physics. It states that any differentiable symmetry of the action² of a physical system has a corresponding conservation law. In other words, every symmetry implies a conserved quantity, and every conserved quantity reveals an underlying symmetry [12]. Conserved quantities, of course, are quantities that are unchanged by transformations of a system. Such quantities include energy, momentum, and electric charge. To continue our example, the translational symmetry of Newtonian mechanics corresponds to the law of conservation of momentum.

The Standard Model possesses a number of symmetries. Among them are: symmetry with respect to translation in time, which corresponds to the conservation of energy, symmetry with respect to translation in space, which corresponds to the conservation of momentum, and symmetry with respect to rotation about a point, which corresponds to the conservation

¹The null operation is technically a symmetry. Doing nothing at all to a system leaves it in a state indistinguishable from the original.

²The action is the integral over time of a system's Lagrangian function.

of angular momentum. The Standard Model also possesses symmetry with respect to gauge transformations. In electrodynamics a gauge transformation adds the gradient of an arbitrary function f to the magnetic vector potential while subtracting the partial derivative of f with respect to time from the electric scalar potential:

$$\begin{aligned}\mathbf{A} &\rightarrow \mathbf{A} + \nabla f \\ \phi &\rightarrow \phi - \frac{\partial f}{\partial t}\end{aligned}$$

This is an extension of the idea that only differences in potential energy have physical significance. Electrodynamical gauge symmetry corresponds to the conservation of electric charge. The other fundamental forces described by the Standard Model, namely the strong and weak nuclear forces, also possess gauge symmetry, corresponding to the conservation of color charge and weak isospin respectively³.

Since the Standard Model describes elementary particles in terms of relativistic quantum fields, it possesses symmetry with respect to Lorentz boosts as required by the special theory of relativity. A boost is the operation of moving from one inertial coordinate system⁴ to another, and boosts are described mathematically by Lorentz transformations, which are rotations of Minkowski spacetime⁵ about the origin. Because of this, Lorentz transformations encompass rotations about the origin in three-dimensional Euclidean space. Now, moving from one inertial system to another may involve a translation of the spacetime origin in addition to a rotation about that origin, so the complete relativistic symmetry is a symmetry with respect to translations (in space and time) and Lorentz rotations. This symmetry, called Poincaré symmetry, corresponds to the relativistic versions of the conservation of energy, momentum, and angular momentum.

In the early 1950s it was shown by Julian Schwinger and others that the relativistic quantum field theories that actually describe nature possess symmetry with respect to the following combined operation: charge conjugation C , which replaces a particle with its antiparticle,

³Like electric charge, these quantities are quantum numbers, and just as a particle which has electric charge participates in electromagnetic interactions by emitting and absorbing photons, so do particles with color charge and weak isospin participate in the strong and weak interactions. In the former case the particles emit and absorb gluons; in the latter case they emit and absorb W and Z bosons.

⁴An inertial system is one in which Newton's first law is valid, and any two inertial systems move at a constant velocity with respect to one another.

⁵This is the four-dimensional coordinate system in which the special theory of relativity is set.

inversion of spatial coordinates P , and time reversal T . Actually, the electromagnetic and strong interactions are symmetric with respect to C , P , and T separately, and they are said to conserve C -parity, parity, and T -parity, respectively. The weak interaction, however, has been shown experimentally to violate C -parity conservation and parity conservation, as well as the product of parities which would be conserved by the combined operation CP . Therefore, the weak interaction is only symmetric with respect to the combined operation CPT .

In addition to the symmetries described so far, the Standard Model also possesses symmetries arising from the fact that the fields describing the elementary fermions may be shifted by an arbitrary phase angle ($\psi \rightarrow e^{i\alpha}\psi$) without changing the resulting physics. Four such phase shifts are possible: all quark fields may be shifted by the same phase; the electron and its neutrino may be shifted by some different phase; and the muon with its neutrino and the τ with its neutrino may likewise be shifted. These four symmetries correspond to the conservation of baryon number B , electron number L_e , muon number L_μ , and τ number L_τ . Since L_e , L_μ , and L_τ , are separately conserved, their sum, lepton number L , is also conserved.

The symmetries with respect to phase shift are called accidental symmetries of the Standard Model because they were not explicitly postulated at the outset of its construction, and there is no compelling physical reason for these numbers to be conserved by nature. Experimentally, the phenomenon of neutrino oscillation [2] shows that the lepton flavor numbers are not in fact conserved by nature. Theoretically, even within the Standard Model non-perturbative quantum effects⁶ violate all of these accidental symmetries, leaving only the combination $B - L$ intact.

⁶These are effects not described in terms of particle exchange.

Table 4: Symmetries of the Standard Model and their corresponding conservation laws.

SYMMETRY	CONSERVED QUANTITY
spacetime translations and boosts	energy momentum angular momentum
gauge transformations	electric charge color charge weak isospin (3 rd comp.)
charge conjugation C spatial inversion P time reversal T	C -parity (strong, EM) parity (strong, EM) T -parity (strong, EM)
<i>N.B.:</i> The weak interaction is symmetric only under the combined operation CPT .	
fermion phase	baryon number electron number muon number τ number

III. LEPTON FLAVOR VIOLATION

The conservation of lepton number is one of the accidental symmetries of the Standard Model. If we assign lepton number $L = +1$ to leptons, $L = -1$ to antileptons, and $L = 0$ to all other particles we find that the net lepton number of all particles going into an interaction equals the net lepton number of all particles coming out of the interaction. In other words the net lepton number of the system is unchanged by the interaction. For example, neutron beta decay can be understood in terms of the following weak interaction:

$$d \rightarrow u + e^- + \bar{\nu}_e$$

One of the neutron's d quarks ($L = 0$) turns into a u quark ($L = 0$) by emitting a W^- boson ($L = 0$), which in turn decays into an electron ($L = +1$) and an electron antineutrino ($L = -1$). The net lepton number is zero before and after the interaction.

As another example, consider the antineutrinos produced by the decay of negatively charged pions: $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$. The pion consists of a u antiquark bound to a d quark, which annihilate one another. This annihilation produces a W^- boson that quickly decays into a muon and a muon antineutrino, so in terms of elementary particles the interaction is

$$\bar{u} + d \rightarrow \mu^- + \bar{\nu}_\mu$$

Again lepton number is conserved ($0+0 = 1-1$). If the muon antineutrinos are now directed at a target of protons, the following reaction occurs:

$$\bar{\nu}_\mu + p \rightarrow \mu^+ + n$$

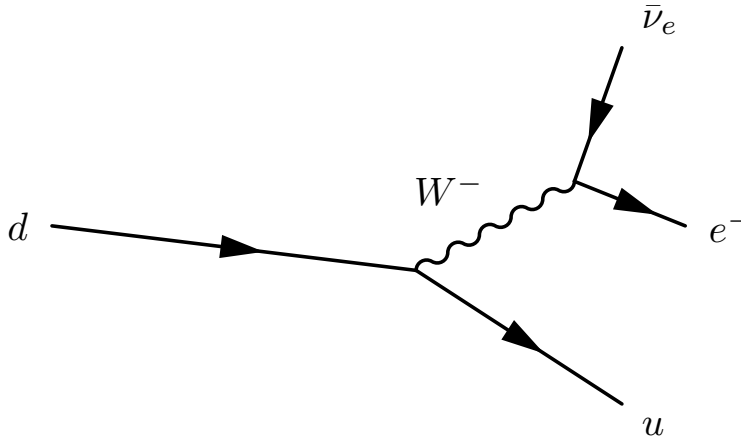


Figure 2: The fundamental interaction behind neutron beta decay.

Or in terms of elementary particles:

$$\bar{\nu}_\mu + u \rightarrow \mu^+ + d$$

The muon antineutrino and one of the proton's u quarks annihilate, creating an antimuon and a d quark in the process. The d quark remains bound to the original proton's other two quarks, and a neutron results. Once again, lepton number is conserved ($-1 + 0 = -1 + 0$). Lepton number would also be conserved by this interaction:

$$\bar{\nu}_\mu + u \rightarrow e^+ + d$$

The only difference is that a positron¹ ($L = -1$) has been substituted for the antimuon ($L = -1$). However, no experiment to date has observed this.

If instead of decaying pions we use decaying neutrons as the source of antineutrinos, then positrons are always produced and antimuons are never observed:

$$n \rightarrow p + e^- + \bar{\nu}_e$$

$$\bar{\nu}_e + p \rightarrow e^+ + n$$

¹Positrons are antielectrons. The name is kept for historical reasons.

It appears that the antineutrinos are somehow associated with the particular type of charged lepton involved in their production. Finally, we note that the following charged lepton decays have never been observed:

$$\mu^- \rightarrow e^- + \gamma$$

$$\tau^- \rightarrow e^- + \gamma$$

$$\tau^- \rightarrow \mu^- + \gamma$$

We conclude that lepton number is separately conserved for each of the three charged-neutral lepton pairs²: e^- with ν_e , μ^- with ν_μ , and τ^- with ν_τ . To all elementary particles we assign a quantum number called electron number (L_e) with $L_e = +1$ for electrons and electron neutrinos, $L_e = -1$ for antielectrons and electron antineutrinos, and $L = 0$ for all other particles. Muon number (L_μ) and τ number (L_τ) are similarly defined. Collectively L_e , L_μ , and L_τ are called lepton flavor numbers or simply lepton flavor, and we say that lepton flavor is conserved by all particle interactions.

Lepton flavor conservation allows us to understand why the following interaction has never been observed:

$$\bar{\nu}_\mu + u \rightarrow e^+ + d$$

Before the interaction we have $L_e = 0$, $L_\mu = -1$, and $L_\tau = 0$; afterward we have $L_e = -1$, $L_\mu = 0$, and $L_\tau = 0$. Both the conservation of electron number and the conservation of muon number have been violated. The neutrinoless decays of charged leptons, which were mentioned above, also violate lepton flavor conservation:

$$l \rightarrow l' + \gamma$$

Here l and l' represent any two charged leptons (or charged antileptons for that matter) of differing flavor.

As shorthand we define lepton flavor violation to be the non-conservation of lepton flavor number by an interaction. The following are all examples lepton flavor violation³:

$$\phi \rightarrow e^- + \mu^+$$

²The pairs are sometimes called lepton families or generations.

³Again we could equally well discuss the charge-conjugate interactions, where all particles are replaced with their corresponding antiparticles.

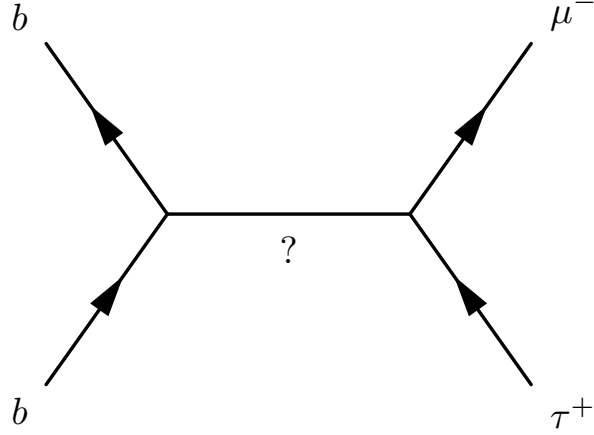


Figure 3: A lepton flavor violating decay. The details of the interaction are unknown.

$$\Upsilon \rightarrow \mu^- + \tau^+$$

$$q\bar{q} \rightarrow l + \bar{l}'$$

Where the last interaction is general and q represents any quark⁴, \bar{q} its antiquark, l a charged lepton, and l' a charged antilepton of differing flavor.

The subject of the analysis described in this dissertation is a search for lepton flavor violation in decays of Υ mesons, using data collected with the CLEO III detector. Specifically we are looking for $\Upsilon(nS) \rightarrow \mu^\mp + \tau^\pm$, where $n = 1, 2, \text{ or } 3$. The Υ is a composite particle made of a b quark bound to a b antiquark. The quarks have their spins aligned, so the Υ has a spin quantum number of 1. The spectroscopic notation S means that the quarks have zero orbital angular momentum relative to one another, and the integer n refers to radial excitation states, where $n = 1$ is the lowest energy Υ state⁵, $n = 2$ is the first radial excitation, $n = 3$ the second, and so on. The lightest four Υ mesons are described [20] in Table 5. The third column shows the width of the resonance, which is inversely proportional to the mean lifetime of the particle, and the fourth column shows the partial width for the decay $\Upsilon(nS) \rightarrow e^- + e^+$. The ratio of the partial width for a particular decay to the width of

⁴Except for the t quark, which is so short-lived that it cannot bind into mesons. In this case we should write $t + \bar{t} \rightarrow l + \bar{l}'$.

⁵The $\Upsilon(1S)$ is not the ground state of the $b\bar{b}$ system. The ground state is the $\eta_b(1S)$, which has antiparallel quark spins.

Table 5: The four lightest Υ mesons.

RESONANCE	MASS	Γ	$\Gamma(e^-e^+)$
$\Upsilon(1S)$	9,460 MeV	53.0 keV	1.340 keV
$\Upsilon(2S)$	10,023 MeV	43.0 keV	0.612 keV
$\Upsilon(3S)$	10,355 MeV	26.3 keV	0.443 keV
$\Upsilon(4S)$	10,579 MeV	20.5×10^3 keV	0.272 keV

the resonance is the probability that a given particle will decay according to the particular decay mode.

With our statistics we do not expect to see lepton flavor violation in decays of $\Upsilon(4S)$ mesons because the $\Upsilon(4S)$ resonance is above the energy threshold for the production of pairs of B mesons⁶. Therefore, it decays into pairs of B mesons almost 100% of the time⁷. Instead we use $\Upsilon(4S)$ data and data collected 40 – 60 MeV below the $\Upsilon(4S)$ energy to optimize the analysis for sensitivity. The search is limited to the $\mu\tau$ -channel. This channel should be the most sensitive to lepton flavor violation, since we expect the unknown mediators of lepton flavor violation to couple more strongly to heavier leptons; otherwise lepton flavor violation, if it exists, would have been observed long ago in electron-positron experiments.

⁶A B meson consists of a b quark bound to some other kind of quark. The lightest B meson is the B^- , which consists of a b quark bound to a u antiquark. Its mass is 5,279 MeV, and twice this falls just shy of the mass of the $\Upsilon(4S)$.

⁷The partial width for $\Upsilon(4S) \rightarrow e^- + e^+$ is already small compared to the width of the $\Upsilon(4S)$, and the partial width for lepton flavor violation, $\Gamma(\mu^\mp \tau^\pm)$, has to be smaller still.

IV. PRIOR WORK

The probability that a given particle will decay according to a particular decay mode is called the branching fraction¹, \mathcal{B} , of the decay. One of the primary concerns of experimental particle physics is the determination of branching fractions, and this analysis in particular seeks to establish a value (or at least an upper limit) for $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm)$, the probability that a single Υ meson will decay into a muon and a τ lepton. Prior to our search, several theoretical and semi-empirical estimates of this branching fraction have been made.

If the lepton flavor violating decay $\Upsilon \rightarrow \mu^\mp + \tau^\pm$ is possible, then by unitarity so is the decay $\tau^\mp \rightarrow \mu^\mp + \mu^- + \mu^+$; therefore, $\tau^\mp \rightarrow \mu^\mp + e^- + e^+$ is also possible. Nussinov, Peccei & Zhang [15] used the experimental upper limits on the branching fractions of the latter decays to constrain $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm)$. Using their model-independent estimate with recent data on neutrinoless τ decays [20], one arrives at $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm) < 10^{-3}$, a number which is quite encouraging to an experimentalist.

Silagadze [18] suggested using lepton flavor violating decays of Υ and J/ψ mesons as a probe of the energy scale of quantum gravity². He estimates, using a four-fermion vertex in the framework of effective field theory [8], the following ratio of partial widths:

$$\frac{\Gamma(\Upsilon \rightarrow \mu^\mp + \tau^\pm)}{\Gamma(\Upsilon \rightarrow \mu^- + \mu^+)} = \frac{1}{2q_b^2} \left(\frac{\alpha'}{\alpha} \right)^2 \left(\frac{M_\Upsilon}{\Lambda} \right)^4$$

Here q_b is the charge of the b quark, α is the fine structure constant, M_Υ is the mass of the Υ meson, Λ is the energy scale of quantum gravity, and α' is the strength of the four-fermion

¹Also called the branching ratio because it is the ratio of the partial width for a particular decay mode to the total width of the resonance.

²Silagadze was interested in quantum gravity, but Λ could equally well represent the energy scale of any new physics beyond the scope of the Standard Model.

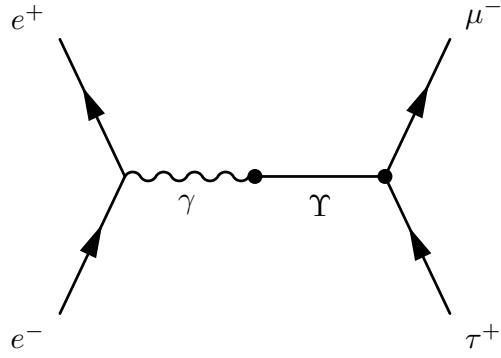


Figure 4: The lepton flavor violating decay $\Upsilon \rightarrow \mu^- + \tau^+$.

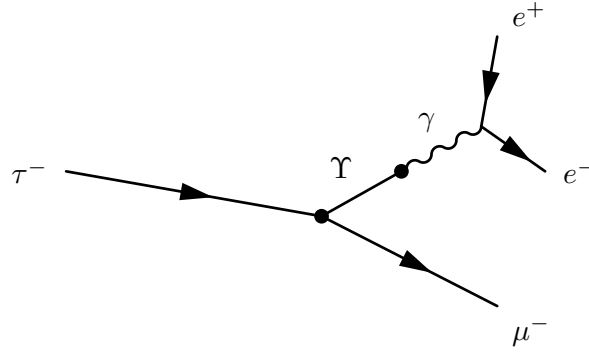


Figure 5: The decay $\tau^- \rightarrow \mu^- + e^- + e^+$.

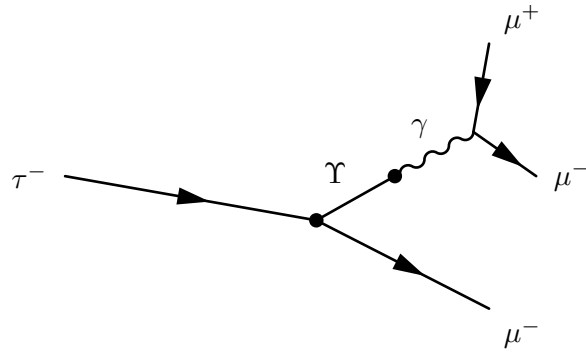


Figure 6: The decay $\tau^- \rightarrow \mu^- + \mu^- + \mu^+$.

Table 6: Summary of relevant prior work.

SEMI-EMPIRICAL ESTIMATE	CL	RESEARCHERS
$\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm) < 10^{-2}$	90%	Nussinov <i>et al.</i>
$\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm) < 10^{-3}$	90%	Nussinov <i>et al.</i> (recent data)
THEORETICAL ESTIMATE		RESEARCHER
$\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm) \approx 2 \times 10^{-5}$		Silagadze
MEASUREMENT	CL	RESEARCHERS
$\mathcal{B}(B^0 \rightarrow \mu^\mp + \tau^\pm) < 3.8 \times 10^{-5}$	90%	Bornheim <i>et al.</i> (CLEO)
$\mathcal{B}(\bar{B}^0 \rightarrow \mu^\mp + \tau^\pm) < 3.8 \times 10^{-5}$	90%	Bornheim <i>et al.</i> (CLEO)
$\mathcal{B}(J/\psi \rightarrow \mu^\mp + \tau^\pm) < 2.0 \times 10^{-6}$	90%	Ablikim <i>et al.</i> (BES)

lepton flavor violating coupling. Solving this equation for Λ yields the following:

$$\Lambda = \left(\frac{1}{2q_b^2} \frac{\Gamma(\Upsilon \rightarrow \mu^- + \mu^+)}{\Gamma(\Upsilon \rightarrow \mu^\mp + \tau^\pm)} \right)^{1/4} \sqrt{\frac{\alpha'}{\alpha}} M_\Upsilon$$

Assuming $\alpha' \approx 1$ and $\Lambda \approx 1$ TeV, these formulae predict $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm) \approx 2 \times 10^{-5}$, which is very interesting in light of the 2.1×10^7 $\Upsilon(1S)$ decays recorded by the CLEO III detector [7, 4].

Previously the CLEO collaboration searched for lepton flavor violation in decays of B mesons [9], while the BES experiment looked for lepton flavor violating decays of the J/ψ [1]. The search described in this dissertation is the first published [14] search for lepton flavor violating decays of Υ mesons.

V. THE CORNELL ELECTRON STORAGE RING

Cornell University has been actively involved in the study of elementary particles since the earliest days of the field [6]. Cornell built its first particle accelerator, a 1-MeV proton cyclotron, in 1935, the same year that Hideki Yukawa of Osaka University predicted the meson later identified as the pion. The following year, Carl Anderson of the California Institute of Technology discovered the muon in cosmic-ray experiments. These achievements, Yukawa's meson theory and Anderson's cosmic-ray experiments, can arguably be considered the birth of particle physics.

Cornell's original cyclotron remained in service until 1956, but its successor was in place well before then. The end of the Second World War brought with it an incredible interest in nuclear and particle physics. Indeed, the mode of research pioneered at Los Alamos National Laboratory during the Manhattan Project continues to dominate physics to this day: the so-called "big science" approach, with dozens or hundreds of physicists collaborating on large experiments. In 1945 Cornell established its Laboratory of Nuclear Studies and began constructing a new building to house a second particle accelerator. Cornell alumnus Floyd R. Newman donated one million dollars for construction, and the F. R. Newman Laboratory was formally dedicated in October of 1948. The Office of Naval Research provided funding for a 300-MeV electron synchrotron, which was completed in 1949.

This first synchrotron was operational until 1954 and was followed in quick succession by a 1-GeV machine (1954 – 1964) and a 2.2-GeV machine (1964 – 1969). These two machines were essentially large-scale upgrades of the original 300-MeV electron synchrotron, with each successive machine being constructed largely out of the parts of its predecessor. However, it soon became clear that the next upgrade would require an entirely new machine housed in an entirely new facility. To increase the energy of a synchrotron, one can either increase

the field strength of the guide magnets, or one can increase the radius of the machine itself. The upgrades to the original 300-MeV synchrotron were essentially magnetic field upgrades, and this approach had reached its practical limit.

The successor, which was to operate at 10 GeV, required a radius of 122 m. The machine's circumference, therefore, was going to be 768 m, or nearly half a mile. In 1965 the laboratory's director, Robert R. Wilson, won approval from Cornell's administration to house the new machine in a tunnel to be dug 12 m beneath Cornell's Alumni Field. The National Science Foundation provided twelve million dollars, and construction began on the R. R. Wilson Synchrotron Laboratory. The new laboratory and its 10-GeV synchrotron were finished in 1968 and formally dedicated in October of that year.

The Cornell 10-GeV Synchrotron, as it came to be known, was built for fixed-target experiments: beams of high-energy particles were directed at stationary targets of various materials, which produced sprays of secondary particles to be measured and analyzed. By the 1970s, however, it was realized that experiments in which two oppositely directed beams were made to collide would greatly increase the energy available for the creation of new particles. In 1975 a proposal to augment the synchrotron, now operating at 12 GeV, with a storage ring was put forward. This would allow beams of electrons and positrons, accelerated by the synchrotron, to be stored for colliding-beam experiments. The \$20.6 million proposal was approved in 1977, and construction began on the Cornell Electron Storage Ring (CESR, pronounced Caesar). By August of 1979, CESR could produce measurable colliding beams, inaugurating what was to become nearly two decades of research for the machine¹.

Fortuitously, CESR was designed to operate at a center-of-mass energy in the range of 3 – 12 GeV. This energy range is ideal for the study of mesons containing charm and bottom quarks, the latter of which were completely unknown when CESR was being designed. The $\Upsilon(1S)$ meson was unexpectedly discovered at the Fermi National Accelerator Laboratory in 1977, and it was quickly understood to be a bound state between a fifth quark, now called the bottom quark, and its antiparticle. The mass of the $\Upsilon(1S)$ is about 9.5 GeV, almost exactly the energy for which CESR was being designed! The $\Upsilon(2S)$ was soon discovered with a mass of 10.0 GeV, and evidence for the $\Upsilon(3S)$ near 10.4 GeV was building. Indeed,

¹CESR's last beam collision occurred on March 3, 2008.



Figure 7: The Cornell Electron Storage Ring and the Wilson Synchrotron Laboratory.

verification of the $\Upsilon(3S)$, and the first evidence for the higher-mass $b\bar{b}$ bound states $\Upsilon(4S)$, $\Upsilon(10860)$, and $\Upsilon(11020)$ came from analyzing collisions produced by CESR.

The operation of CESR is as follows: first, electrons emitted from a heated filament are accelerated in a linear accelerator (linac) to an energy of 300 MeV. Positrons are produced by causing some of the accelerated electrons to strike a tungsten plate located at an intermediate point along the linac. This produces a spray of electrons, positrons, and X-rays, from which the positrons are selected and focused. The positrons are then accelerated to 150 MeV along the remaining length of the linac. From here, the 300-MeV electrons and 150-MeV positrons are fed into the synchrotron, where electrons circulate in the counter-clockwise direction when viewed from above². The synchrotron accelerates the electrons and positrons to the desired final energy, which can be as high as 6 GeV. Once the particles are at the desired energy they are transferred to the storage ring. The entire acceleration cycle is repeated 60 times per second for about 10 minutes, first clockwise with positrons, then counterclockwise with electrons, until the required beam currents are built up in the storage ring. When the beams are ready³, they are made to pass through each other at a point at the south end of the ring called the interaction point. The electrons and positrons collide and annihilate one another, and the energy released in this process creates subatomic particles, which must be detected and analyzed. This is the job of the CLEO detector.

²The positrons, of course, circulate in the opposite direction.

³Circulating beams of high-energy electrons and positrons produce copious amounts of X-rays, which are themselves valuable for research. Cornell established a separate laboratory, the Cornell High-Energy Synchrotron Source (CHESS), to study these X-rays.

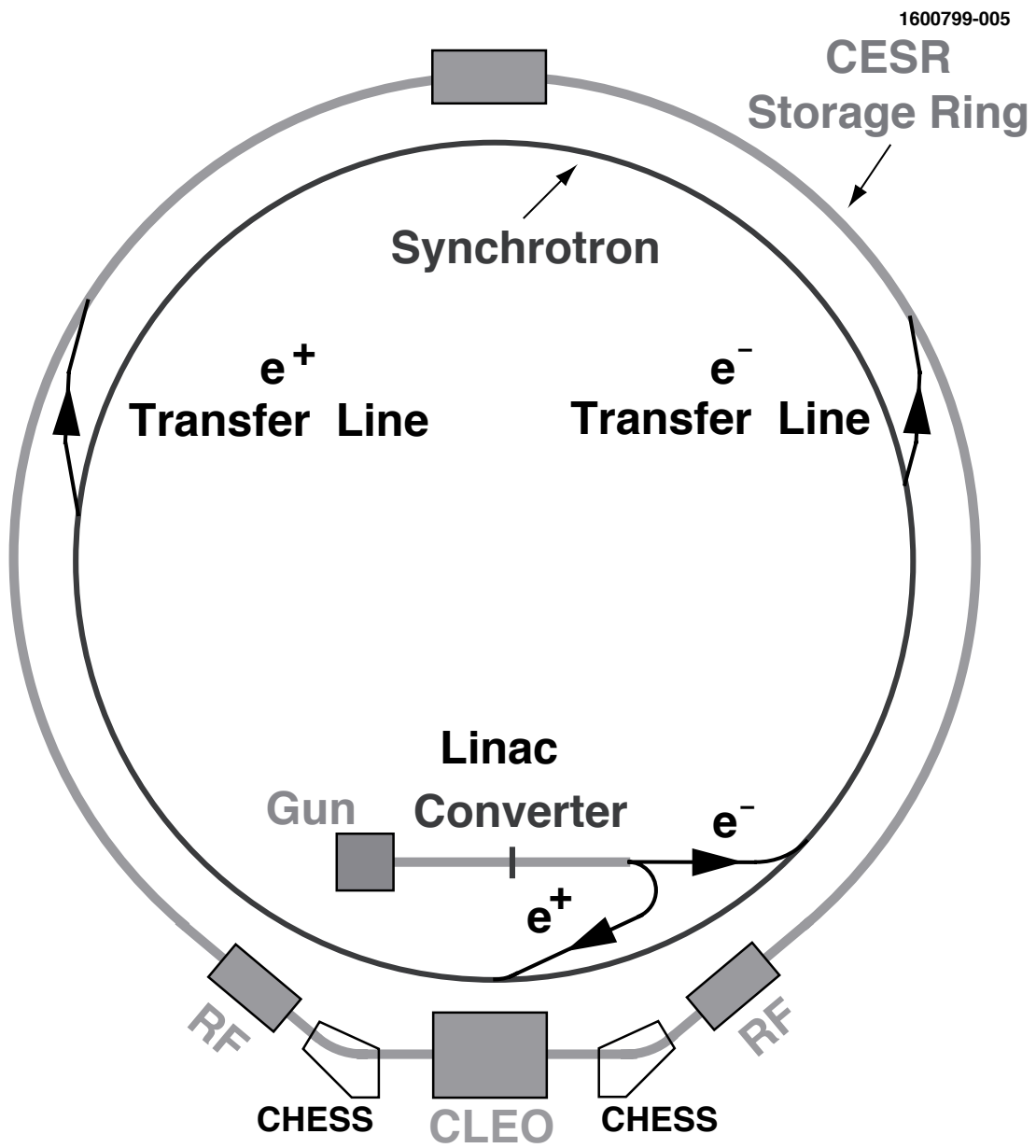


Figure 8: Schematic of the CESR accelerator complex.

VI. THE CLEO DETECTOR

The CLEO¹ detector is a general-purpose particle detector [19], which was used to detect and measure the subatomic particles resulting from CESR's high-energy electron-positron collisions. CLEO was in operation from October 1979 until March 2008, and during this time the detector underwent two major and three minor upgrades. The latest version of CLEO is CLEO-c, which started taking data in 2003. CLEO is also the name of the collaboration of physicists who study the data taken with the detector.

The data used in our analysis were taken with the version of the detector known as CLEO III². CLEO is a hermetic detector; by design it covers as much of the solid angle around CESR's interaction point as possible. The different subsystems of the detector have differing angular coverage, so we use the subsystem with the least coverage to quantify the overall coverage of the detector. This subsystem is the ring-imaging Cherenkov detector, which covers 83% of the 4π -sr solid angle around the interaction point.

The innermost piece of CLEO is the beam pipe, which surrounds CESR's interaction point. The beam pipe is a double-walled beryllium cylinder of inner diameter 21.165 mm, outer diameter 22.330 mm, and length 300.0 mm³. The inner surface of the pipe is coated with 10 μm of gold, and the 0.5-mm annular region between the walls of the pipe is filled with coolant.

Surrounding the beam pipe is the first of the detector's subsystems, which in CLEO

¹Rare among large experiments, CLEO is not an acronym; it is, rather, a pun. The detector is named after Cleopatra, whose companion was Julius Caesar.

²CLEO-c differs from CLEO III in only two respects: first, the innermost subsystem of CLEO-c is a stereo drift chamber, whereas CLEO III used a silicon vertex detector. Second, CLEO-c's solenoidal magnetic field is maintained at 1.0 T; it was held at 1.5 T in CLEO III.

³CLEO's coordinate system is a cylindrical system (ρ, ϕ, z) whose origin is at the interaction point. The $+z$ -axis points in the direction of the electron beam. In this system the beam pipe extends from $z = -0.150$ m to $z = +0.150$ m.

III was the silicon vertex detector (SVD). The SVD used 447 double-sided silicon wafers to detect the passage of charged particles. It was used for high-resolution particle tracking near the interaction point. The wafers were arranged into four cylinders concentric with the beam pipe. The individual wafers were identical and measured 52.6 mm long by 27.0 mm wide by 0.3 mm thick. The first cylinder contained 21 wafers and was 25.0 mm in radius by 160 mm long. The second contained 40 wafers and was 27.5 mm in radius by 213 mm long. The third contained 126 wafers and was 70.0 mm in radius by 373 mm long, and the fourth contained 260 wafers and was 101.0 mm in radius by 533 mm long. The SVD covered 93% of 4π sr.

The CLEO III drift chamber (DR3) is the next detector subsystem, and it also is used for charged particle tracking. The DR3 is cylindrical and concentric with the beam pipe. It contains thousands of conducting wires which are parallel or nearly parallel to the axis of the cylinder, and its volume is filled with a mixture of helium and propane. The DR3, like all drift chambers, works by measuring the charge deposited on high-voltage sense wires by gas ions, which are created when a high-energy charged particle passes through the drift chamber's gas. The sense wires are held at a high electrostatic potential relative to field wires, which are grounded. The ratio of field wires to sense wires in the DR3 is 2:1. Negatively charged ions accelerate toward the sense wires, and the accumulated charge is measured. The DR3 consists of an inner section with stepped endplates and an outer section with conical endplates. The inner section extends from 132 – 342 mm in radius and contains 1,696 axial⁴ sense wires. The outer section extends from 370 – 790 mm in radius and contains 8,100 stereo⁵ sense wires. The DR3 covers 93% of 4π sr.

The ring-imaging Cherenkov detector (RICH) is the next subsystem [3]. It is used for particle identification. When a charged particle passes through an insulating medium at a speed greater than the speed of light in that medium Cherenkov radiation results. The photons radiated by the charge when it enters the medium do not move fast enough to get out of the way of the charge itself, and the radiation piles up, forming a conical wavefront⁶ whose

⁴These are parallel to the beam pipe.

⁵These make a slight angle (between 21 and 28 mrad) with respect to the axis of the DR3. This allows the z -coordinate of a charge deposit to be measured.

⁶A similar thing happens when an airplane exceeds the speed of sound.

precise shape depends on the speed of the charged particle. By measuring these Cherenkov photons with a suitable photodetector, the shape of the wavefront, and hence the speed of the charged particle, can be determined. Combining this measurement of speed with an independent measurement of the particle's momentum allows the mass of the charged particle to be determined, which in turn allows the particle species to be identified. CLEO's RICH uses lithium fluoride as its Cherenkov radiator. It extends from 800 – 975 mm in ρ , is 2.4 m long in z , and covers 83% of 4π sr.

Next comes CLEO's electromagnetic crystal calorimeter (CC), which is used to measure the energy of charged particles and photons [13]. The CC consists of a barrel section and two end caps, and it contains 7,784 cesium iodide crystals. Charged particles entering the crystals excite atoms in the material, causing them to emit light. This light is then collected by silicon photodiodes mounted on the rear face of the crystal. High energy photons are detected by the showers of secondary charged particles they produce⁷. Each crystal is 300 mm long and measures 50 mm by 50 mm at the front face. The barrel consists of 6,144 crystals that are tapered toward the front face⁸ and arranged so that their front faces point toward the interaction point. The end caps each contain 820 crystals that are rectangular in cross section and arranged parallel to the beam pipe. The barrel has an inner radius of 1.02 m, an outer radius of 1.32 m, and extends 3.26 m in z at the inner radius. The end caps extend from 0.434 m to 0.958 m in ρ . Their front faces are at $z = \pm 1.308$ m, and their back faces are at $z = \pm 1.608$ m. The CC covers 95% of 4π sr.

The next of CLEO's components is not a detector; it is the superconducting solenoid. This produces a magnetic field, which is uniform to 0.2% over the tracking volume⁹. CLEO III used a 1.5-T magnetic field. The solenoidal magnetic field is parallel to the $+z$ -axis and causes charged particles to move in helical trajectories. The radius, R , of the helix is proportional to the component particle's momentum, p_{\perp} , normal to the magnetic field, B : $p_{\perp} = |q|BR$, where q is the particle's charge. This is the well known cyclotron formula. Thus, determining the radius of a charged particle's trajectory through a uniform magnetic field allows one to determine its momentum. The solenoid and its surrounding cryostat extend

⁷These secondaries then excite atoms in the crystal and cause them to emit light.

⁸There are 24 slightly different tapered shapes.

⁹The tracking volume contains the SVD and the DR3.

from 1.4 – 1.6 m in ρ and are 3.8 m long in z .

The muon detector (MU) is the final subsystem [10]. As its name implies, it is used to track muons. The MU uses arrays of proportional chambers embedded in the solenoid's return iron, which essentially blocks all other types of particles. A proportional chamber is a gas-filled tube down the center of which runs a single high voltage anode wire, with the tube itself being used as the cathode. It works similarly to a drift chamber in that it measures the charge deposited on the anode wire by gas ions. As in a drift chamber, high-energy charged particles passing through the proportional chamber ionize its gas¹⁰. Negatively charged gas ions accelerate toward the anode, and the accumulated charge is measured. Like the CC, the MU has a barrel section and two end caps. A single plane of chambers consists of three layers of counters, which in turn consist of eight adjacent chambers. An individual chamber consists of a copper-beryllium anode wire, plated with 50 μm of silver, run down the center of a square cell measuring 9 mm on a side. Three of the four walls of the cell are coated with graphite and used as cathodes, and eight adjacent cells are placed in a polyvinyl chloride enclosure to make a single counter, which is 83 mm wide by 10 mm thick by 5 m long. The counters are sandwiched between copper cathode strips and a copper ground. The copper cathodes are 80 mm wide and run perpendicular to the anode wires. They are used to determine the z -coordinate of a hit. Finally, three layers of counters separated by foam form a detection plane. There are three planes of chambers in the barrel at depths of 0.36, 0.72, and 1.08 m in the iron and one plane of chambers in each of the two end caps at $z = \pm 2.7$ m. The MU covers 85% of 4π sr.

¹⁰The MU uses the same gas supply as the DR3.

CLEO-III

2230104-001

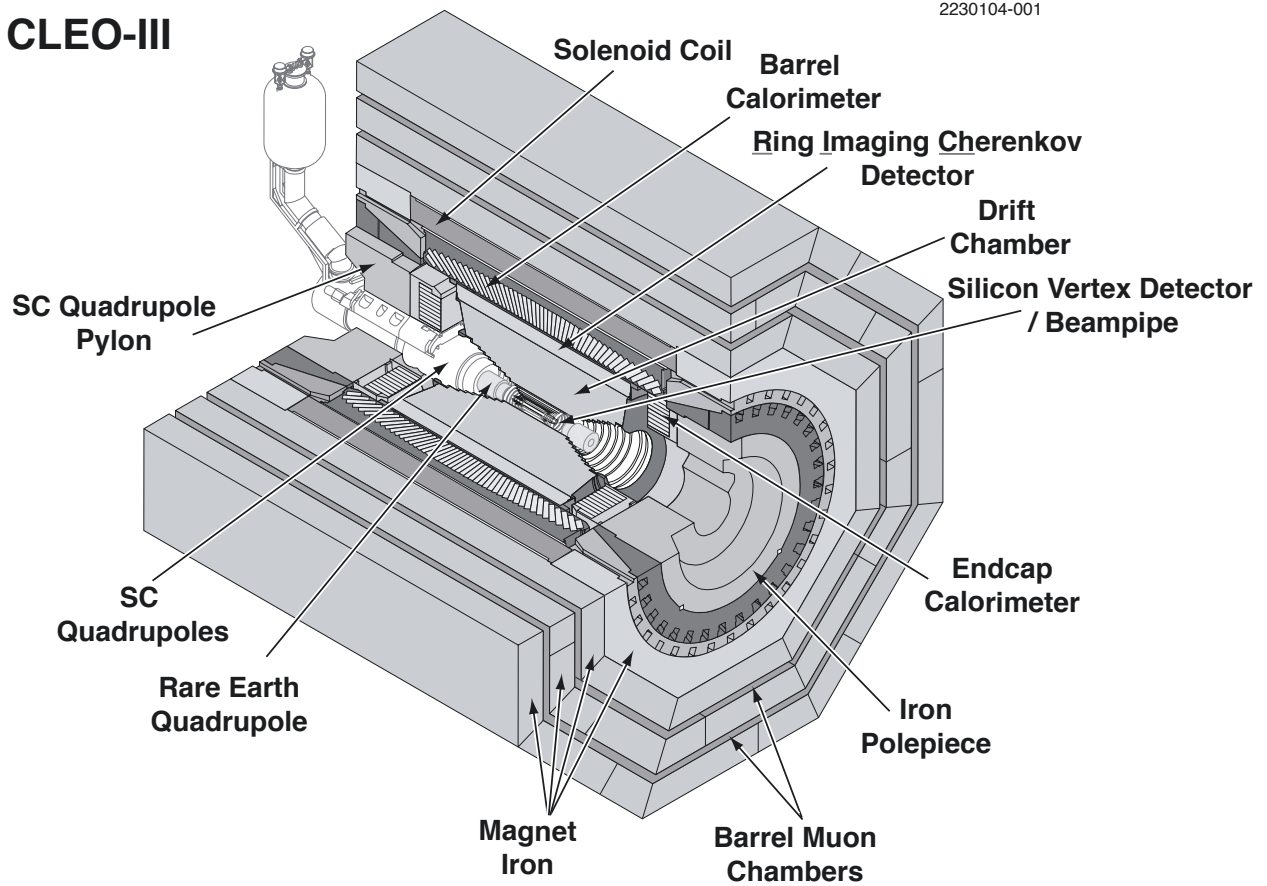


Figure 9: The CLEO III detector, which was in operation from 2000 until 2003.

Table 7: CLEO's upgrades. The abbreviations are used by the CLEO collaboration.

VERSION	DATES	COMPONENTS
I	1979 – 1986	inner proportional chamber (IZ) main drift chamber (DR) outer drift chamber (OZ) gas Cherenkov detector (CV) $\frac{dE}{dx}$ proportional chambers (DX) time-of-flight scintillators (TF) shower detectors (RS, ES, CS) muon drift chambers (MU)
I.V	1986 – 1988	straw tube drift chamber (IV) inner drift chamber (VD, replaced IZ in 1984) new main drift chamber (DR2)
II	1989 – 1995	new straw tube drift chamber (PT, replaced IV) inner drift chamber (VD) main drift chamber (DR2) new time-of-flight scintillators (TF) crystal calorimeter (CC) muon proportional chambers (MU)
II.V	1995 – 1999	silicon vertex detector (SVX, replaced PT)
III	2000 – 2003	new vertex detector (SVD, replaced VD and SVX) new main drift chamber (DR3) ring-imaging Cherenkov detector (RICH) crystal calorimeter (CC) muon proportional chambers (MU)
c	2003 – 2008	stereo drift chamber (ZD, replaced SVD)

VII. PARTICLE DETECTION AND RECONSTRUCTION

The number of individual detection elements in CLEO is of the order of 10^6 . When CESR is running and energetic beams of electrons and positrons collide in the heart of CLEO, the subatomic particles which result travel outward from the interaction point and through the detector. These particles may cause charge to be deposited on the sense wires of the drift chamber. One of them may knock an electron loose from a cesium ion in the calorimeter, which in turn will knock more electrons loose, creating a shower of secondary electrons. The particles leaving the interaction point may create Cherenkov photons in the RICH; they might induce voltages on the vertex detector or deep in the muon system. Whenever and however a particle interacts with an individual detection element it is said to leave a hit. All of CLEO's detection elements work by converting hits into measurable voltages¹. Reading out, storing, and processing these data presents a formidable task.

The first step is to decide that an event worth recording is taking place. This is done by programmable circuitry called the global level-one (L1) trigger. This circuitry uses pattern-recognition algorithms on hits coming from the drift chamber and calorimeter to decide whether or not an event is taking place. It looks for hits in the drift chamber aligned into possible tracks, and it looks at the location of hits in the calorimeter as well as the amount of energy deposited by the resulting showers. The specific trigger criteria used by CLEO III are shown in Table 8. The numbers in the table refer to tracks in the drift chamber, which can be in the axial or stereo portions, or showers in the calorimeter barrel; the amount of energy deposited by the showers is indicated by low, medium, or high. The L1 trigger issues a yes-no decision every 42 ns.

Once an event has been accepted by the L1 trigger, the individual detector channels need

¹This implies that calibrating CLEO is no simple job.

Table 8: The criteria used by the CLEO III global level-one trigger.

NAME	DEFINITION
Hadronic	$N_{\text{axial}} > 1$ and $N_{\text{barrel, low}} > 0$
μ Pair	two collinear (back-to-back) stereo tracks
Barrel Bhabha	two collinear high-energy showers in the CC barrel
Endcap Bhabha	two collinear high-energy showers in the CC end caps
Electron + Track	$N_{\text{axial}} > 0$ and $N_{\text{barrel, med}} > 0$
τ / Radiative	$N_{\text{stereo}} > 1$ and $N_{\text{barrel, low}} > 0$
Two Track	$N_{\text{axial}} > 1$
Random	random 1 kHz source

to be digitized and stored for analysis. When CESR was running at the $\Upsilon(4S)$ resonance, approximately 50% of the events accepted by the L1 trigger were beam-gas and beam-wall interactions: unwanted interactions with residual gas in the beam pipe or with the material of the beam pipe itself. Such events are clearly unsuitable for physics analysis, and a second “trigger”, the level-three (L3) decision², was used to discard them before writing the data to disk. The L3 decision was implemented in software and looked at a subset of the digitized hits to determine whether or not an event was worth saving for analysis. Telltale signatures of these unwanted events include energetic bremsstrahlung photons³ or lots of hadrons coming from the beam pipe.

Next the raw detector hits recorded for an event need to be converted into meaningful measurements. Particle trajectories are derived primarily from hits in the drift chamber and silicon vertex detector, but the location of calorimeter showers and hits in the muon detector also provide useful positional information. Particle energies are measured by the calorimeter, and momenta come from the curvature of the particles’ trajectories. The RICH and the muon detector aid in particle identification: the RICH directly measures a particle’s

²Earlier versions of CLEO had a level-two trigger, but CLEO III didn’t use it.

³The word is German for “braking radiation”. It refers to the radiation emitted by a charged particle when it decelerates in matter.

Table 9: The performance of the CLEO III data acquisition system.

PARAMETER	PERFORMANCE
maximum read-out rate	150 Hz (data) 500 Hz (random trigger)
average event size	25 kB
average read-out time	30 μ s
transfer bandwidth	6 MB/s
L3 throughput	130 Hz

velocity, which allows mass to be determined from an independent measurement of the particle’s momentum, and the muon system identifies probable muons⁴.

The process of converting hits into measurements is called event reconstruction, and CLEO’s event reconstruction happened in two stages. The first stage, accomplished by software called `pass1`, happened immediately after the conclusion of a data-taking run⁵. The `pass1` software analyzed only a fraction of the hits recorded in order to quickly provide information for quality-control and calibration purposes⁶.

The second stage of event reconstruction was done by software called, aptly enough, `pass2`. The `pass2` software analyzed all hits recorded for an event, and was therefore much slower⁷. It also made one final decision about whether or not to keep an event. This “trigger”, the level-four (L4) decision, looked at the fully reconstructed event and classified it based on the number of reconstructed tracks and the number, location, and energy of calorimeter showers. If an event had fewer than two tracks coming from the origin, L4 discarded it.

⁴Any particle which enters the muon detector has a high probability of being a muon.

⁵There was also a version called online `pass1` that processed a subset of the recorded events for the event display, which was used for real-time monitoring during a run.

⁶The output of `pass1` was typically ready in a few hours’ time.

⁷All of the runs in a data-taking period constitute a data set, which would be processed by `pass2` at the end of the period. The number of individual runs in a data set was typically of the order of 10^3 , and running `pass2` over an entire data set took about a month.

Table 10: The CLEO III data sets used in our analysis.

NAME	RUNS	EVENTS	RESONANCE	DATES
data5	257	37,214,027	$\Upsilon(4S)$	8/24/00 – 9/11/00
data6	536	54,509,053	$\Upsilon(4S)$	10/12/00 – 11/19/00
data7	499	69,289,552	$\Upsilon(4S)$	11/20/00 – 12/18/00
data8	473	77,059,984	$\Upsilon(4S)$	12/19/00 – 1/16/01
data9	553	75,156,478	$\Upsilon(4S)$	1/17/01 – 2/6/01
data10	636	82,956,536	$\Upsilon(4S)$	2/7/01 – 3/5/01
data11	528	76,638,319	$\Upsilon(4S)$	3/13/01 – 4/3/01
data12	565	87,789,019	$\Upsilon(4S)$	4/4/01 – 5/1/01
data13	552	85,660,104	$\Upsilon(4S)$	5/2/01 – 5/29/01
data14	623	113,418,172	$\Upsilon(4S)$	5/30/01 – 6/26/01
data16	452	49,845,830	$\Upsilon(3S)$	11/8/01 – 12/11/01
data17	568	74,314,425	$\Upsilon(3S), \Upsilon(1S)$	12/12/01 – 1/21/02
data18	596	89,036,857	$\Upsilon(1S)$	1/22/02 – 3/12/02
data19	432	60,807,413	$\Upsilon(1S)$	3/14/02 – 4/15/02
data21	480	64,841,127	$\Upsilon(2S)$	5/22/02 – 7/9/02
data22	376	38,622,158	$\Upsilon(3S), \Upsilon(2S)$	7/10/02 – 8/12/02
data23	207	27,626,508	$\Upsilon(2S)$	9/14/02 – 10/7/02
data25	213	31,746,192	$\Upsilon(2S)$	10/23/02 – 11/5/02
data27	218	33,241,389	$\Upsilon(2S)$	11/19/02 – 12/3/02

VIII. THE SAMPLES

Scattering experiments are frequently characterized by their luminosities. In this context luminosity is the number of incident particles per unit area per unit time. The luminosity of an accelerator is measured by observing an interaction that has a well known scattering cross section¹, *e.g.* Bhabha scattering ($e^- + e^+ \rightarrow e^- + e^+$). Measuring the luminosity then becomes a matter of measuring the rate at which Bhabha scattering occurs:

$$L = \frac{1}{\sigma} \frac{dN}{dt}$$

where σ is the cross section for Bhabha scattering.

The time-integrated luminosity

$$\mathcal{L} = \frac{N}{\sigma}$$

is used to quantify the size of a data sample. Here N is the total number of interactions of a particular kind in the sample, and σ is the cross section for the interaction.

This analysis uses a subset of the data collected by the CLEO III detector between August 24, 2000 and December 3, 2002. Our data samples contain 20.8 million $\Upsilon(1S)$ resonant decays, 9.3 million $\Upsilon(2S)$ decays, and 5.9 million $\Upsilon(3S)$ decays [7, 4]. The integrated luminosities of these samples are 1.1 fb^{-1} , 1.3 fb^{-1} , and 1.4 fb^{-1} respectively².

We also use the CLEO III $\Upsilon(4S)$ sample, which has an integrated luminosity of 6.4 fb^{-1} , as well as a 2.3-fb^{-1} sample collected 60 MeV below the $\Upsilon(4S)$ resonance³. The $\Upsilon(4S)$

¹The cross section is used to express the probability of an interaction. The name comes from classical scattering theory where it represents the cross sectional area of the target normal to the trajectory of the incident particle.

²Since cross sections are areas they should be expressed in units of m^2 ; integrated luminosities should therefore be in units of m^{-2} . However, these units are inconveniently large, so the barn ($1 \text{ b} = 10^{-28} \text{ m}^2$) and its submultiples (*e.g.* mb, nb, fb) are used instead.

³From the so called hadronic continuum.

Table 11: The samples used in this analysis.

DATA	BEAM ENERGY	\mathcal{L}	N	RESONANCE
signal data	4,730 MeV	1.1 fb ⁻¹	20.8 × 10 ⁶	$\Upsilon(1S)$
signal data	5,010 MeV	1.3 fb ⁻¹	9.3 × 10 ⁶	$\Upsilon(2S)$
signal data	5,180 MeV	1.4 fb ⁻¹	5.9 × 10 ⁶	$\Upsilon(3S)$
background data	5,290 MeV	6.4 fb ⁻¹	6.0 × 10 ⁶	$\Upsilon(4S)$
background data	5,260 MeV	2.3 fb ⁻¹		continuum
MONTE CARLO	BEAM ENERGY	\mathcal{L}	N	SIMULATED DECAY
background MC	5,290 MeV	23.3 fb ⁻¹	21.3 × 10 ⁶	$\Upsilon(4S) \rightarrow \tau^\mp + \tau^\pm$
background MC	5,290 MeV	1.1 fb ⁻¹	1.2 × 10 ⁶	$\Upsilon(4S) \rightarrow \mu^\mp + \mu^\pm$
signal MC	4,730 MeV		30.0 × 10 ³	$\Upsilon(1S) \rightarrow \mu^\mp + \tau^\pm$
signal MC	5,010 MeV		30.0 × 10 ³	$\Upsilon(2S) \rightarrow \mu^\mp + \tau^\pm$
signal MC	5,180 MeV		21.8 × 10 ³	$\Upsilon(3S) \rightarrow \mu^\mp + \tau^\pm$

and continuum data are used to quantify background processes, to measure detection and reconstruction efficiencies, and to estimate systematic errors. Being able to use actual data for these purposes instead of having to rely solely on Monte Carlo simulation is a noteworthy aspect of this analysis.

In addition to the measured data, we do in fact use a number of Monte Carlo samples. To study the background arising from the decay of Υ mesons into pairs of charged leptons we use a Monte Carlo sample of 21.3 million τ pairs with an integrated luminosity of 23.3 fb⁻¹ and a Monte Carlo sample of 1.2 million muon pairs with an integrated luminosity of 1.1 fb⁻¹. These samples have a beam energy⁴ corresponding to our $\Upsilon(4S)$ data. We also use three Monte Carlo samples to study the detection signatures of signal events. These samples consist of 30,000 lepton flavor violating decays of the $\Upsilon(1S)$, 30,000 of the $\Upsilon(2S)$, and 21,780 of the $\Upsilon(3S)$. Their integrated luminosities, of course, cannot be specified.

⁴The beam energy is the energy per particle of the incident electrons and positrons.

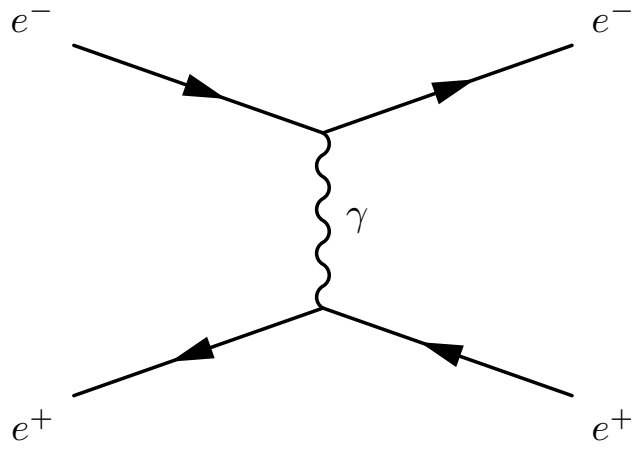


Figure 10: One of the leading-order contributions to Bhabha scattering.

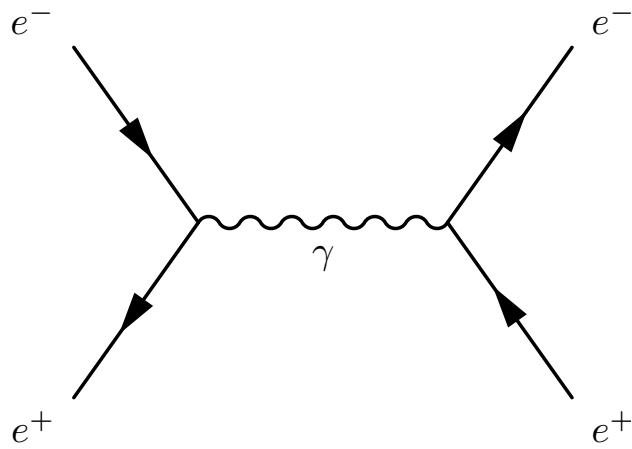


Figure 11: The other leading-order contribution to Bhabha scattering.

IX. SELECTION CRITERIA

The detection signature of our signal events is a muon¹ whose momentum is very close to the beam energy, E_b . A straightforward kinematic calculation shows that these signal muons should have a momentum of $0.965E_b$ at the $\Upsilon(1S)$ resonance, $0.968E_b$ at the $\Upsilon(2S)$, and $0.970E_b$ at the $\Upsilon(3S)$. Real data, of course, will show a signal peak that is smeared somewhat because of the finite resolution of the detector. In order to select possible signal events from our data samples a number of criteria were developed. We are looking for $\Upsilon \rightarrow \mu^\mp + \tau^\pm$, so we select events that have exactly two reconstructed charged particle tracks of opposite curvature², corresponding to the muon and the τ . Furthermore, we identify the τ through its decay to an electron: $\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau$, making this an exclusive analysis³. For identifying signal events the quantities of interest are $x = p_\mu/E_b$ and $y = p_e/E_b$, the momenta, normalized to beam energy, of the muon and electron candidates respectively.

Electron and muon identification criteria are applied to the two tracks comprising an event. First, we require that exactly one of the tracks be tagged as a high-quality muon candidate, which means that it penetrates the muon detector to at least the second plane of chambers, which is at a depth of 0.72 m in the magnet iron, leaving hits on at least two of the three layers counters comprising the planes at and below the penetration depth⁴. To identify the electron candidate we look at the ratio E/p for the second track. This is the ratio of the energy deposited in the calorimeter along the track's projection to the momentum measured from the track's curvature. Electrons will lose a lot of energy in the calorimeter,

¹Or an antimuon; charge-conjugation is implied throughout this chapter.

²Therefore, they have opposite electric charge.

³Since $\mathcal{B}(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau) = 18\%$, we are *excluding* 82% of our sample from the outset.

⁴In other words, the muon candidate has to go deep into the magnet iron and leave lots of hits. Any particle that can do this is almost certainly a muon.

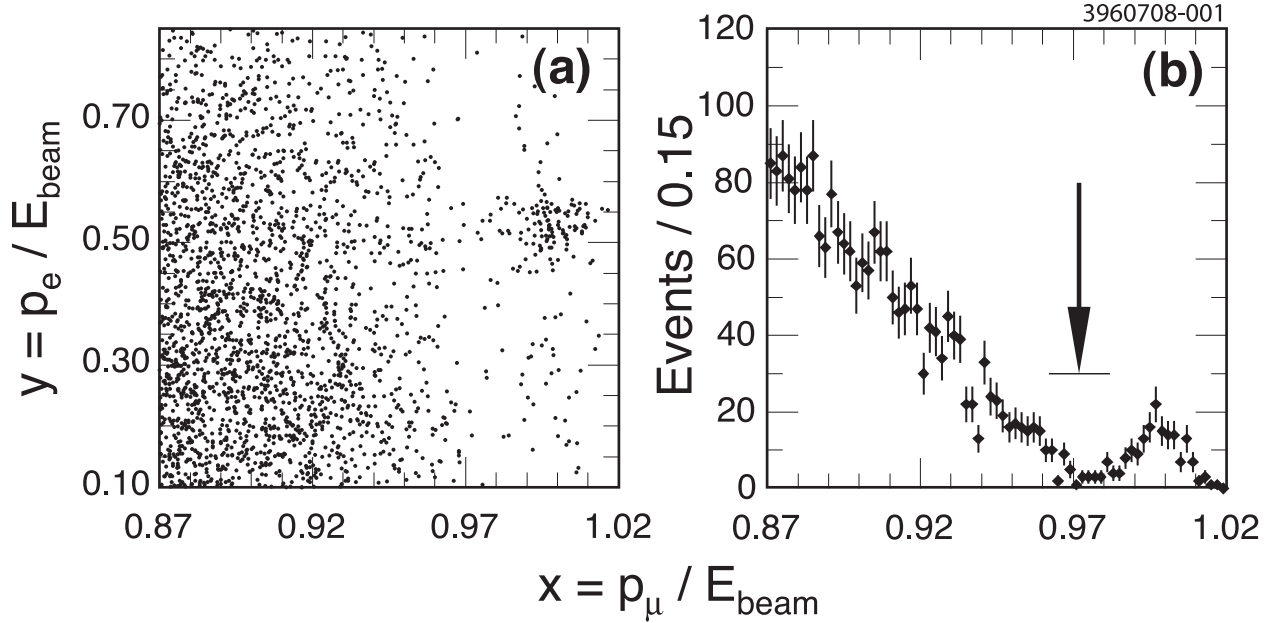


Figure 12: (a) The scatter plot of y vs. x ; (b) its binned x projection for calibration data.

so requiring $0.85 < E/p < 1.10$ for the second track helps ensure that it is, in fact, due to an electron. Additionally we require that the electron candidate penetrate no deeper into the muon detector than the first plane of chambers, which is at 0.36 m. Real electrons rarely make it deep into the muon detector, so this requirement helps reduce the possibility of our electron candidate being a misidentified muon. The result of these criteria is that one of our tracks should now be very likely caused by a muon and probably not by an electron, while the other should be very likely caused by an electron and probably not by a muon.

Now, because of the logistics of data storage *circa* A.D. 2000, there is another set of selection criteria that needs to be mentioned. Events measured by CLEO III were sorted into various “subcollections” based on criteria tailored to different physics analyses. For example, the τ subcollection contains events likely to be electron-positron annihilation to τ pairs. It is intended for use by researchers studying τ physics at the Υ resonances; this analysis uses the τ subcollection. To be included in this subcollection an event must satisfy at least one of several sets of criteria. The relevant set for our analysis includes the following

requirements:

- The event must have at least two reconstructed tracks.
- The 2nd highest track momentum in the event must be less than 85% of the beam energy.
- The total energy detected for the event must be greater than 20% of the center-of-mass energy⁵.
- Total energy detected in the calorimeter alone for an event must be less than 85% of the center-of-mass energy.

The first of these subcollection requirements is consistent with our two-track criterion. To meet the second we require $0.10 < y < 0.85$, and for the third we require $0.87 < x < 1.02$. The final subcollection requirement is met when we select events with a deeply penetrating muon candidate.

No selection criteria are perfect, and the ones described so far select more than just signal events. A certain number of background events — events which mimic the signal in some way — are tagged by our selection criteria. The biggest contribution to our background events comes from electron-positron annihilation to τ pairs: $e^- + e^+ \rightarrow \tau^- + \tau^+$, where afterward one of the τ leptons decays into a muon and the other decays into an electron: $\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau$ while $\tau^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\tau$. Since τ leptons are extremely short-lived⁶, they will decay close to the interaction point, and the event will consist of a muon, a τ lepton, and some neutrinos. Therefore, events such as the one just described will be misidentified as signal by our selection criteria. Muon pairs: $e^- + e^+ \rightarrow \mu^- + \mu^+$, where afterward one of the muons decays into an electron⁷, also contribute to the background.

There is one other source of background events which requires some explanation. Suppose a pair of muons is produced with an energetic photon: $e^- + e^+ \rightarrow \mu^- + \mu^+ + \gamma$. Such events satisfy our E/p criterion when the photon is detected in the calorimeter along the projection of one of the reconstructed muon tracks. If the photon carries off enough of the event's energy the muon whose track is intersected by the photon will fail to enter the muon detector, and it will be misidentified as an electron candidate. Because of the requirement on E/p , these

⁵This is twice the beam energy in a symmetric collider such as CESR: $E_{\text{cm}} = 2E_b$.

⁶They have a mean lifetime of 290.6 fs.

⁷This is less likely because the mean lifetime of a muon is 2.197 μs ; most muons make it out of the detector before decaying.

Table 12: Processes contributing to our background (charge-conjugation is implied).

LABEL	BACKGROUND	DESCRIPTION
$\tau\tau$	τ lepton pairs	$e^- + e^+ \rightarrow \tau^- + \tau^+$ $\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau$ $\tau^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\tau$
$\mu\mu(e)$	muon pairs with decay	$e^- + e^+ \rightarrow \mu^- + \mu^+$ $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$
$\mu\mu(\gamma)$	muon pairs with radiation	$e^- + e^+ \rightarrow \mu^- + \mu^+$ $\mu^- \rightarrow \mu^- + \gamma$

events cluster near $y = 0.53$. For such events E is the energy of the photon, which is approximately half of the beam energy, combined with nearly 0.3 GeV deposited in the calorimeter by the misidentified muon.

Eventually, attempting to suppress the background events with additional selection criteria reaches a point of diminishing returns and a more sophisticated approach is required. This is where maximum-likelihood fitting comes in. To use this method we look for measured quantities whose distributions for signal events are significantly different from their distributions for backgrounds. For example, at the $\Upsilon(1S)$ resonance the momentum distribution of our signal muons is sharply peaked around $0.965E_b$, while the momentum distribution of the muons in our τ pairs background falls off linearly and becomes zero for momenta above $0.963E_b$. Therefore, we can use x as one of the variables in a multi-dimensional likelihood fit. In fact, in terms of separating signal events from background events, x is the most important variable in this analysis. Our exclusive analysis uses the four fitting variables summarized in Table 13. The ionization energy loss of the electron candidate dE/dx has not yet been discussed. This quantity is a measure of consistency with the theoretically predicted value of dE/dx for real electrons and is interpreted according to the hypothesis that the electron candidate really is an electron. It is expressed in units of standard deviation. The limits we impose on the fitting variables constitute additional selection criteria.

Table 13: The variables used in our multi-dimensional maximum-likelihood fits.

VARIABLE	DESCRIPTION	RANGE
$x = p_\mu/E_b$	normalized momentum of muon candidate	$0.87 < x < 1.02$
$y = p_e/E_b$	normalized momentum of electron candidate	$0.10 < y < 0.85$
E/p	ratio of electron's energy to its momentum	$0.85 < E/p < 1.10$
dE/dx	electron's specific ionization energy loss	$-3.0\sigma < dE/dx < 3.0\sigma$

Once selection criteria are chosen, their efficiency at selecting signal events must be determined. We do this by applying the criteria to signal Monte Carlo samples. The results of this process are tabulated in Tables 14, 15, and 16. Note that the CLEO III trigger is effectively a set of selection criteria which has high, but not perfect, efficiency, and that the overall efficiency of our selection criteria is about 9%. This is mostly due to the fact that this is an exclusive analysis, *i.e.* the selection criteria include $\mathcal{B}(\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau)$. Finally, the effect of applying all selection criteria to the $\Upsilon(4S)$ data is shown in Figure 12, where the location of the hypothetical signal peak is indicated by the arrow, where the width of the horizontal bar at its tip is $\pm\sigma(x)$. The data in these plots are the background events whose distributions in our fit variables are needed for maximum-likelihood fitting.

Table 14: Selection efficiencies at the $\Upsilon(1S)$, determined from signal Monte Carlo.

CRITERION	BEFORE	AFTER	EFFICIENCY	CUMULATIVE
two tracks	30,000	21,174	70.6%	70.6%
one muon	21,174	15,021	70.9%	50.1%
one electron	15,021	3,337	22.2%	11.1%
opposite charges	3,337	3,335	99.9%	11.1%
L1 trigger	3,335	3,129	93.8%	10.4%
L3 trigger	3,129	3,129	100.0%	10.4%
L4 trigger	3,129	3,118	99.6%	10.4%
τ subcollection	3,118	3,040	97.5%	10.1%
e is not μ	3,040	3,039	100.0%	10.1%
$0.10 < y < 0.85$	3,039	2,793	91.9%	9.3%
$0.87 < x < 1.02$	2,793	2,728	97.7%	9.1%
$-3.0 < dE/dx < 3.0$	2,728	2,650	97.1%	8.8%

Table 15: Selection efficiencies at the $\Upsilon(2S)$, determined from signal Monte Carlo.

CRITERION	BEFORE	AFTER	EFFICIENCY	CUMULATIVE
two tracks	30,000	21,250	70.8%	70.8%
one muon	21,250	15,059	70.9%	50.2%
one electron	15,059	3,403	22.6%	11.3%
opposite charges	3,403	3,403	100.0%	11.3%
L1 trigger	3,403	3,209	94.3%	10.7%
L3 trigger	3,209	3,209	100.0%	10.7%
L4 trigger	3,209	3,199	99.7%	10.7%
τ subcollection	3,199	3,110	97.2%	10.4%
e is not μ	3,110	3,109	100.0%	10.4%
$0.10 < y < 0.85$	3,109	2,846	91.5%	9.5%
$0.87 < x < 1.02$	2,846	2,786	97.9%	9.3%
$-3.0 < dE/dx < 3.0$	2,786	2,704	97.1%	9.0%

Table 16: Selection efficiencies at the $\Upsilon(3S)$, determined from signal Monte Carlo.

CRITERION	BEFORE	AFTER	EFFICIENCY	CUMULATIVE
two tracks	21,780	15,417	70.8%	70.8%
one muon	15,417	10,952	71.0%	50.3%
one electron	10,952	2,458	22.4%	11.3%
opposite charges	2,458	2,457	100.0%	11.3%
L1 trigger	2,457	2,313	94.1%	10.6%
L3 trigger	2,313	2,313	100.0%	10.6%
L4 trigger	2,313	2,309	99.8%	10.6%
τ subcollection	2,309	2,247	97.3%	10.3%
e is not μ	2,247	2,247	100.0%	10.3%
$0.10 < y < 0.85$	2,247	2,051	91.3%	9.4%
$0.87 < x < 1.02$	2,051	2,010	98.0%	9.2%
$-3.0 < dE/dx < 3.0$	2,010	1,954	97.2%	9.0%

X. THE METHOD OF MAXIMUM LIKELIHOOD

To estimate the numbers of signal events in our samples after selection criteria have been applied, we use V. Savinov’s software implementation of the multidimensional unbinned extended maximum likelihood (EML) method. This software is called FELIX.

The EML method was first proposed by Fermi and is described in R. Barlow’s excellent book [5]. It takes as its input a sample of N events and a set functions $\mathcal{P}_l(\{z\}, \{\alpha\}_l)$ describing the probability densities of the events. Each contribution to the sample needs to be described by a corresponding function \mathcal{P} . These functions depend on the set of fitting variables $\{z\}$ and a set of parameters $\{\alpha\}$ describing the multidimensional shape of \mathcal{P} . Our analysis uses the four fitting variables described previously, so in the notation given above $\{z\} = \{x, y, E/p, dE/dx\}$. Also, we parameterize each variable separately, so

$$\mathcal{P}_l(\{z\}, \{\alpha\}_l) = \mathcal{P}_l(x, \{\xi\}_l)\mathcal{P}_l(y, \{\eta\}_l)\mathcal{P}_l(E/p, \{\zeta\}_l)\mathcal{P}_l(dE/dx, \{\omega\}_l)$$

Here \mathcal{P}_l represents the overall probability density function of the l th contribution to the sample of events; $\{\xi\}_l$ is the set of parameters describing the x -shape of the l th contribution, with the parameter sets $\{\eta\}_l$, $\{\zeta\}_l$, and $\{\omega\}_l$ similarly describing the shape of the l th contribution in y , E/p , and dE/dx respectively. There are four components contributing to our event samples: the three background components summarized in Table 12, and a possible signal component. For each of these four contributions, four distributions need to be parameterized: x , y , E/p , and dE/dx . The parameters of all sixteen of the probability density functions used in fitting are measured from on-resonance $\Upsilon(4S)$ data¹. The off-resonance $\Upsilon(4S)$ data, taken 60 MeV below the resonance, are used as a control sample.

¹Even signal, since y , E/p , and dE/dx are going to have the same distributions as our background components with real electrons, and the width of the peak in x is just the momentum resolution of the detector. The location of the peak, remember, comes from the kinematics of the interaction.

Table 17: The shapes used in our maximum-likelihood fits.

	τ PAIRS	μ PAIRS (DEC.)	μ PAIRS (RAD.)	SIGNAL
x	1 st order poly.	Crystal Ball	Crystal Ball	Gaussian
y	2 nd order poly.	2 nd order poly.	Crystal Ball	2 nd order poly.
E/p	Crystal Ball	Crystal Ball	1 st order poly.	Crystal Ball
dE/dx	Gaussian	Gaussian	Gaussian	Gaussian

Several of our distributions are described by the ‘‘Crystal Ball’’ function [11]:

$$\mathcal{C}(z, \{A, \bar{z}, \sigma, \alpha, \beta\}) = \begin{cases} Ae^{-(\bar{z}-z)^2/2\sigma^2} & \text{if } z > \bar{z} - \alpha\sigma \\ A \frac{(\frac{\beta}{\alpha})^\beta e^{-\alpha^2/2}}{(\frac{\bar{z}-z}{\sigma} - \frac{\beta}{\alpha} - \alpha)^\beta} & \text{if } z \leq \bar{z} - \alpha\sigma \end{cases}$$

This function matches a power-law tail to a Gaussian below the mean of the Gaussian. It is parameterized by normalization A , Gaussian mean \bar{z} , Gaussian width σ , parameter α , which determines where the function starts to diverge from the Gaussian, and parameter β , which determines the length of the tail.

Once we have our fitting shapes we use them to construct the likelihood function

$$L = \frac{1}{N!} \exp\left(-\sum_j N_j\right) \prod_i \sum_j N_j \mathcal{P}_j(\{z\}_i, \{a\}_j)$$

where N is the total number of events in our sample and N_j is the number of events in the j th component. The EML method finds the set of values N_j that maximize L and does *not* require that the sum of N_j over all components be equal to N .

We proceed as follows:

- First, we measure the parameters of our fitting shapes by performing binned maximum-likelihood fits to the on-resonance $\Upsilon(4S)$ data using the MN_FIT and MINUIT software packages.

- Then we perform an unbinned extended maximum-likelihood fit to the on-resonance $\Upsilon(4S)$ data with **FELIX**, using the parameters obtained in the previous step. We include signal shapes to make sure we have not introduced a bias.
- Next we fit the off-resonance $\Upsilon(4S)$ control sample with **FELIX**, first without and then with the signal component. We add signal Monte Carlo events to the control sample to make sure that **FELIX** can find the correct number of signal events.
- Finally, we fit the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ data samples with **FELIX**.

Two complications to this straightforward procedure need to be explained. First, the distribution of x for the τ pairs background needs to be “smeared” by convolving it with a Gaussian function describing the finite momentum resolution of the detector. The x -shape of the τ pairs background decreases linearly and becomes zero above a certain value:

$$\mathcal{P}_{\tau\tau}(x, \{A_{\tau\tau}, X\}) = A_{\tau\tau}(1 - x/X)$$

Here $A_{\tau\tau}$ is the normalization parameter and X is the value of x above which $\mathcal{P}_{\tau\tau}(x)$ is zero. The parameter X is determined by the kinematics of the τ lepton’s decay into a muon and has the values $X = 0.963$ at the $\Upsilon(1S)$ resonance, $X = 0.967$ at the $\Upsilon(2S)$ resonance, and $X = 0.970$ at the $\Upsilon(3S)$ resonance. These values fix the horizontal intercept of the distribution. However, because of the finite resolution of the detector, there is a small chance of detecting τ pairs events with values of x above this kinematic boundary. The convolution mentioned above allows us to take this into account. We use our background of muon pairs to measure the detector’s momentum resolution, which is just the width of the Gaussian part of the Crystal Ball function, and find that it is 0.86%. This value is also used as the width of the signal peak.

The other complication is that y and E/p are strongly correlated for the muon pairs with radiation. Figure 13 shows the distribution of E/p vs. y for radiative muon pairs in our calibration sample. The distribution of y for this component is described by a Crystal Ball function $\mathcal{C}(y)$ whose mean depends on E/p . To deal with this we introduce the following two-dimensional probability density, where for clarity we define $u = E/p$:

$$\mathcal{P}_{\mu\mu(\gamma)}(y, u) = A_{\mu\mu(\gamma)}h(y, u)\mathcal{C}(y, \{1, \bar{y}(u), 0.0065, 1.0, 6.0\})$$

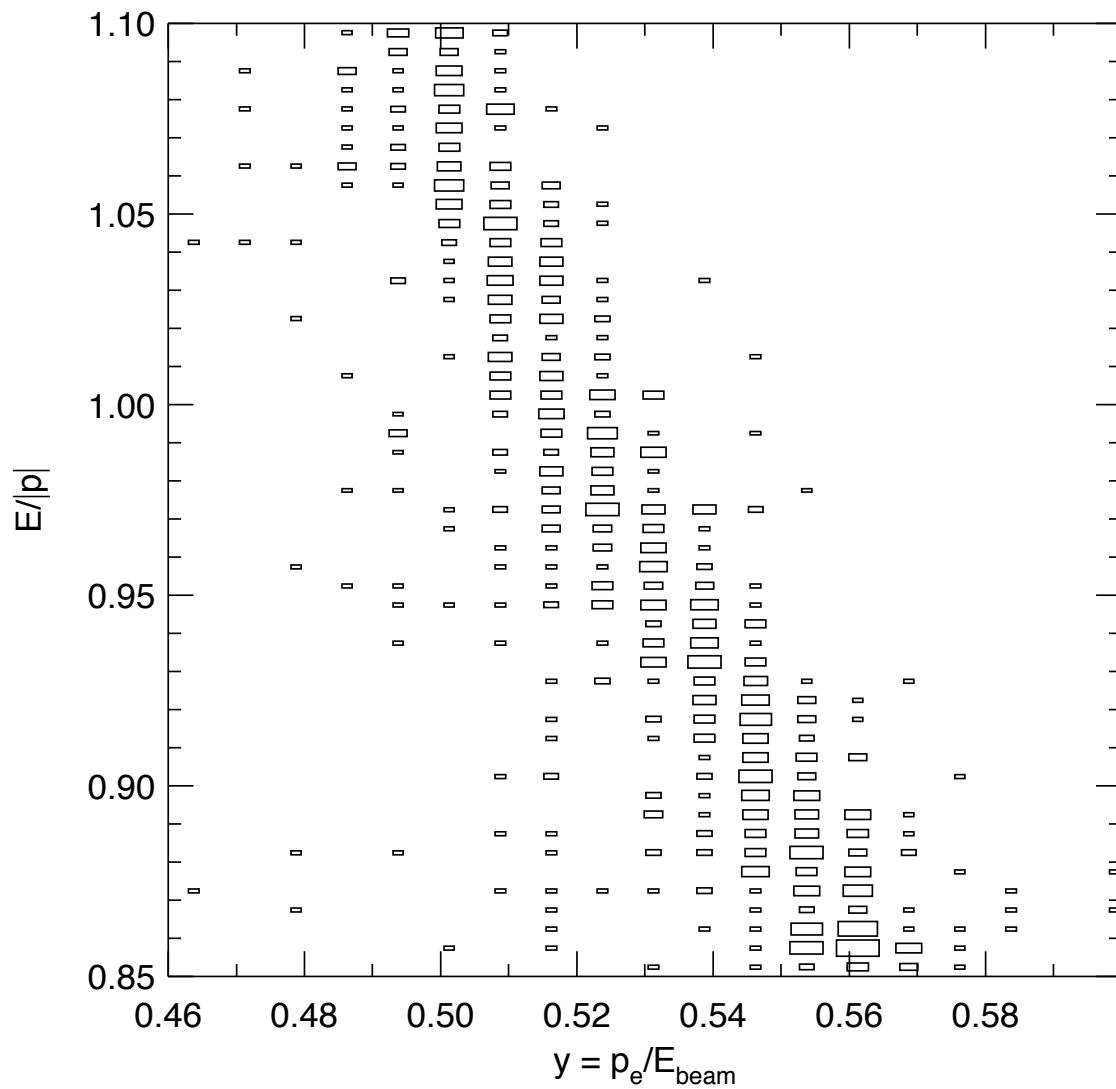


Figure 13: Correlation of E/p and y for radiative muon pairs in our calibration sample.

This has normalization parameter $A_{\mu\mu(\Upsilon)}$ and uses the auxiliary function

$$h(y, u) = 1 - 0.81338 (u \cos \theta - y \sin \theta)$$

which is a linear function of the rotated variable $u' = -y \sin \theta + u \cos \theta$. Figure 13 shows that the rotation angle is $\theta = \arctan[(0.56 - 0.50)/(1.10 - 0.85)]$, or approximately 13.5° . The dependence of \bar{y} on u is described by

$$\bar{y}(u) = 0.50 + (1.10 - u) \tan \theta$$

Using $u' = y \cos \theta + u \sin \theta$ to eliminate u and simplifying, we find

$$\frac{\bar{y} - y}{\sigma} = \frac{(0.50 \cos \theta + 1.10 \sin \theta) - y'}{\sigma \cos \theta}$$

which makes it clear that

$$\mathcal{C}(y, \{1, \bar{y}(u), 0.0065, 1.0, 6.0\}) = \mathcal{C}(y', \{1, \bar{y}', 0.0065 \cos \theta, 1.0, 6.0\})$$

where $\bar{y}' = 0.50 \cos \theta + 1.10 \sin \theta$. So, in terms of the rotated variables y' and u' , the probability density given above becomes

$$\mathcal{P}_{\mu\mu(\Upsilon)}(y', u') = A_{\mu\mu(\Upsilon)} h(u') \mathcal{C}(y', \{1, 0.74, 0.0063, 1.0, 6.0\}).$$

The results of our unbinned EML fits are shown in Table 18. They are consistent with $N_{\text{sig.}} = 0$ in all studied samples. Since no signal was observed, we may set an upper limit on the branching fraction $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm)$. In a perfect, error-free world this would simply be

$$\mathcal{B} < \frac{1}{N}$$

where N is the total number of Υ decays in our sample². However, there are many sources of error in this analysis, and these must now be described.

²Since a perfect detector would have a 100% detection efficiency, N would be the true number of Υ mesons produced by CESR.

Table 18: The results of our fits.

SAMPLE	COMPONENTS	$N_{\tau\tau}$	$N_{\mu\mu(e)}$	$N_{\mu\mu(\gamma)}$	$N_{\text{sig.}}$	$N_j < 0$
$\Upsilon(4S)$	3	$1,568 \pm 40$	66 ± 9	60 ± 8	–	yes
$\Upsilon(4S)$	4	$1,571 \pm 40$	70 ± 10	60 ± 8	-6 ± 5	yes
$\Upsilon(4S)$	4	$1,568 \pm 40$	66 ± 8	60 ± 8	$0 + 2$	no
continuum	3	584 ± 25	18 ± 5	21 ± 5	–	yes
continuum	4	584 ± 25	18 ± 5	21 ± 5	0 ± 4	yes
continuum	4	584 ± 25	18 ± 5	21 ± 5	$0 + 4$	no
$\Upsilon(1S)$	3	492 ± 23	27 ± 6	13 ± 4	–	yes
$\Upsilon(1S)$	4	491 ± 23	26 ± 6	13 ± 4	3 ± 4	yes
$\Upsilon(1S)$	4	491 ± 23	26 ± 6	13 ± 4	$2 + 4$	no
$\Upsilon(2S)$	3	463 ± 22	22 ± 6	14 ± 4	–	yes
$\Upsilon(2S)$	4	462 ± 22	22 ± 6	13 ± 4	1 ± 4	yes
$\Upsilon(2S)$	4	462 ± 22	22 ± 6	13 ± 4	$1 + 4$	no
$\Upsilon(3S)$	3	451 ± 22	18 ± 5	7 ± 3	–	yes
$\Upsilon(3S)$	4	450 ± 22	17 ± 5	7 ± 3	2 ± 4	yes
$\Upsilon(3S)$	4	450 ± 22	17 ± 5	7 ± 3	$2 + 4$	no

XI. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties in this analysis arise from several sources and fall into three categories. First there is an overall systematic uncertainty associated with the detection efficiency. The largest contributions to this are listed in Table 19. The most important of these is the error on the trigger's efficiency. Signal Monte Carlo studies indicate that the trigger is 94% efficient, and studies of the production cross section for τ pairs suggest that the systematic error on this efficiency is 5%. V. Savinov studied two-track selection efficiency as part of his PhD thesis [17] and found the error on this to be 4%. The track reconstruction efficiency of CLEO III is 3% for muons and 1% for electrons. The systematic error on muon identification was studied by V. Savinov and S. Nam while implementing the muon detector's response for use in Monte Carlo simulation. This was found to be 2%. The selection criteria for the τ subcollection are well understood and highly efficient; it is unlikely that the error on their efficiency would exceed 2%. The spread of the efficiencies in Tables 14, 15, and 16 indicate that the error on our signal Monte Carlo statistics is 2%. The L3 and L4 software triggers are very nearly 100% efficient, so we estimate the uncertainty on this to be 1%. Our electron identification is based on loose selection criteria on E/p and dE/dx . The uncertainty on the efficiency of these criteria would not be large, and we estimate it to be 1%. All these contributions combine to give an overall systematic error on the detection efficiency of 8%.

We verify this estimate by measuring the cross section for electron-positron annihilation to τ lepton pairs in our $\Upsilon(4S)$ calibration sample, with a detection efficiency obtained from generic τ pairs Monte Carlo samples. This measurement, shown in Figure 14, agrees with the expected 0.92 nb to within 4%, with a statistical uncertainty of 5%.

For signal events there is an uncertainty in the shape of y and in the efficiency of the

Table 19: Contributions to the systematic uncertainties on the detection efficiency.

SOURCE	UNCERTAINTY	COMMENT
trigger	5%	L1 trigger
event selection	4%	
track reconstruction	3%	two tracks
muon identification	2%	
event preselection	2%	τ subcollection
signal MC statistics	2%	not used for shapes
software trigger	1%	L3 and L4 decisions
electron identification	1%	

y -region selection criteria arising from the uncertainty in the τ lepton’s polarization, which is *a priori* unknown for the lepton flavor violating decay we seek. With respect to the unpolarized τ leptons used in this analysis two extremes are possible: the electron arising from the decay of the τ could be boosted forward, or it could be boosted backward. In the former case the y -selection efficiency is 3% lower than the unpolarized case, and in the latter it is 3% higher. We therefore use 3% as the systematic uncertainty on the y -selection efficiency.

Next is the error on the integrated luminosities of our data samples, which leads to uncertainties in the numbers of Υ mesons present in the samples. This in turn contributes to the error on the upper limit we wish to set for $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm)$. The luminosities of the CLEO III data sets have a systematic error of 2%.

Finally, there is some uncertainty in our results arising from errors on the parameters of our fitting shapes. These we take into account in our procedure for evaluating the upper limit of $\mathcal{B}(\Upsilon \rightarrow \mu^\mp + \tau^\pm)$. We perform 1,000 EML fits for each data sample using the parameters measured from $\Upsilon(4S)$ data as described previously, but we vary the parameters according to the Gaussian uncertainties in their values. This yields sets of 1,000 values each for $N_{\tau\tau}$,

$N_{\mu\mu(e)}$, $N_{\mu\mu(\gamma)}$, and $N_{\text{sig.}}$, the latter of which we now write in terms of the branching fraction:

$$N_{\text{sig.}} = \epsilon N_{\Upsilon} \mathcal{B}$$

where ϵ is the detection efficiency and N_{Υ} is the number of Υ mesons present in the sample. We then integrate the likelihood function over the three background components, leaving it as a function of \mathcal{B} , which we also vary according to the Gaussian uncertainties in ϵ (8.5%, including τ polarization error) and N_{Υ} (2%). We now have 1,000 likelihood functions which we normalize and sum.

The results of this procedure are likelihood contours, which are themselves normalized, like the one shown in Figure 15. This figure shows: (a) the binned x projection of the results of the EML fit to $\Upsilon(1S)$ data, which are the points with the error bars. The solid line indicates the result of the fit; the shaded areas show τ pair, μ pair, and signal contributions to the fit. The dashed line shows the hypothetical signal of 100 lepton flavor violating events superimposed on the results of the fit, and (b) the distribution of the likelihood function as a function of branching fraction for $\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}$. The 95% CL upper limit on \mathcal{B} corresponds to the value of \mathcal{B} that gives 95% of the area under the contour in the physical, *i.e.* nonnegative, region of \mathcal{B} . This procedure yields the following upper limits on $\mathcal{B}(\Upsilon \rightarrow \mu^{\mp} + \tau^{\pm})$ at the 95% confidence level:

$$\begin{aligned} \mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 6.0 \times 10^{-6} \\ \mathcal{B}(\Upsilon(2S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 14.4 \times 10^{-6} \\ \mathcal{B}(\Upsilon(3S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 20.3 \times 10^{-6}. \end{aligned}$$

$\chi^2 = 1.2$ for 4 - 1 d.o.f., C.L. = 75.0%

Errors		Parabolic	Minos	
Function 1: Polynomial of Order 1				
NORM	0.88057	$\pm 5.0000E-02$	- 5.0000E-02	+ 5.0000E-02
* POLY01	0.0000	± 0.000	- 0.000	+ 0.000
* OFFSET	5.0000	± 0.000	- 0.000	+ 0.000

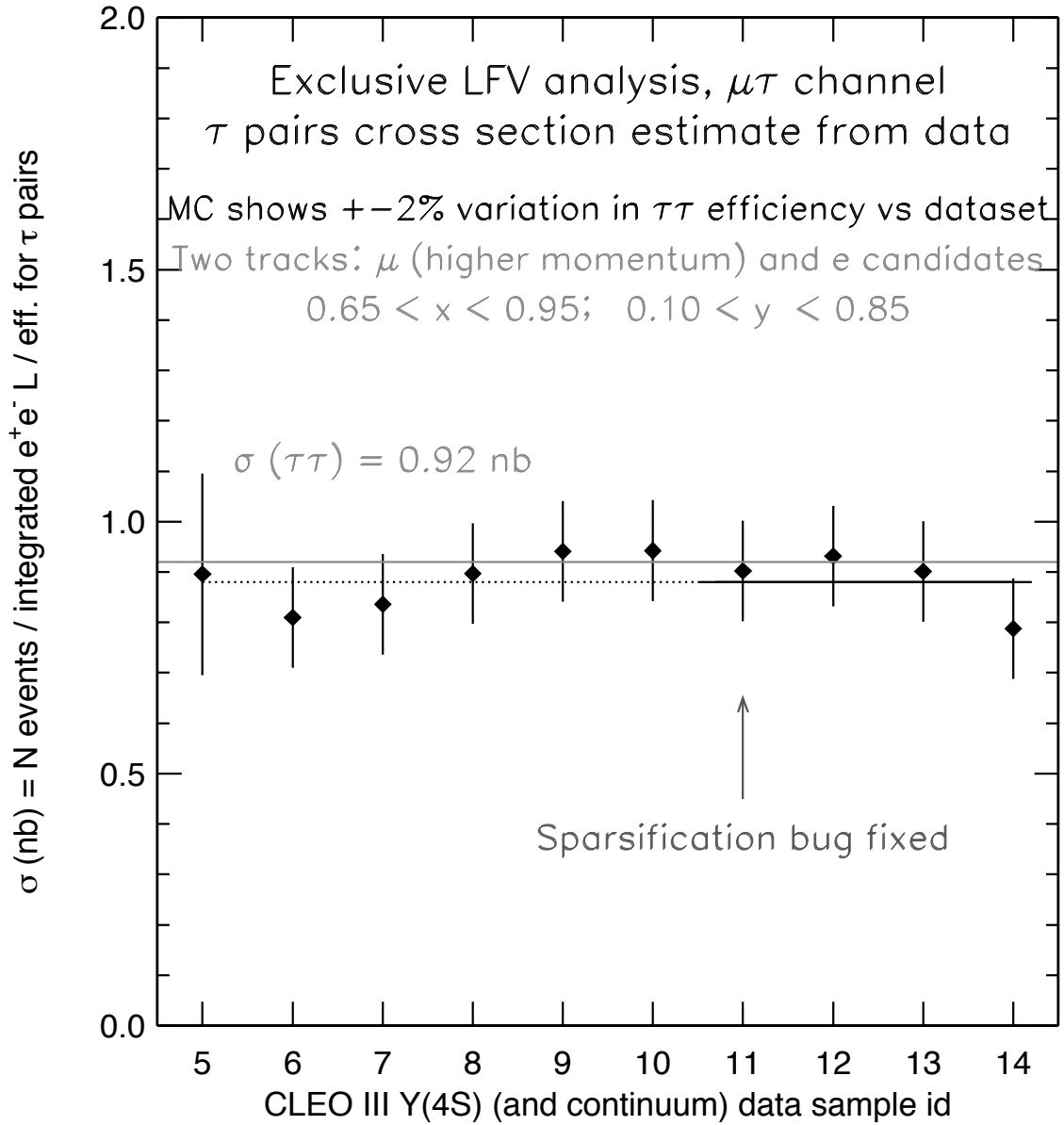


Figure 14: A measurement of the τ pairs production cross section from data.

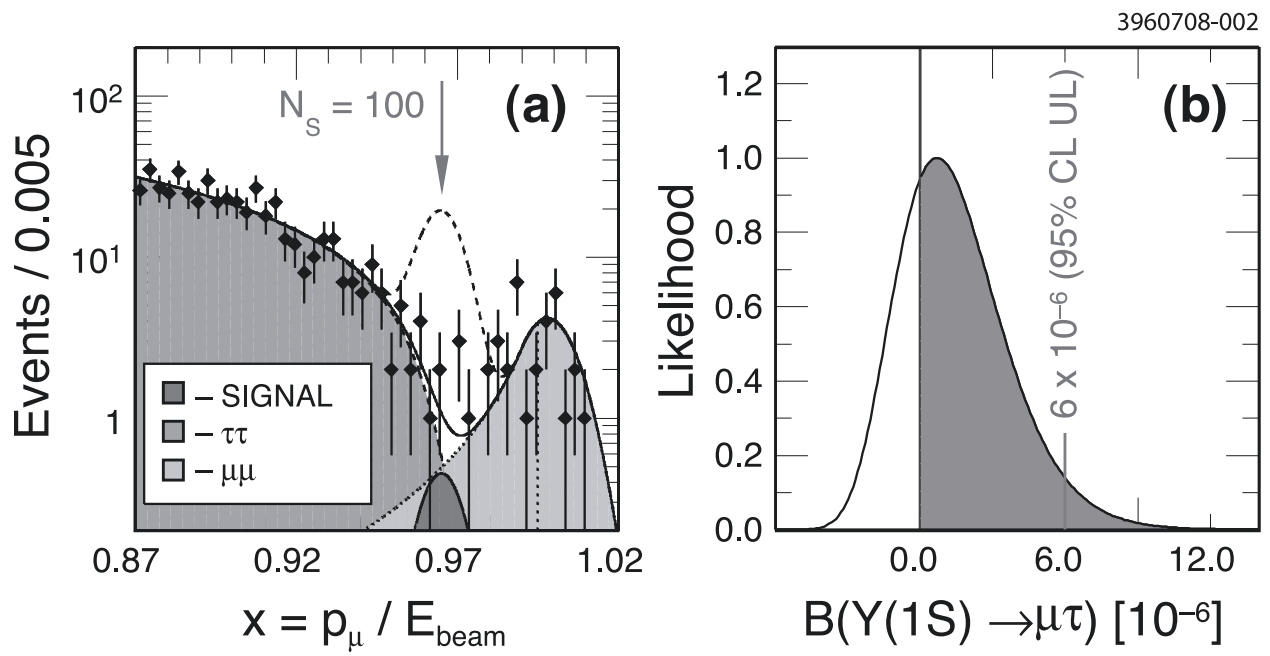


Figure 15: The results of the EML fit to $\Upsilon(1S)$ data.

XII. THE INCLUSIVE ANALYSIS

The detection efficiency of our analysis is fairly low (8.9%) in part because we identify the τ lepton in signal events through its decay to an electron, a process that occurs only 18% of the time. Therefore, an inclusive analysis, one that does not discard four out of every five signal events from the outset, was investigated to see if it would have better sensitivity.

Two criteria were used to select signal events from the data samples. First, events were required to have at least two reconstructed charged particle tracks, and second, at least one of the tracks was required to be a high-quality muon candidate as defined in Chapter 9. These criteria select all of the backgrounds that were present in the inclusive analysis as well as electron-positron annihilation to muon pairs: $e^- + e^+ \rightarrow \mu^- + \mu^+$, and muons arising from the cosmic ray bombardment of the upper atmosphere¹.

From the two or more tracks in a selected event we define the following: Track 1 is the muon candidate with the highest momentum, and Track 2 is the track with the highest momentum of the remaining tracks. Tracks 1 and 2 are then used to define the following quantities:

$$\begin{aligned}x &= p_1/E_b \\y &= p_2/E_b \\ \cos \alpha &= \mathbf{p}_1 \cdot \mathbf{p}_2 / (p_1 p_2)\end{aligned}$$

Angle α is the angle between the momentum vectors of the two tracks of interest; it can be used for background suppression.

¹This happens way more often than a graduate student would think.

Table 20: Inclusive selection efficiencies at the $\Upsilon(1S)$.

CRITERION	BEFORE	AFTER	EFFICIENCY	CUMULATIVE
τ subcollection	30,000	22,664	76%	76%
at least one muon	22,664	19,154	85%	64%
$ \cos \alpha < 0.95$	19,154	9,038	47%	30%
$0.13 < p_{\perp}/E_{\text{cm}} < 1.10$	9,038	7,605	84%	25%
χ^2 per degree of freedom < 5	7,605	7,605	100%	25%
fitting region	7,605	6,009	79%	20%

To suppress background events arising from electron-positron annihilation to muon pairs and cosmic rays we require $|\cos \alpha| < 0.95$. This works because these backgrounds have collinear tracks, which our signal, because of the τ lepton's decay, will not necessarily have. The decay of the τ will also introduce a neutrino, which means that a signal event will have some undetected energy and momentum. Therefore, to suppress radiative muon pairs, which should be fully reconstructed, we look at the perpendicular component of the missing event momentum and require $0.13 < p_{\perp}/E_{\text{cm}} < 1.10$.

Table 20 shows that the overall detection efficiency is 20%. Note that since all events in the τ subcollection have at least two reconstructed charged particle tracks and have passed all levels of the trigger, the first row of the table includes those efficiencies. To estimate the sensitivity of the inclusive analysis we count the number of background events remaining in the signal region: $0.94 < x < 0.99$, after all selection criteria have been applied to the on-resonance $\Upsilon(4S)$ calibration sample. We then scale this number to the $\Upsilon(1S)$ resonance to estimate the number of background events in the signal region of our $\Upsilon(1S)$ data sample:

$$N_{\text{bkg.}} = 1.6 N_{\text{bkg., } \Upsilon(4S)} \left(\frac{\mathcal{L}_{\Upsilon(1S)}}{\mathcal{L}_{\Upsilon(4S)}} \right) \left(\frac{10,579 \text{ MeV}}{9,460 \text{ MeV}} \right)^2.$$

The number of signal events in the signal region is

$$N_{\text{sig.}} = \epsilon N_{\Upsilon} \mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}).$$

We want to know the smallest value of the branching fraction \mathcal{B} that will still give us enough signal events to make $\sqrt{N_{\text{sig.}}^2/N_{\text{bkg.}}} > 3$. This is

$$\mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}) > \frac{3\sqrt{N_{\text{bkg.}}}}{\epsilon N_{\Upsilon}}.$$

In other words the inclusive analysis will be able to see signal events if the lepton flavor violating branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm})$ is at least as large as $3\sqrt{N_{\text{bkg.}}}/(\epsilon N_{\Upsilon})$, where $N_{\text{bkg.}}$ is estimated from the number of events remaining in our calibration data sample that fall within the signal region. This procedure was also used to optimize the selection criteria on $|\cos \alpha|$ and p_{\perp}/E_{cm} for sensitivity. This is shown in Figure 16, where first $|\cos \alpha|$ is optimized and then p_{\perp}/E_{cm} . After imposing one final requirement on the quality of the fit for the muon candidate's track², we find that the inclusive selection criteria are sensitive to

$$\mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}) > (8 \pm 1) \times 10^{-6}$$

assuming errors similar to those in the exclusive analysis.

The effect of applying the optimized selection criteria to calibration data and to $\Upsilon(1S)$ signal Monte Carlo is shown in Figure 17. In the figure the vertical scale of the one-dimensional x -distributions is events/0.001, and the vertical scale of the one-dimensional y -distributions is events/0.005. We find that there is only one variable suitable for EML fitting: x , where the fitting region in x is the same as the one used in the exclusive analysis: $0.87 < x < 1.02$. There is only one fitting variable because we have no requirements on Track 2 other than that it exists. The background selection criteria remove events with $y \gtrsim 0.60$ in an x -dependent manner, so we impose one final requirement before fitting: $0.10 < y < 0.60$. However, we conclude that the inclusive analysis does not give enough improvement to justify fitting.

²For the muon candidate's track χ^2 per degree of freedom must be less than 5.

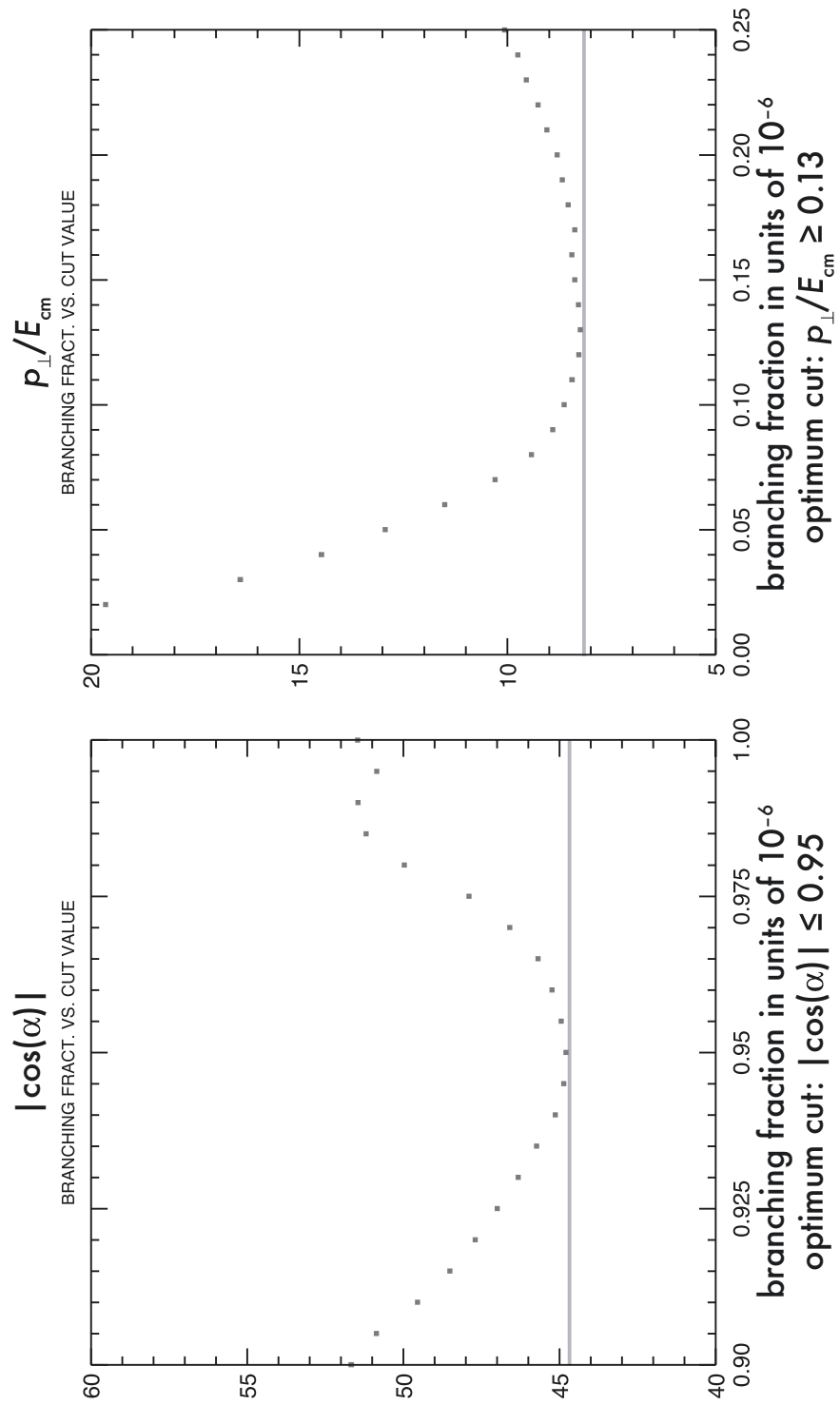


Figure 16: Optimization of selection criteria in the inclusive analysis.

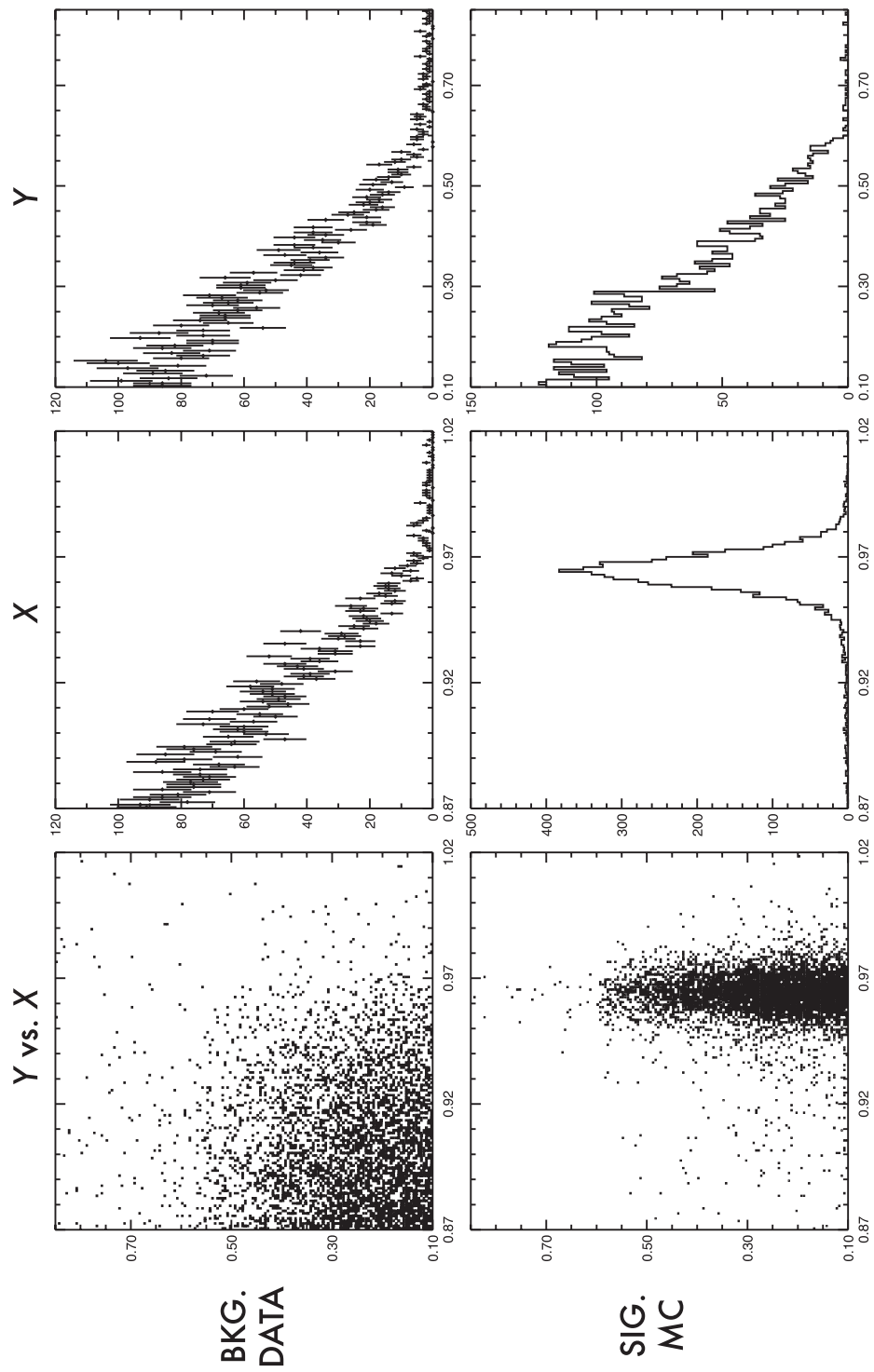


Figure 17: Scatter plots of y vs. x and their 1-d projections in the inclusive analysis.

XIII. CONCLUSIONS

The exclusive analysis, which identified the τ lepton in signal events through its decay to an electron, observed no signal events, and we used it to establish the following upper limits at the 95% confidence level:

$$\begin{aligned}\mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 6.0 \times 10^{-6} \\ \mathcal{B}(\Upsilon(2S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 14.4 \times 10^{-6} \\ \mathcal{B}(\Upsilon(3S) \rightarrow \mu^{\mp} + \tau^{\pm}) &< 20.3 \times 10^{-6}.\end{aligned}$$

Errors on the fitting parameters, detection efficiencies, and sample luminosities were accounted for in deriving these numbers. The inclusive analysis only looks for the signal's muon and consequently has much larger backgrounds. Its optimized selection criteria are sensitive to:

$$\mathcal{B}(\Upsilon(1S) \rightarrow \mu^{\mp} + \tau^{\pm}) > (8 \pm 1) \times 10^{-6}$$

assuming errors similar to those in the exclusive analysis.

We searched for physics beyond the scope of the Standard Model using a machine that was originally designed before the b quark was discovered. Interestingly, all but one of the necessary parameters for this analysis were measured from actual data. No signal was observed, so the conservation of the lepton flavor numbers remains in tact for now. Observing how this accidental symmetry is broken — and it is broken; neutrino oscillation tells us this — will have to wait for more energetic experiments.

APPENDIX

PLOTS

This appendix contains numerous plots related to fitting. The first group (Figures 18 – 26) illustrates the measurement of our fitting parameters from the calibration data.

- Figure 18 shows projections of the on-resonance $\Upsilon(4S)$ data, *i.e.* our calibration sample, onto each of the four fit variables.
- Figure 19 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of x below 0.96. This result supports the conclusion that a straight line is a good approximation for the x -shape of our τ pairs background.
- Figure 20 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of x between 0.87 and 0.96. This result shows that the y -shape for real electrons is well approximated by a 2nd-order polynomial.
- Figure 21 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of x between 0.87 and 0.96. This result shows that the shape of E/p for real electrons can be described by the Crystal Ball function.
- Figure 22 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of x between 0.87 and 0.96. This result shows that the shape of dE/dx for real electrons can be described by a Gaussian function.
- Figure 23 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of y between 0.45 and 0.60. These are events that contain one high-quality muon and one poorly reconstructed muon. This fit shows that

the x -shape of such events can be described by the Crystal Ball function.

- Figure 24 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the region of x between 0.96 and 1.02. These are radiative muon pairs. The fit is done in the rotated variable y' , which was described in Chapter 10, and it shows that the y' -shape of this background can be described by the Crystal Ball function.
- Figure 25 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the following regions: x between 0.96 and 1.02, y between 0.45 and 0.60, and dE/dx below -1.0 . These are radiative muon pairs. The fit is done in the rotated variable u' , which was described in Chapter 10, and it shows that a straight line is a good approximation for the u' -shape of this background.
- Figure 26 shows the results of a binned maximum-likelihood fit to our calibration data. The fit is performed for the following regions: x between 0.96 and 1.02 and y between 0.45 and 0.60. These are radiative muon pairs. This fit shows that the shape of dE/dx for this background is well described by a Gaussian function.

The next group (Figures 27 – 30) illustrates the detection signature of signal events. Note that the signal's peak is determined kinematically and its spread is determined by the detector resolution. We also use the detector resolution to smear the shape of x for the τ pairs background.

- Figure 27 shows the results of a binned maximum-likelihood fit to $\Upsilon(1S)$ signal Monte Carlo. This result supports the conclusion that the detection signature of signal events is a monochromatic muon whose momentum distribution is smeared by the detector's resolution. We do not use these results to fit our signal data. Instead we use a Gaussian with σ_x measured from our calibration data.
- Figure 28 is the same as the previous except that it shows $\Upsilon(2S)$ signal Monte Carlo.
- Figure 29 is the same as the previous except that it shows $\Upsilon(3S)$ signal Monte Carlo.
- Figure 30 shows the results of a binned maximum-likelihood fit to our calibration data. The events have been selected to enhance radiative muon pairs. This is our measurement of the detector's momentum resolution σ_x .

Next (Figures 31 – 35) we show the results of our unbinned extended maximum-likelihood (EML) fits.

- Figure 31 shows projections of the results of a three-component (no signal) unbinned EML fit in four dimensions along with binned $\Upsilon(4S)$ calibration data for comparison. Solid line (black) shows all components; dashed line (light grey) shows τ pairs; dotted line (medium grey) shows μ pairs with radiation; dot-dashed line (dark grey) shows μ pairs with decay. Also shown (upper left only) is a hypothetical signal peak containing 100 events.
- Figure 32 is the same as the previous except that it shows continuum control data.
- Figure 33 shows projections of the results of a four-component unbinned EML fit in four dimensions along with binned $\Upsilon(1S)$ signal data for comparison. Solid line (black) shows all components; dashed line (light grey) shows τ pairs; dotted line (medium grey) shows μ pairs with radiation; dot-dashed line (dark grey) shows μ pairs with decay. Also shown (upper left only) is a hypothetical signal peak containing 100 events.
- Figure 34 is the same as the previous except that it shows $\Upsilon(2S)$ signal data.
- Figure 35 is the same as the previous except that it shows $\Upsilon(3S)$ signal data.

The final group of plots (Figures 36 – 40) illustrates the determination of upper limits.

- Figure 36 shows the profile of the likelihood function *vs.* LFV branching fraction for $\Upsilon(4S)$ calibration data. (a) No correlations among fit variables and no systematic uncertainties are taken into account. (b) Correlations are accounted for but not systematics. (c) Correlations and systematic uncertainties in PDF shapes are accounted for. (d) Correlations, PDF systematics, and efficiency systematics are accounted for. (e) Correlations and all systematic uncertainties are accounted for.
- Figure 37 is the same as the previous except that it shows continuum control data.
- Figure 38 is the same as the previous except that it shows $\Upsilon(1S)$ signal data.
- Figure 39 is the same as the previous except that it shows $\Upsilon(2S)$ signal data.
- Figure 40 is the same as the previous except that it shows $\Upsilon(3S)$ signal data.

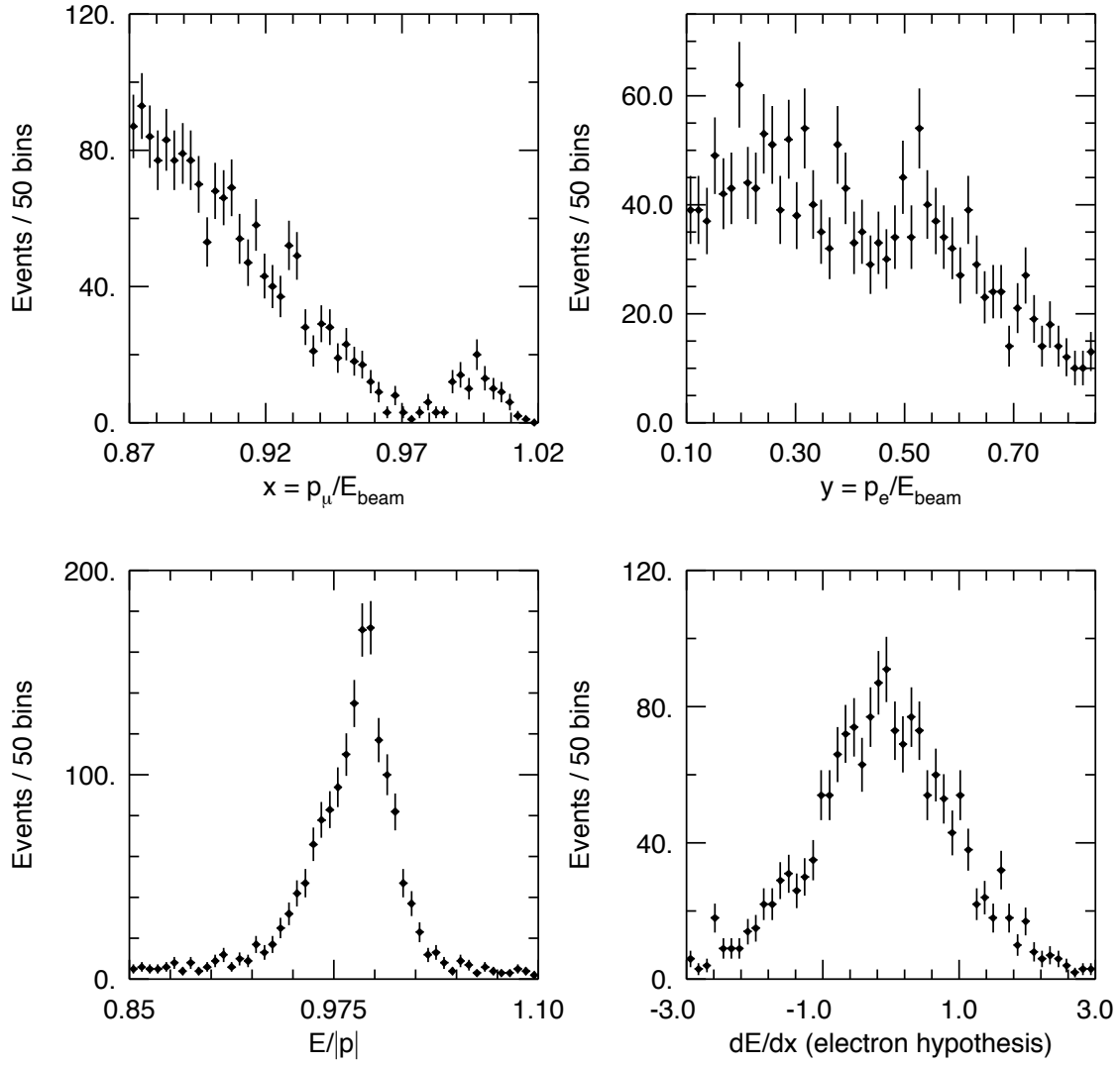


Figure 18: Projections of the on-resonance $\Upsilon(4S)$ data onto each of our fit variables.

Likelihood = 64.5
 $\chi^2 = 63.8$ for 72 - 1 d.o.f.,
 Errors
 Function 1: x for $\tau\tau$

		Parabolic	Minos	C.L.= 71.6%
A	385.07	± 9.630	- 9.548 + 9.712	
* B	0.97100	± 0.000	- 0.000 + 0.000	
* ALPHA	0.0000	± 0.000	- 0.000 + 0.000	

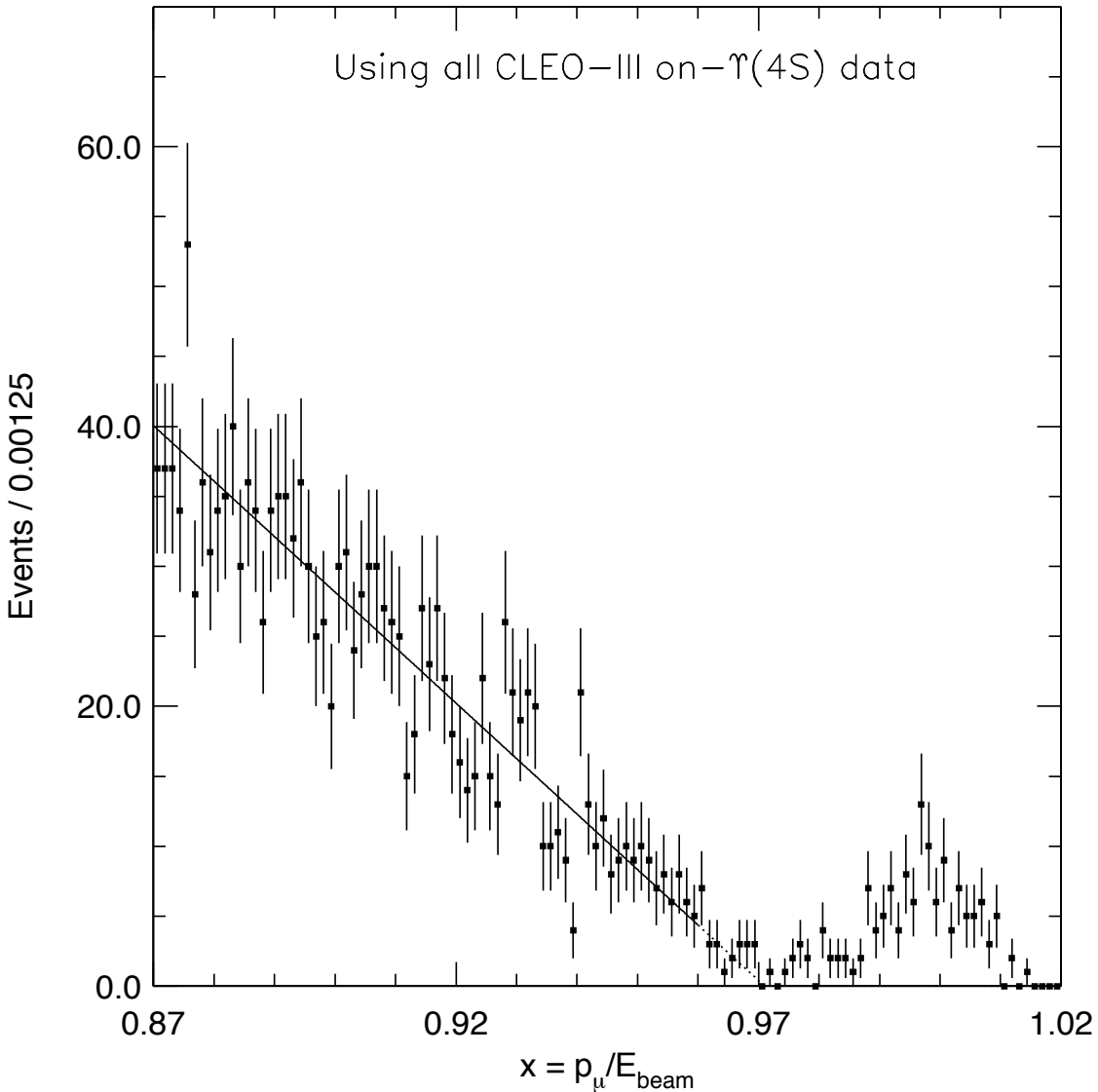


Figure 19: The shape of x for our τ pairs background.

Likelihood = 89.7
 $\chi^2 = 91.2$ for 75 - 3 d.o.f., C.L. = 6.2%
 Errors

		Parabolic	Minos	
Function 1: Chebyshev Polynomial of Order 2				
NORM	2051.9	± 58.20	- 57.45	+ 58.97
CHEB01	-0.58446	$\pm 3.8604E-02$	- 3.8193E-02	+ 3.9009E-02
CHEB02	-0.11689	$\pm 4.1710E-02$	- 4.1917E-02	+ 4.1485E-02

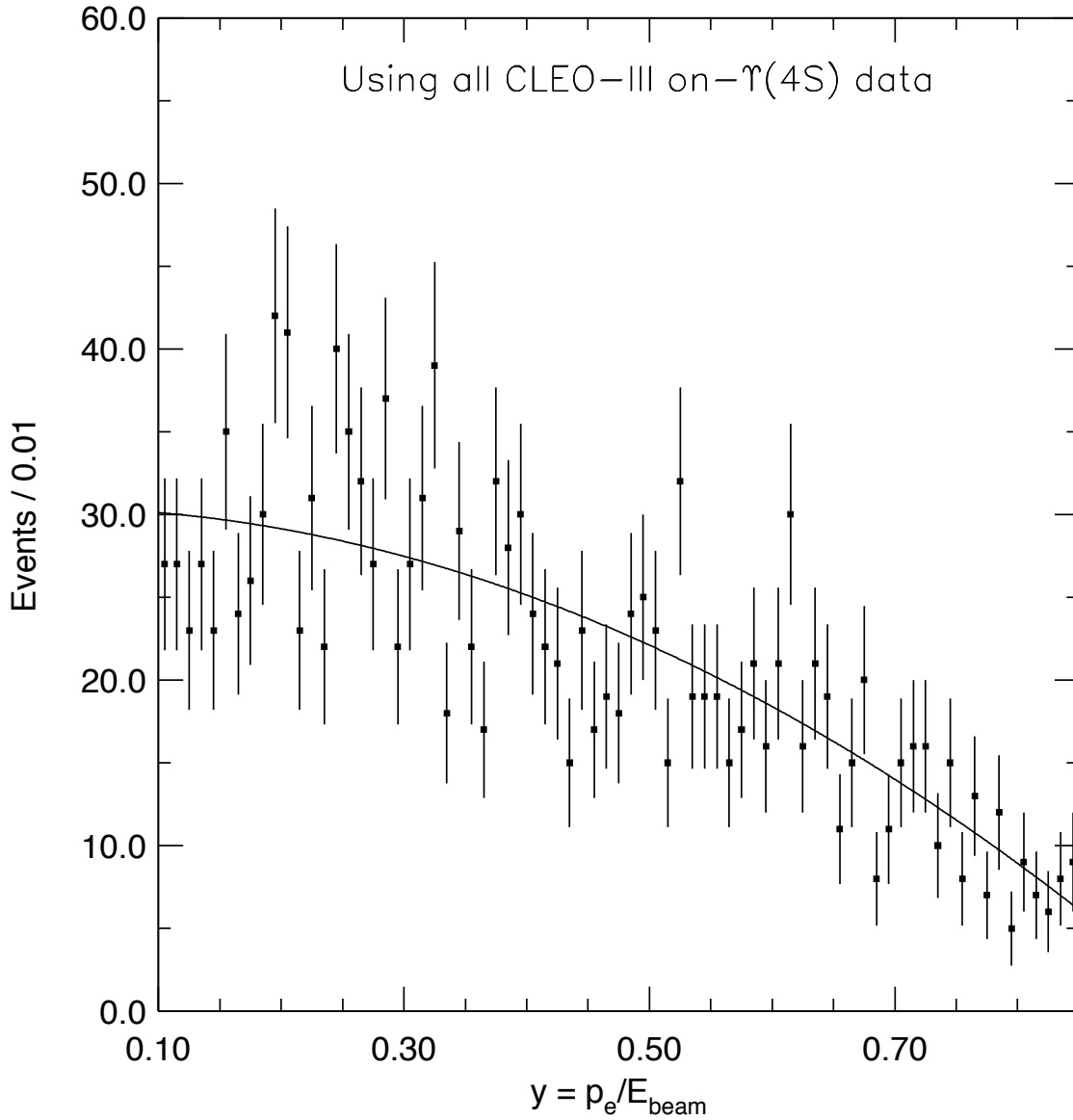


Figure 20: The shape of y for real electrons.

Likelihood = 93.7
 $\chi^2 = 85.0$ for 100 - 5 d.o.f., C.L. = 75.9%

Errors		Parabolic	Minos	
Function 1: CB line shape				
NORM	1414.8	± 42.99	- 42.42	+ 43.08
MEAN	0.99142	$\pm 6.4835E-04$	- 6.5192E-04	+ 6.4572E-04
SIGMA	1.80197E-02	$\pm 5.9889E-04$	- 5.8831E-04	+ 6.0939E-04
* ALPHA	1.0000	± 0.000	- 0.000	+ 0.000
* BETA	9.0000	± 0.000	- 0.000	+ 0.000
Function 2: Chebyshev Polynomial of Order 1				
NORM	747.79	± 96.43	- 94.16	+ 99.03
CHEB01	-3.82168E-02	± 0.1528	- 0.1497	+ 0.1545

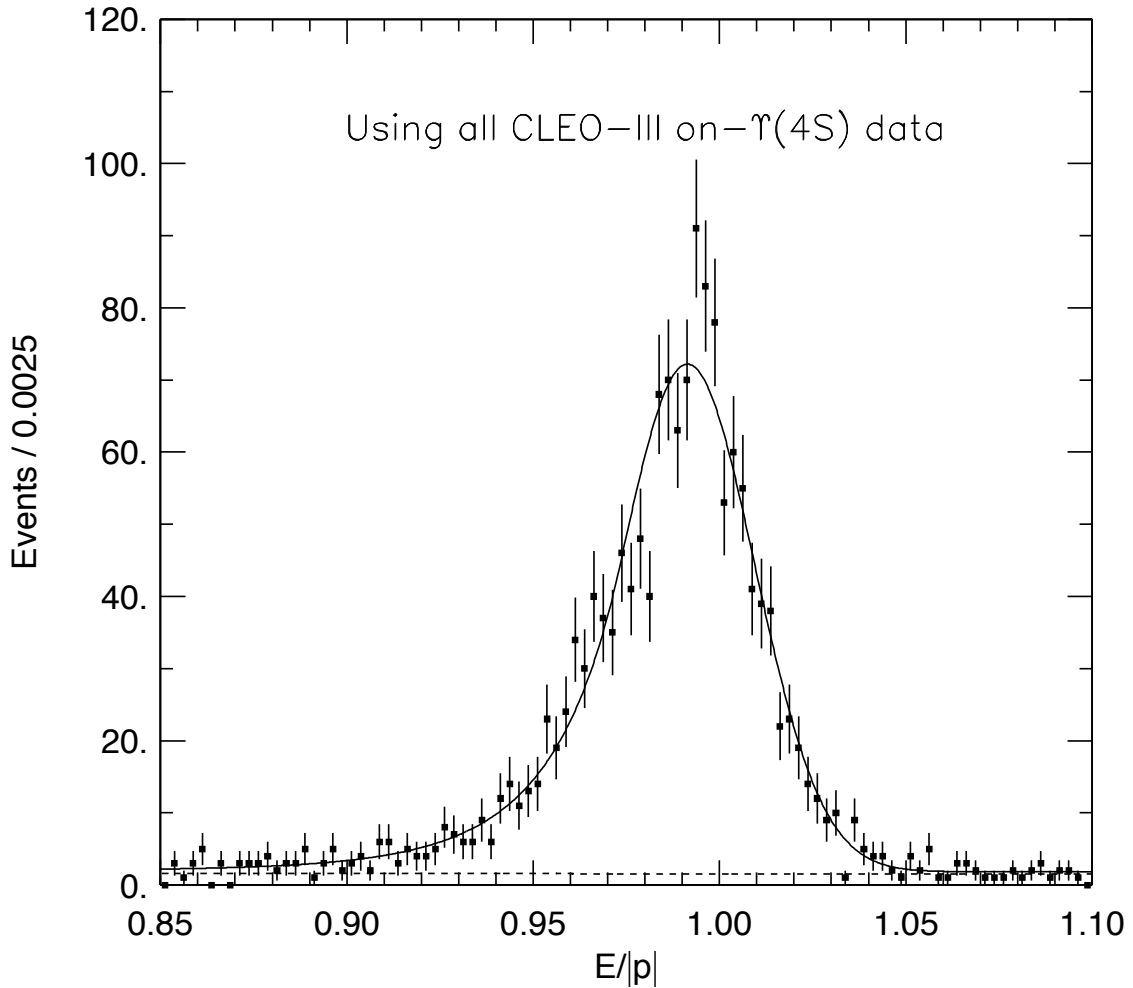


Figure 21: The shape of E/p for real electrons.

Likelihood = 57.5
 $\chi^2 = 58.8$ for 50 - 3 d.o.f.,
 Errors Parabolic Minos C.L.= 11.7%
 Function 1: Gaussian (sigma)
 AREA 1562.3 \pm 40.73 - 40.38 + 41.09
 MEAN -3.66864E-02 \pm 3.0013E-02 - 3.0335E-02 + 2.9735E-02
 SIGMA 1.0379 \pm 2.5009E-02 - 2.4352E-02 + 2.5711E-02

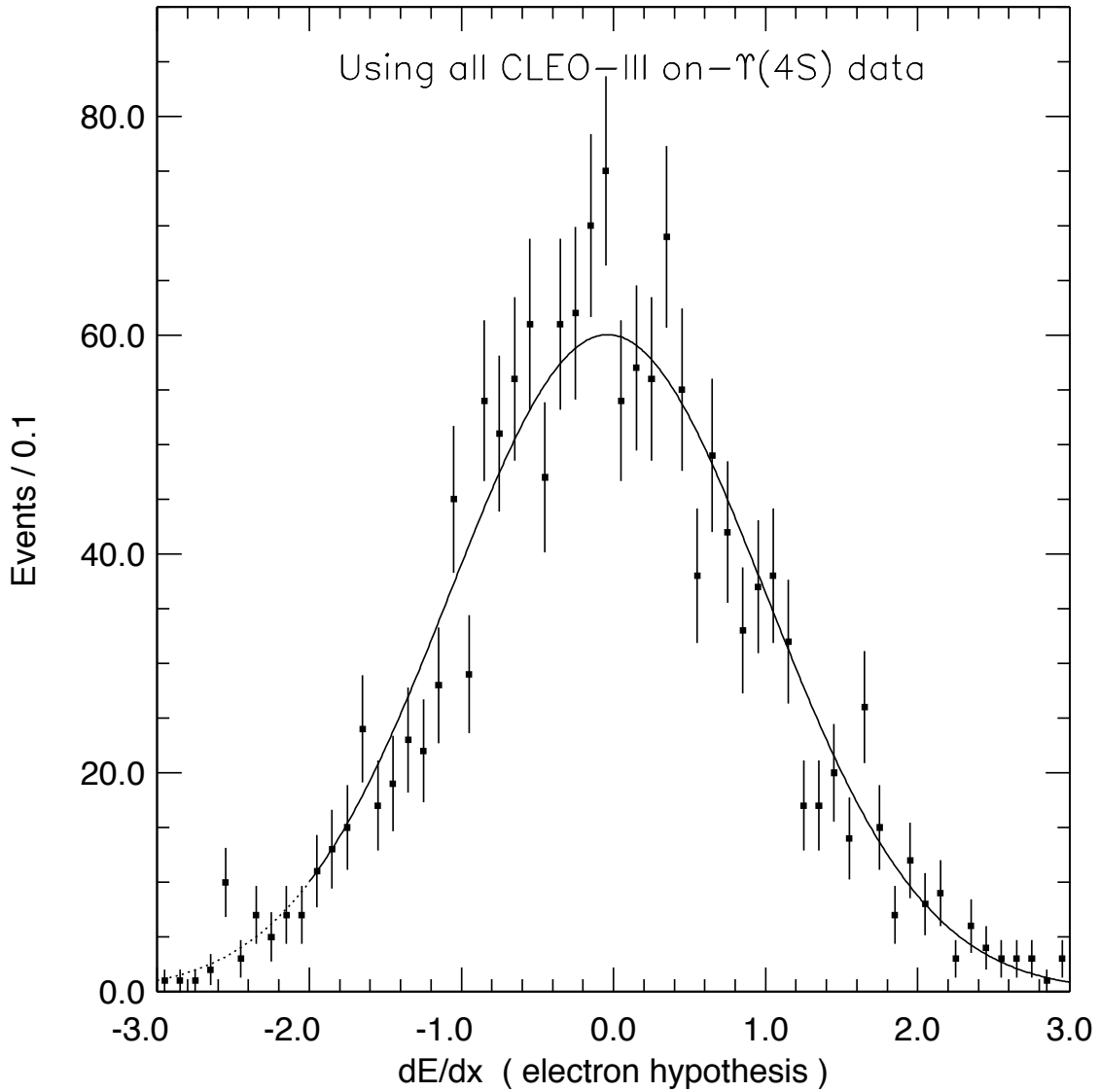


Figure 22: The shape of dE/dx for real electrons.

Likelihood = 152.2

$\chi^2 = 150.1$ for 120 - 5 d.o.f.,

C.L. = 1.6%

Errors	Parabolic	Minos	
Function 1: CB line shape			
A	922.21 ± 35.63	- 33.79 + 39.61	
MEAN	0.99818 ± 4.4095E-04	- 5.4580E-04 + 5.1335E-04	
SIGMA	8.67925E-03 ± 3.8676E-04	- 4.3468E-04 + 4.8469E-04	
ALPHA	0.92955 ± 1.7298E-02	- 0.1011 + 0.1180	
BETA	2.0420 ± 0.1967	- 0.3279 + 0.4068	

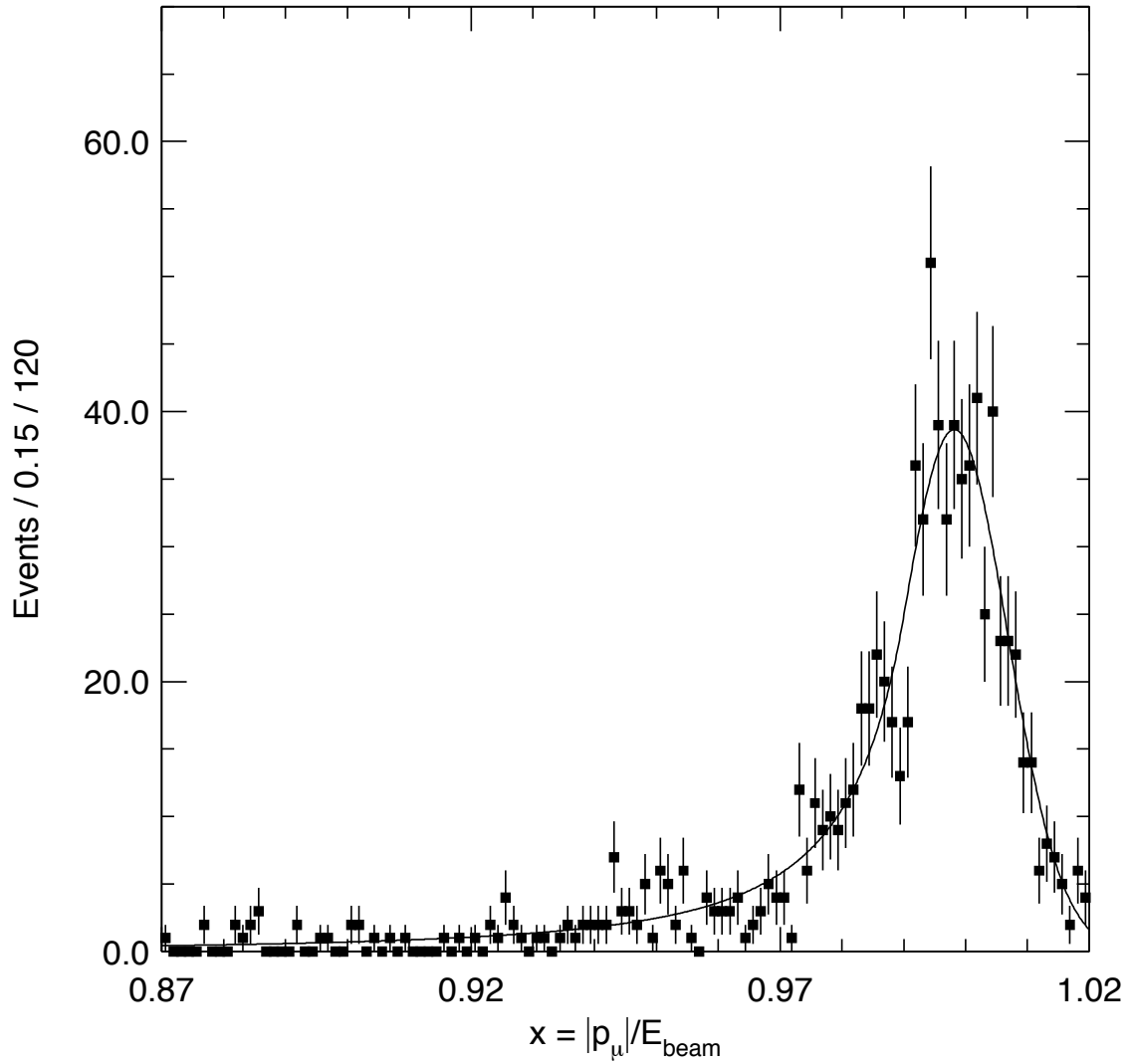


Figure 23: The shape of x for μ pairs.

Likelihood = 42.9
 $\chi^2 = 42.4$ for 32 - 3 d.o.f.,
 Errors
 Function 1: CB line shape
 A 734.47 \pm 27.99 - 27.63 + 28.36
 MEAN 0.56318 \pm 2.9954E-04 - 3.0185E-04 + 2.9891E-04
 SIGMA 6.43317E-03 \pm 2.5477E-04 - 2.4681E-04 + 2.6422E-04
 * ALPHA 1.0000 \pm 0.000 - 0.000 + 0.000
 * BETA 6.0000 \pm 0.000 - 0.000 + 0.000
 C.L.= 5.2%

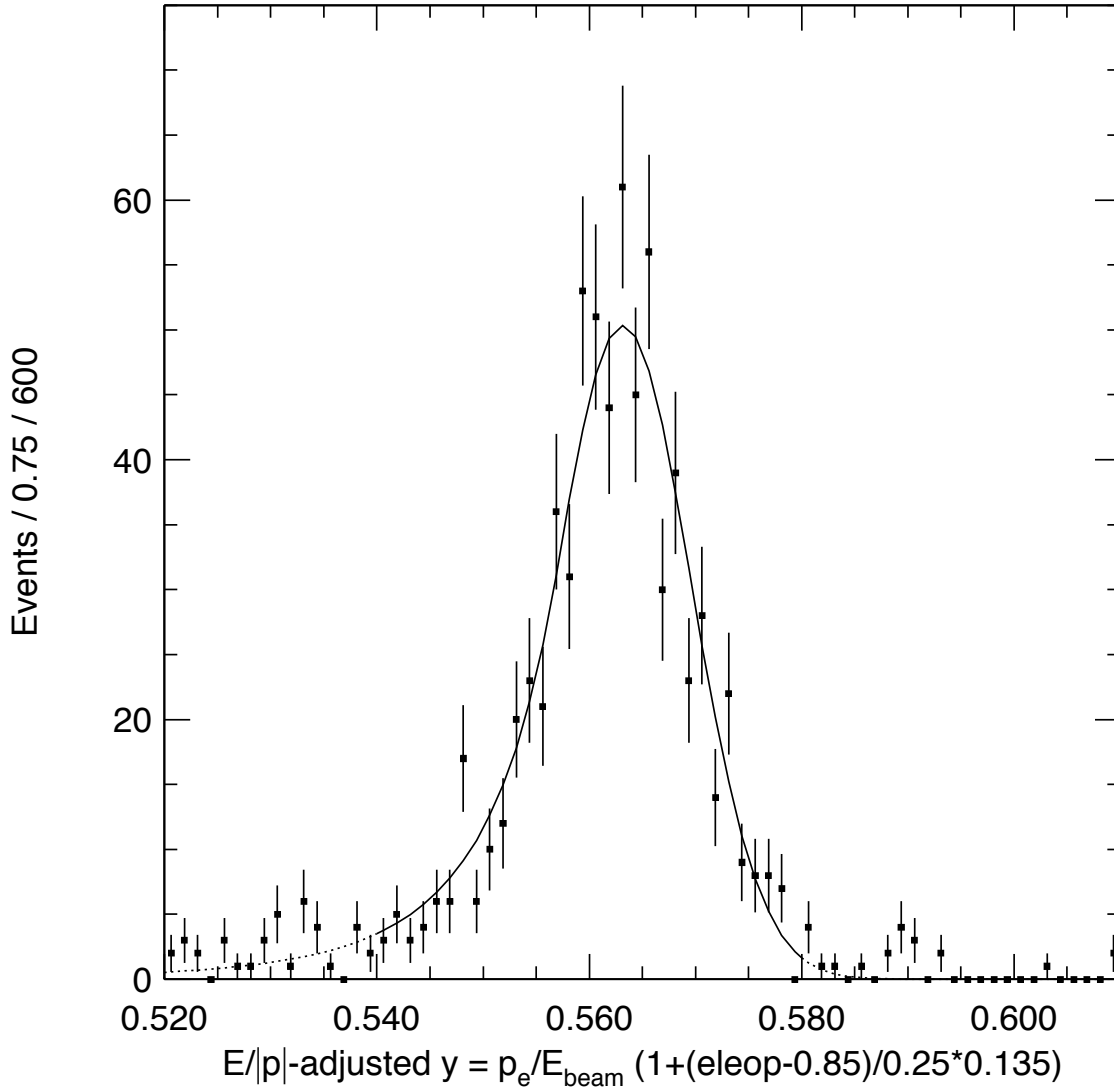


Figure 24: The shape of the rotated variable y' for radiative μ pairs.

Likelihood = 13.9
 $\chi^2 = 13.3$ for 12 - 2 d.o.f.,
 Errors Parabolic Minos C.L.= 20.7%
 Function 1: Line = NORM*(1+SLOPE*X)
 NORM 184.17 \pm 6.250 - 26.07 + 26.08
 SLOPE -0.81338 \pm 1.1419E-02 - 4.8389E-02 + 6.4988E-02

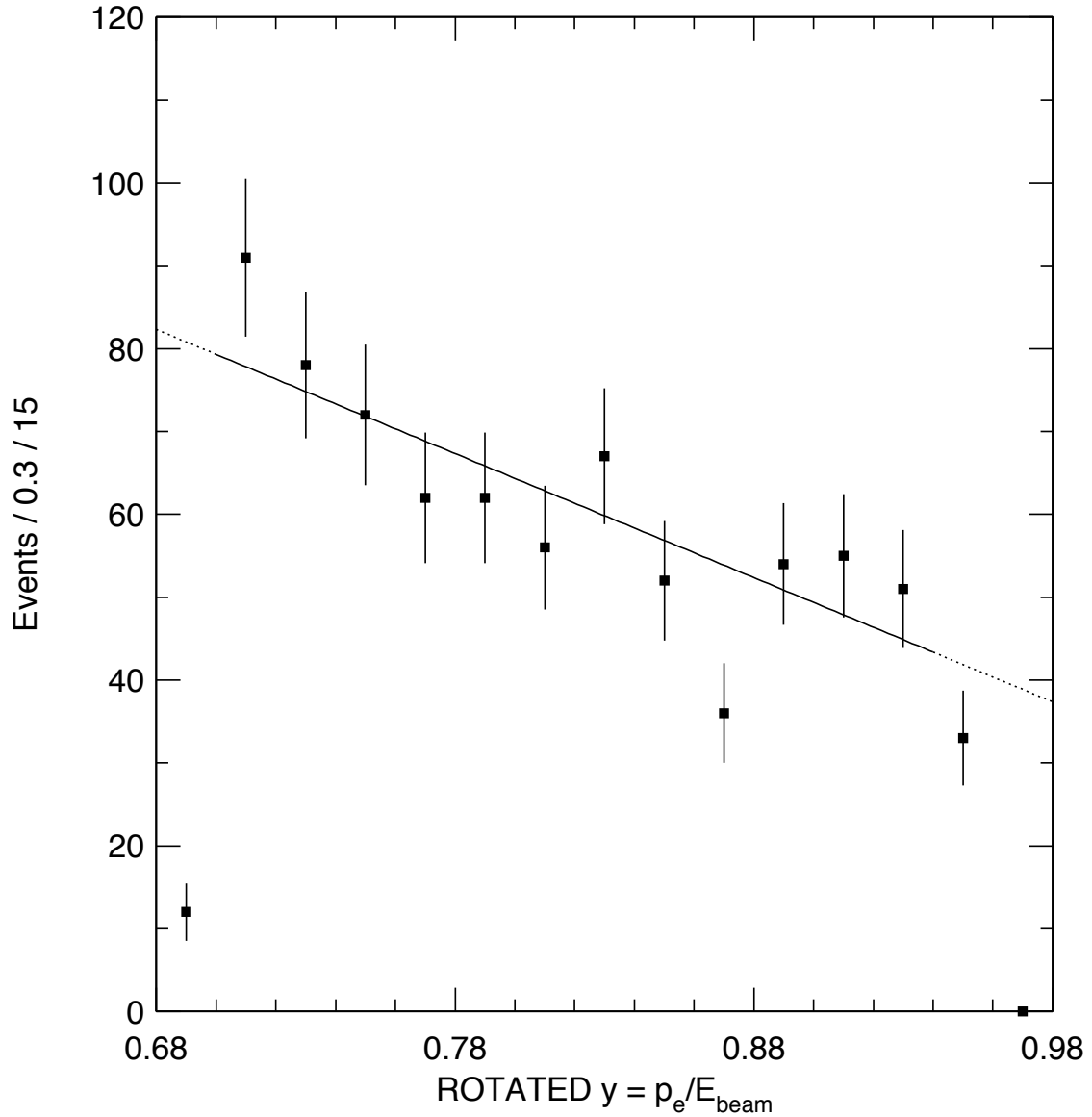


Figure 25: The shape of the rotated variable u' for radiative μ pairs.

Likelihood = 26.2
 $\chi^2 = 26.1$ for 40 - 3 d.o.f.,
 Errors Parabolic Minos C.L.= 91.1%
 Function 1: Gaussian (sigma)
 AREA 814.31 \pm 30.74 - 30.24 + 31.26
 MEAN -2.4182 \pm 4.2526E-02 - 4.3618E-02 + 4.1652E-02
 SIGMA 0.96624 \pm 3.8235E-02 - 3.6414E-02 + 4.0293E-02

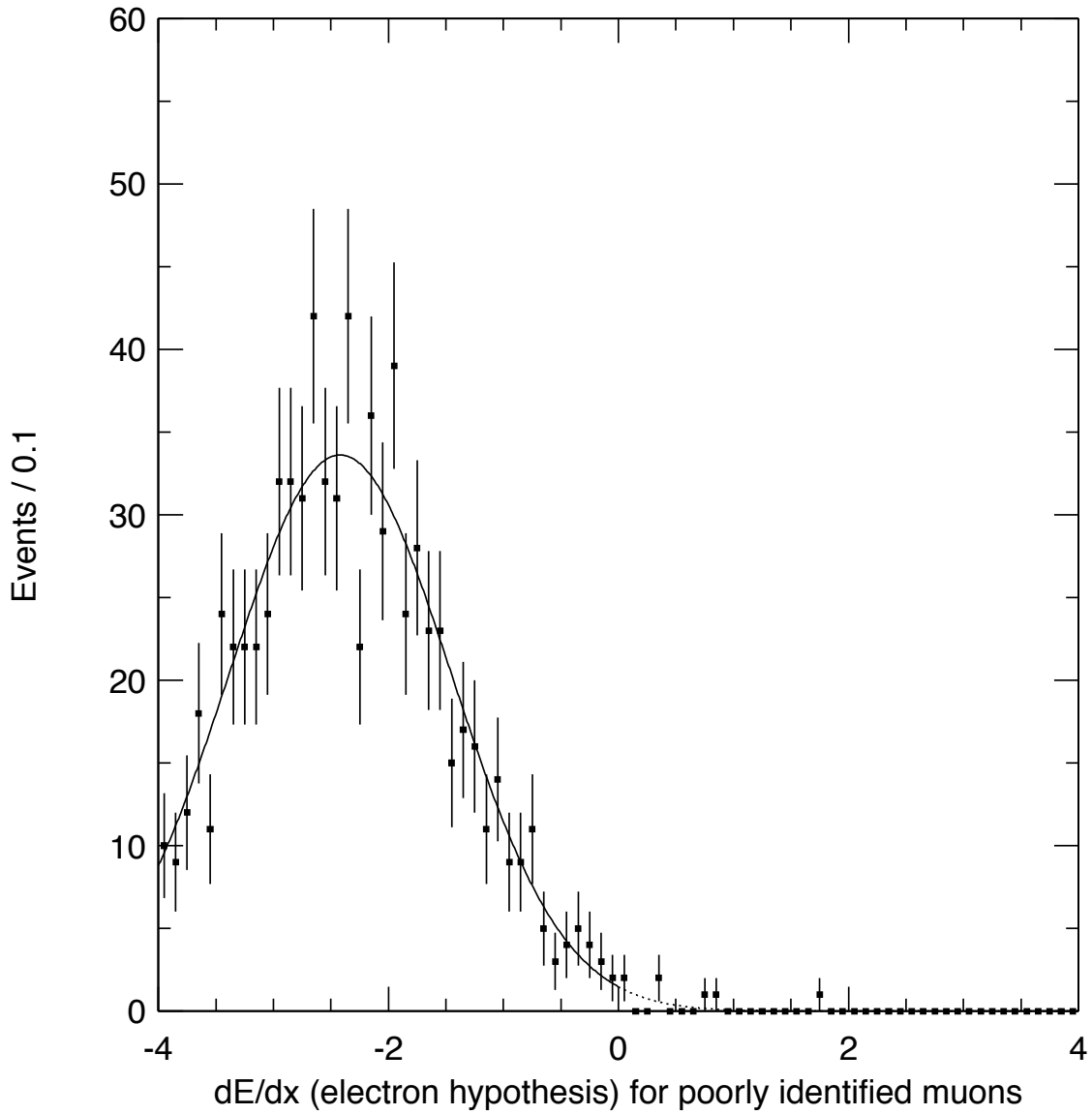


Figure 26: The shape of dE/dx for radiative μ pairs.

Likelihood = 19.4
 $\chi^2 = 19.5$ for 24 - 3 d.o.f.,
 Errors Parabolic Minos C.L.= 55.1%
 Function 1: Gaussian (sigma)

		±	-	+
AREA	2475.4	50.79	50.43	51.15
MEAN	0.96465	1.4505E-04	1.4521E-04	1.4499E-04
SIGMA	6.60704E-03	1.2744E-04	1.2459E-04	1.3044E-04

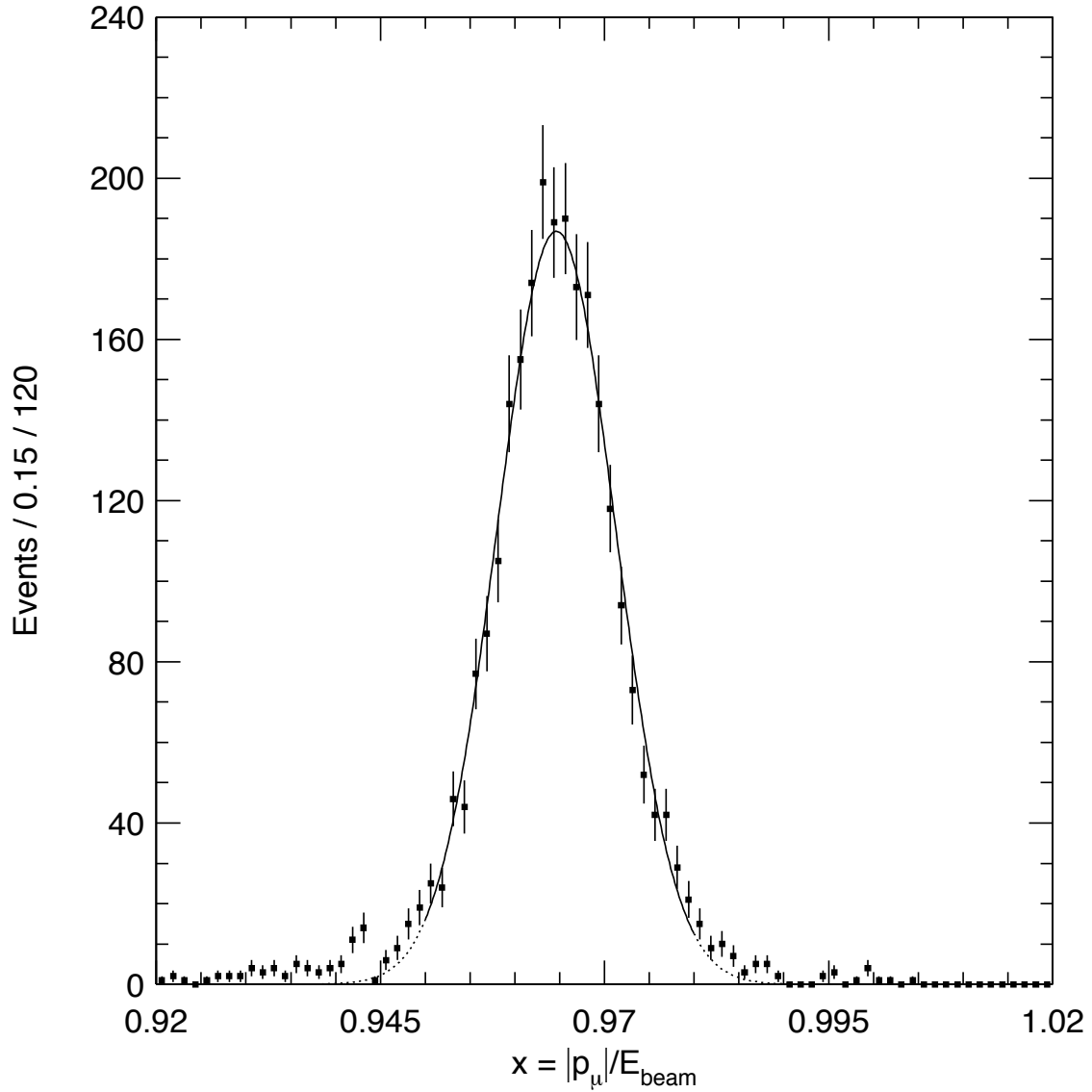


Figure 27: The shape of x for $\Upsilon(1S)$ signal Monte Carlo.

Likelihood = 34.8
 $\chi^2 = 36.4$ for 24 - 3 d.o.f.,
 Errors Parabolic Minos C.L.= 2.0%
 Function 1: Gaussian (sigma)

		Parabolic	Minos	
AREA	2508.5	± 52.41	- 52.00	+ 52.82
MEAN	0.96836	$\pm 1.6790E-04$	- 1.6834E-04	+ 1.6763E-04
SIGMA	7.31912E-03	$\pm 1.5997E-04$	- 1.5566E-04	+ 1.6457E-04

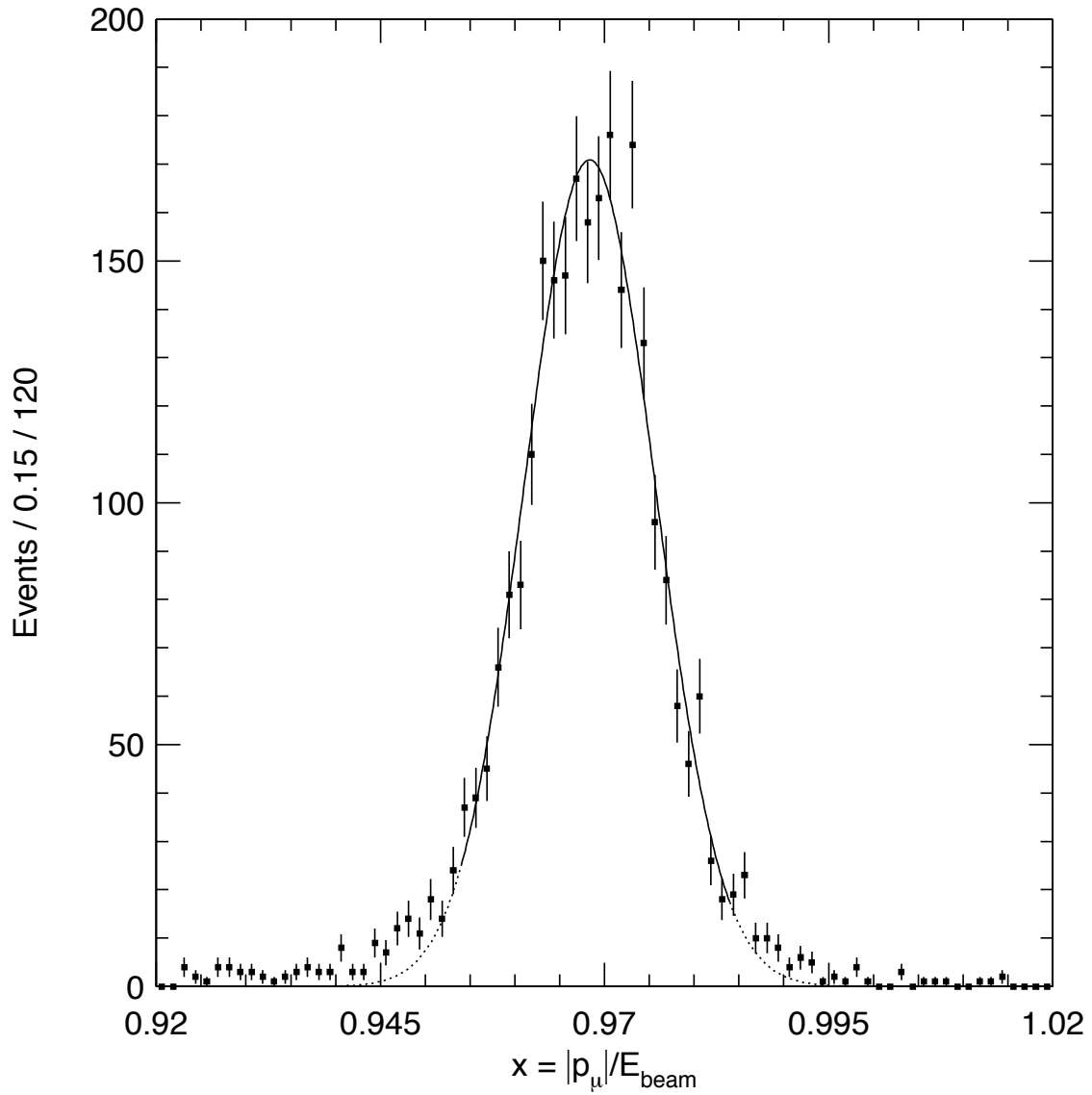


Figure 28: The shape of x for $\Upsilon(2S)$ signal Monte Carlo.

Likelihood = 28.7
 $\chi^2 = 28.8$ for 24 - 3 d.o.f.,
 Errors Parabolic Minos C.L.= 11.9%
 Function 1: Gaussian (sigma)

		Parabolic	Minos	
AREA	1796.8	± 44.28	- 43.85	+ 44.72
MEAN	0.97021	$\pm 1.9689E-04$	- 1.9694E-04	+ 1.9711E-04
SIGMA	7.28895E-03	$\pm 1.8706E-04$	- 1.8136E-04	+ 1.9321E-04

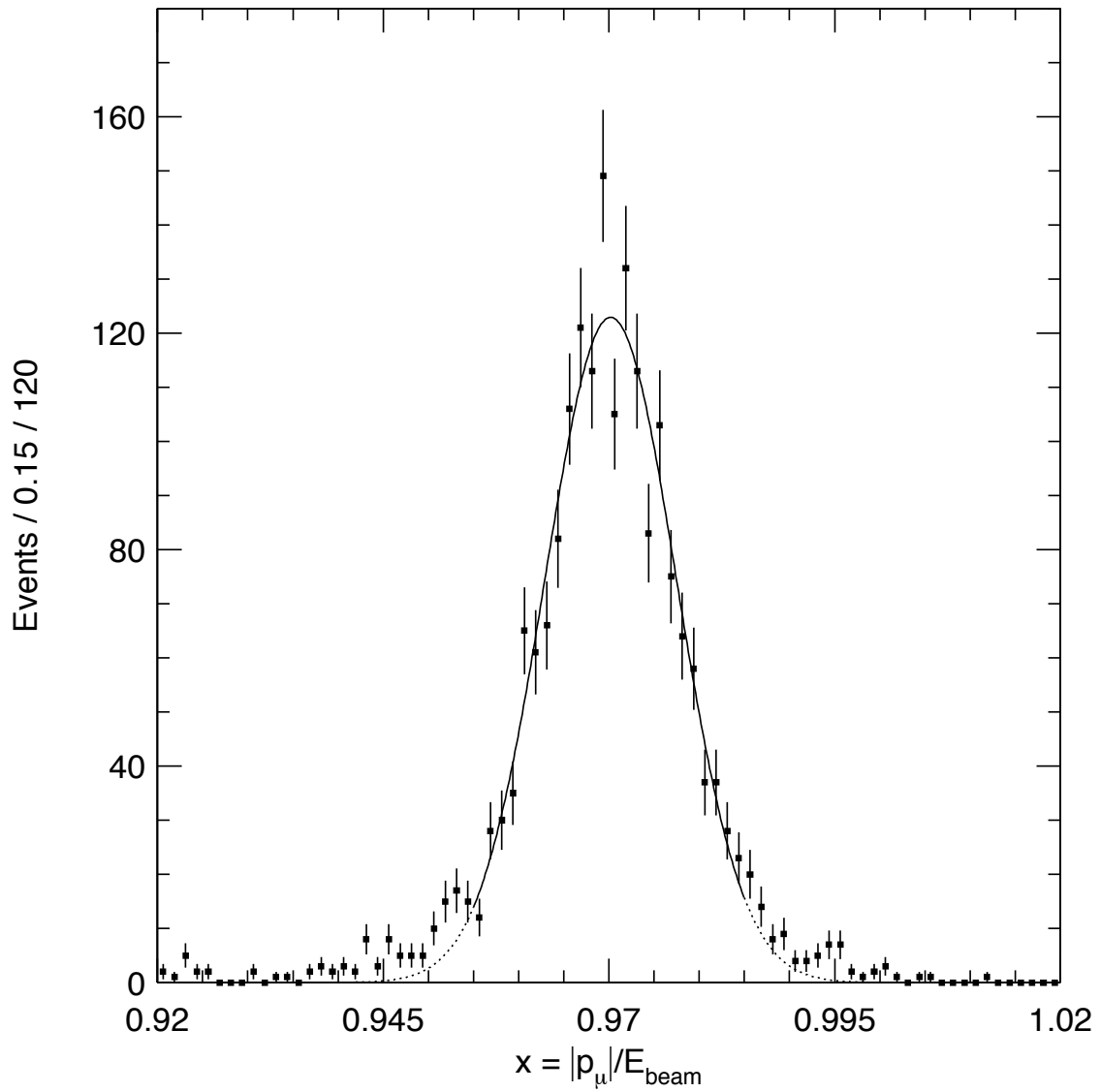


Figure 29: The shape of x for $\Upsilon(3S)$ signal Monte Carlo.

Likelihood = 75.4
 $\chi^2 = 76.0$ for 80 - 4 d.o.f.,
 Errors
 Function 1: Crystal Ball line shape
 Parabolic Minos
 'A' 11321. \pm 246.3 - 226.1 + 275.1
 'MEAN' 0.99821 \pm 1.4023E-04 - 1.3985E-04 + 1.4071E-04
 'SIGMA' 8.64671E-03 \pm 1.9981E-04 - 1.9656E-04 + 2.0363E-04
 * 'ALPHA' 1.0000 \pm 0.000 - 0.000 + 0.000
 'BETA' 3.2372 \pm 0.6770 - 0.5685 + 0.8427

C.L.= 47.8%

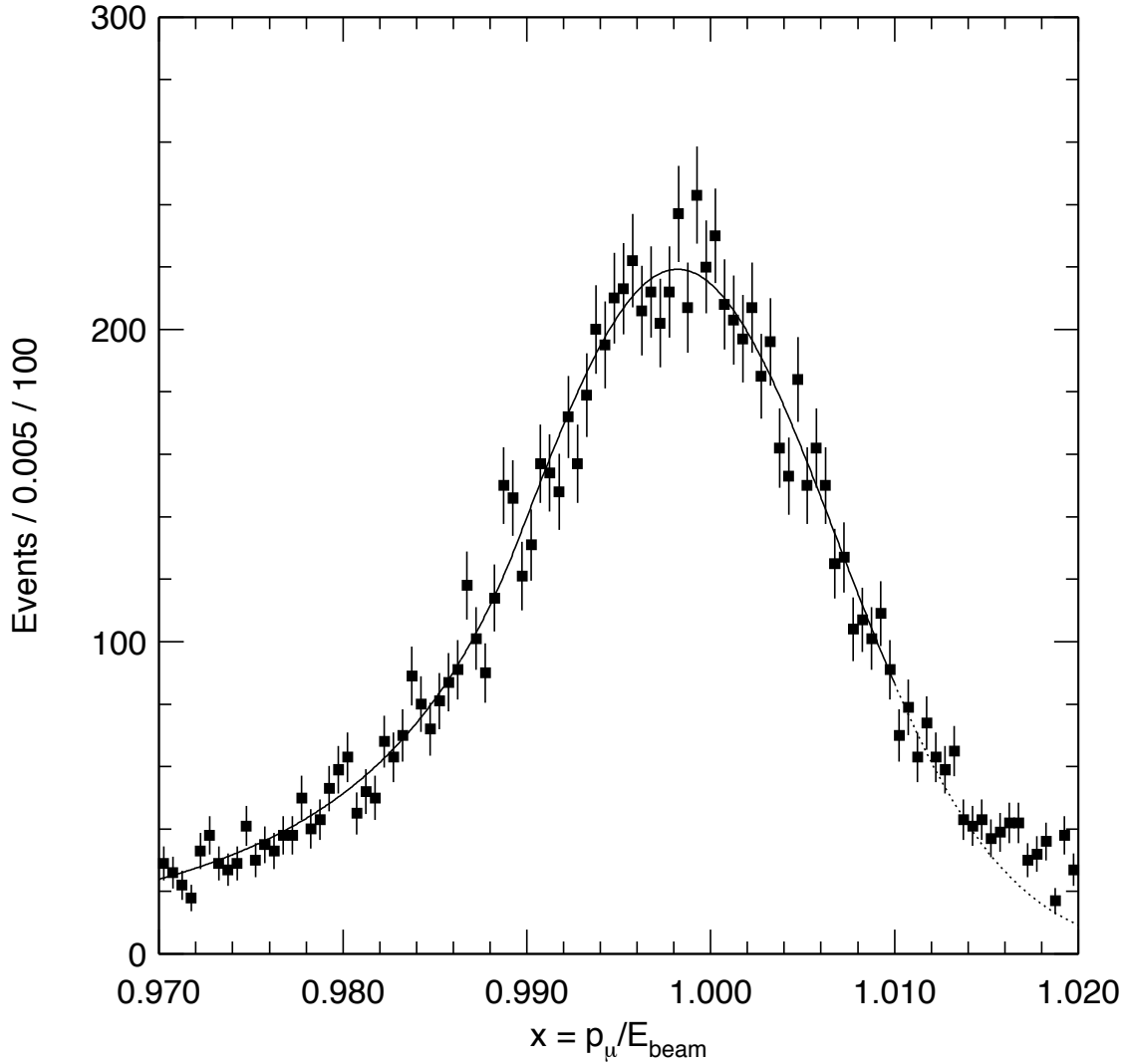


Figure 30: The smearing in x , measured from our $\Upsilon(4S)$ calibration data.

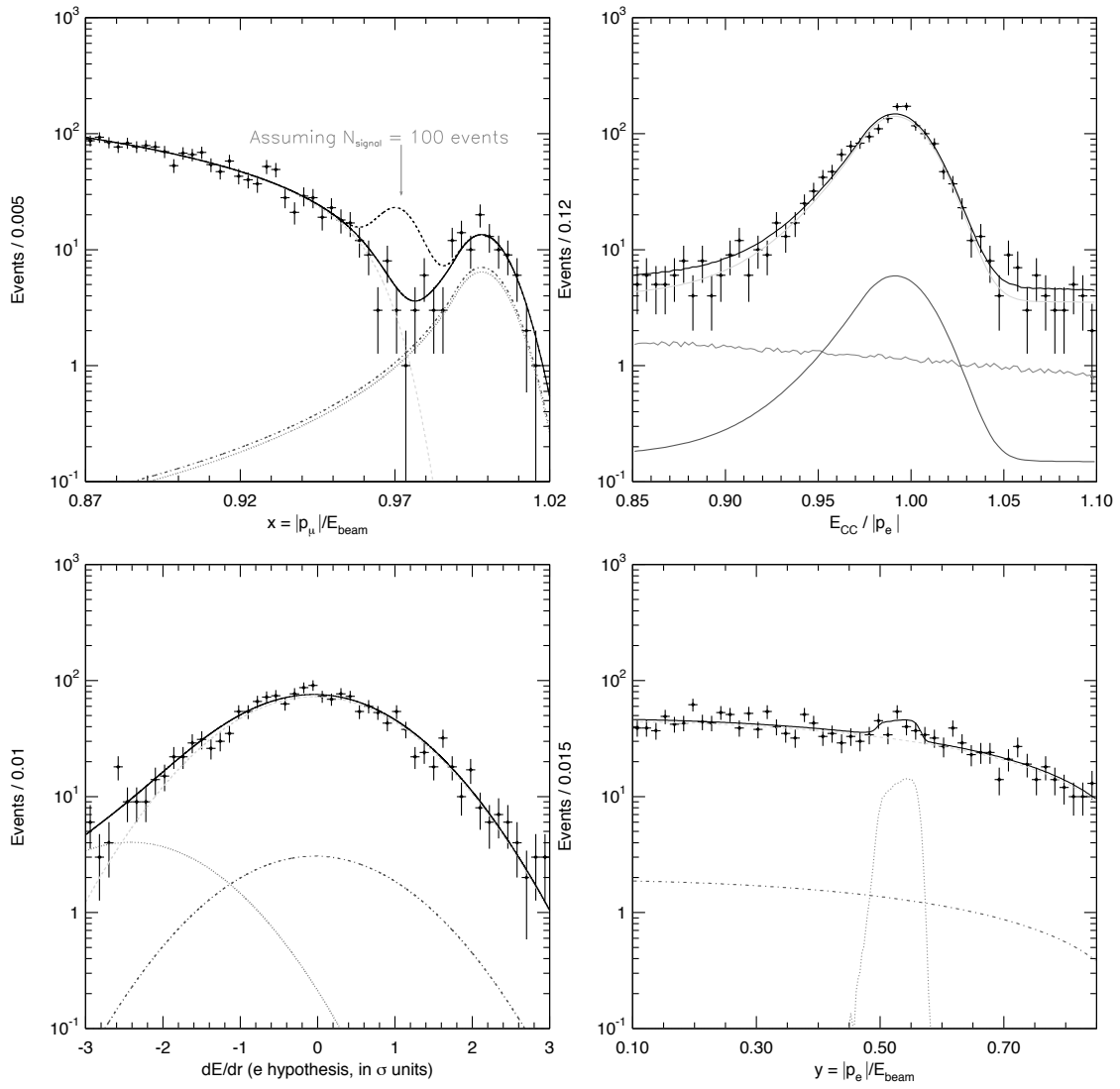


Figure 31: The results of a three-component fit to our calibration data.

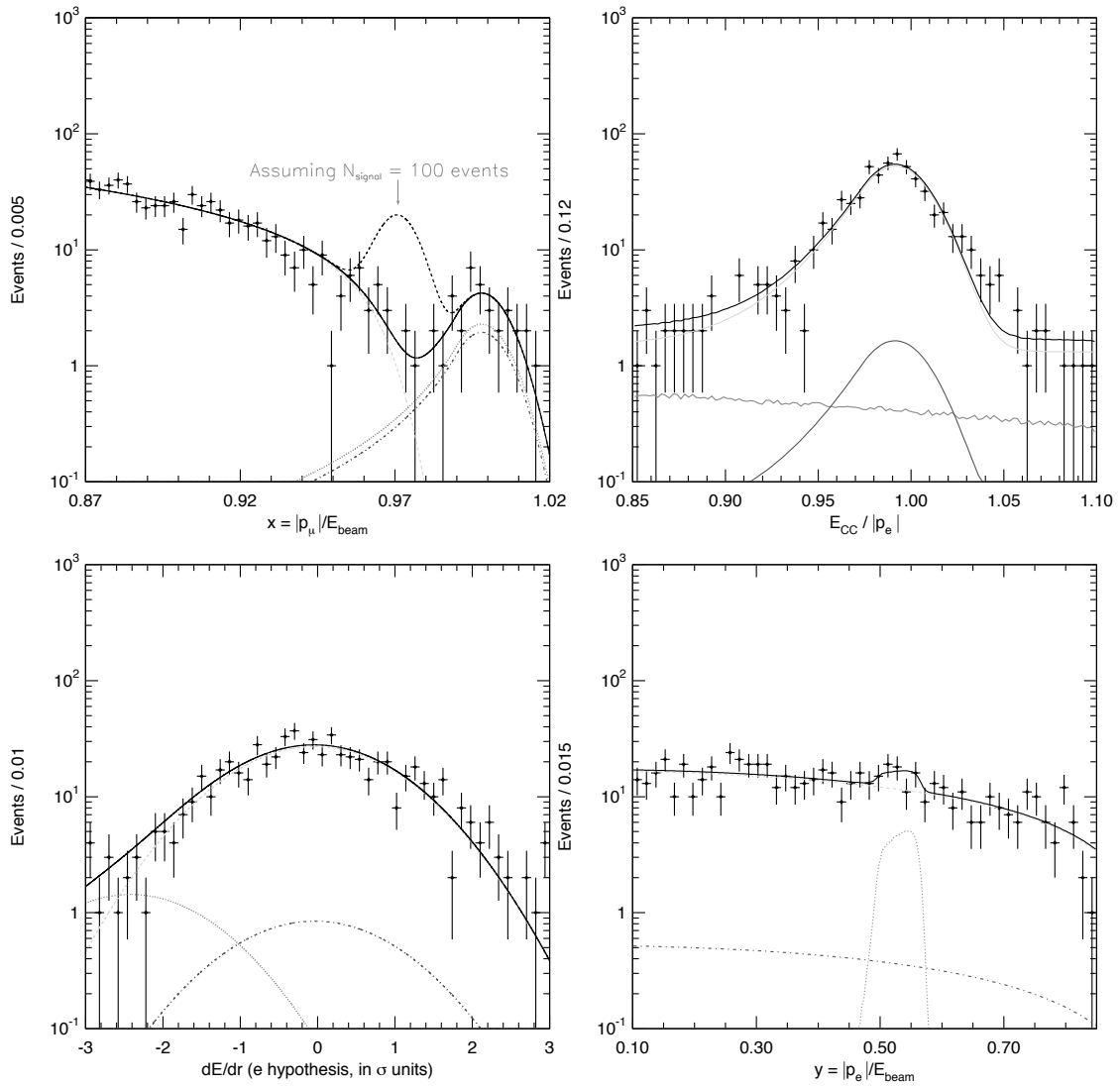


Figure 32: The results of a three-component fit to our control data.

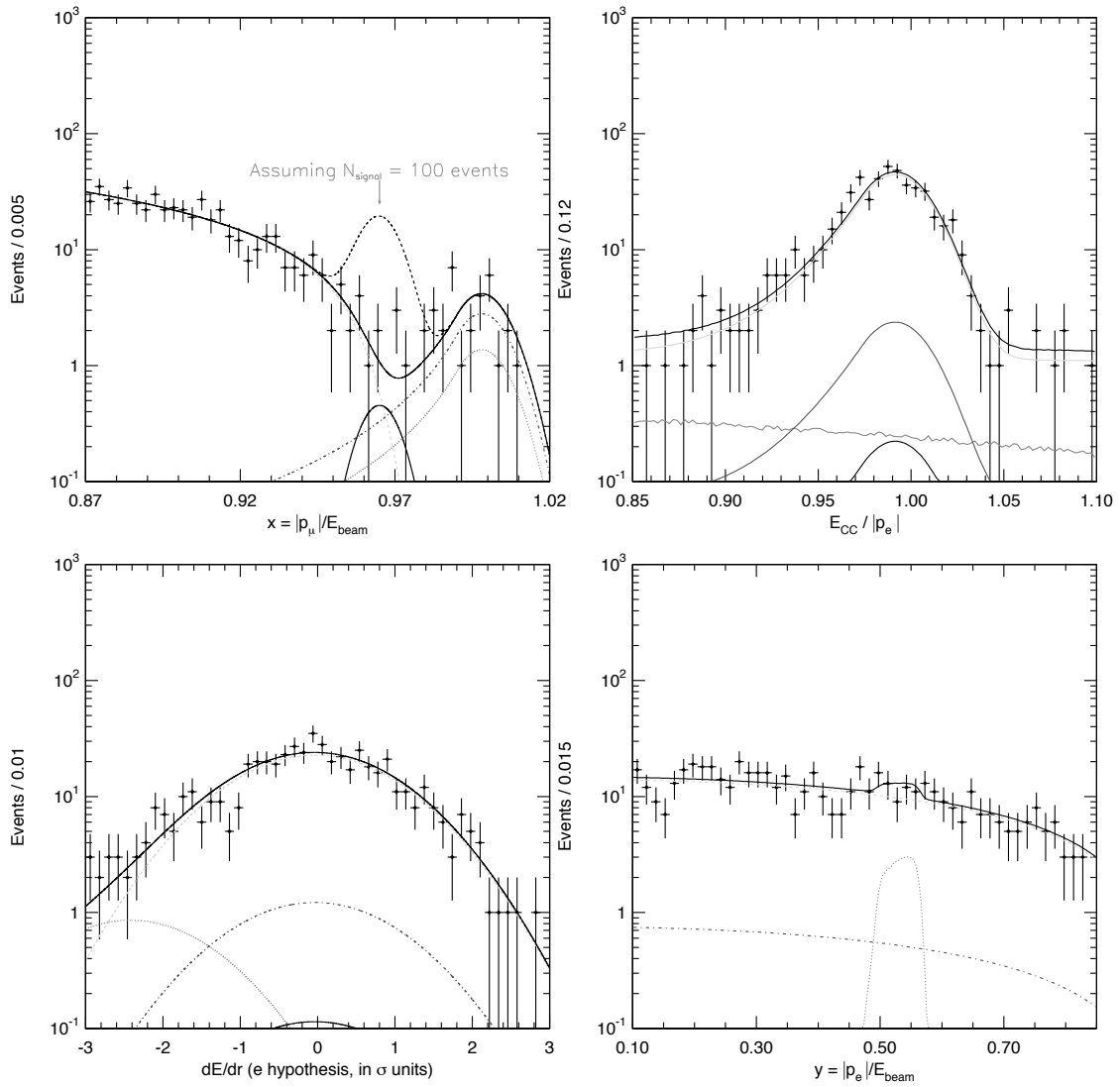


Figure 33: The results of a four-component fit to the $\Upsilon(1S)$ signal data.

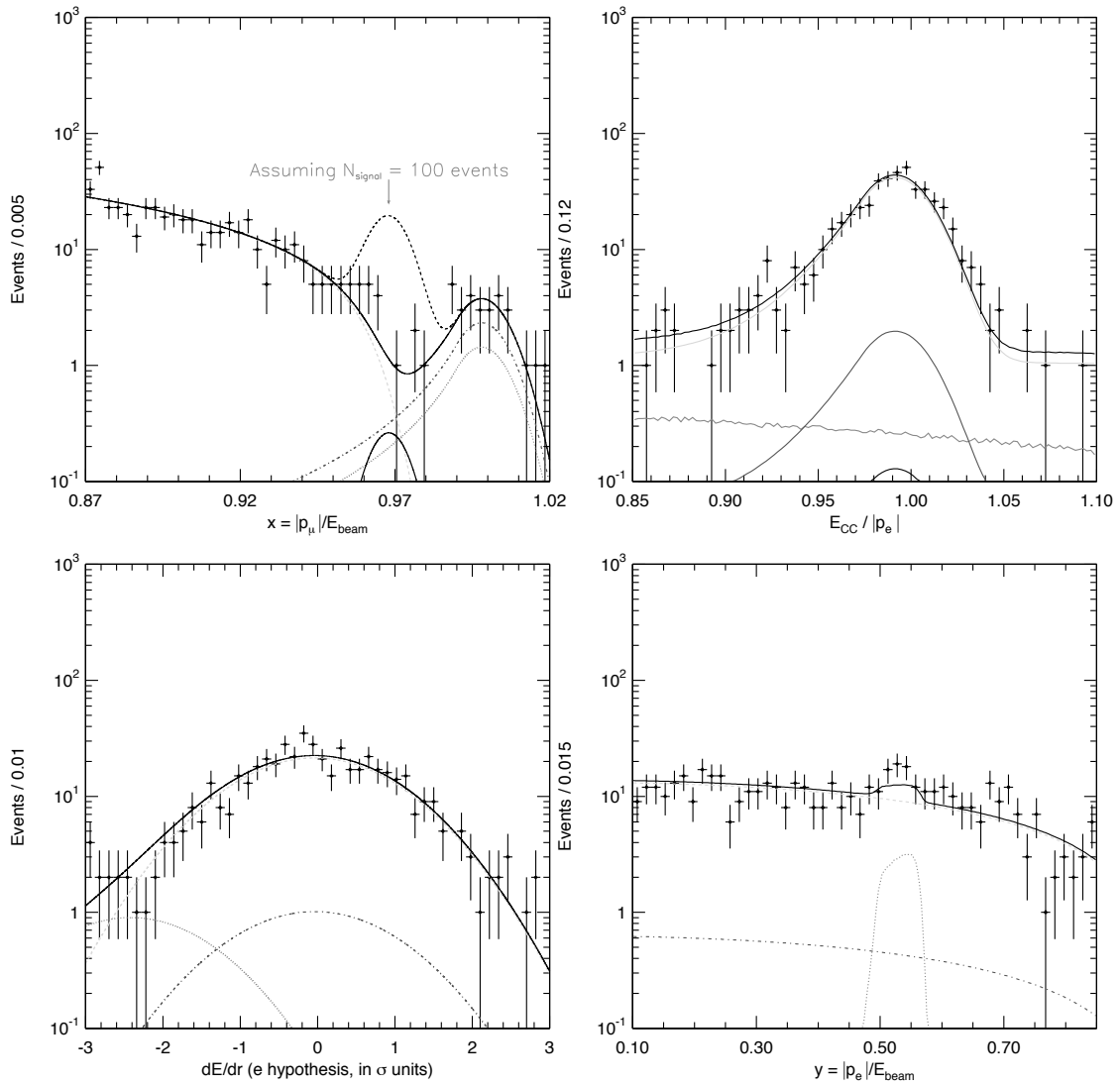


Figure 34: The results of a four-component fit to the $\Upsilon(2S)$ signal data.

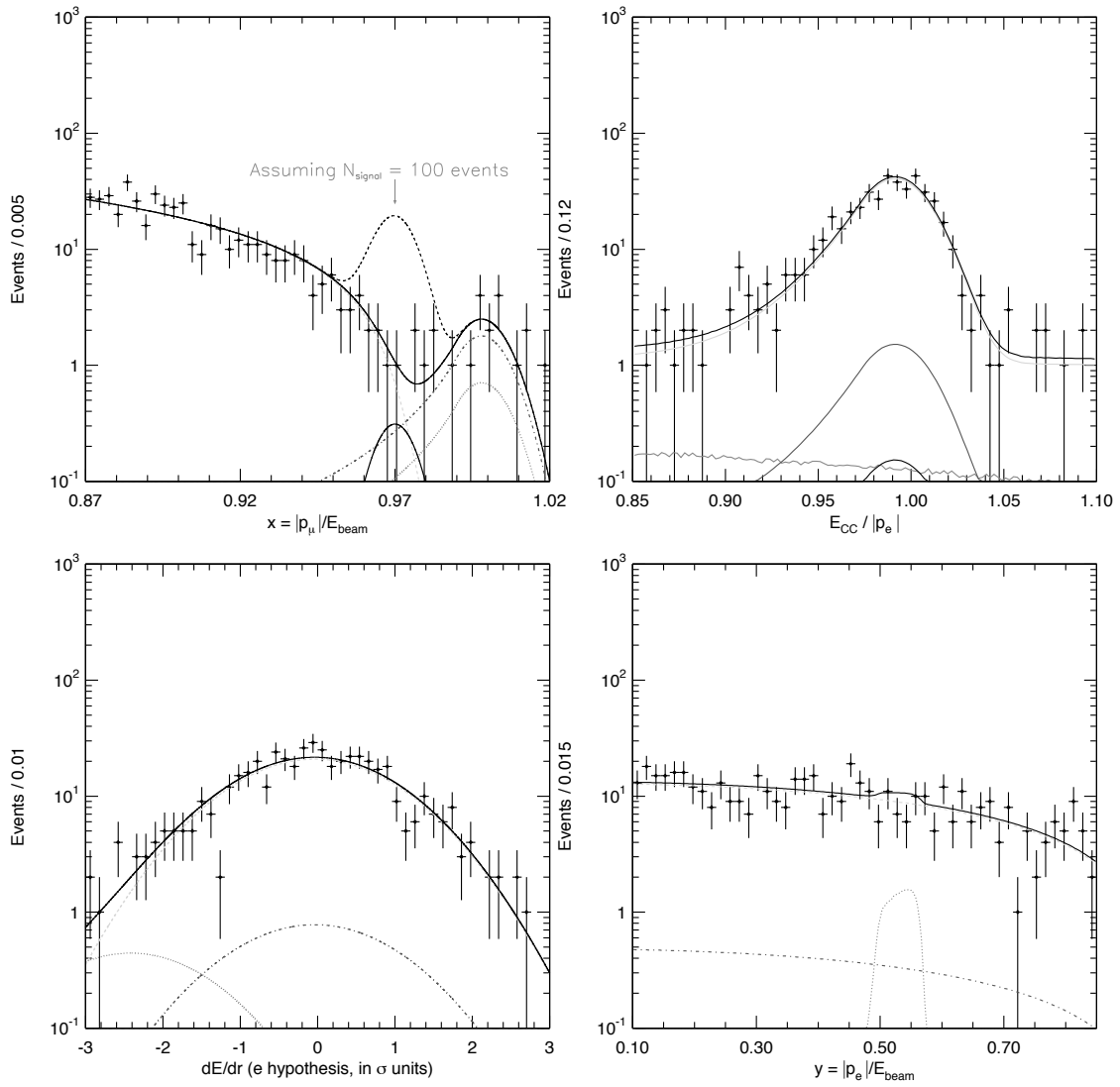


Figure 35: The results of a four-component fit to the $\Upsilon(3S)$ signal data.

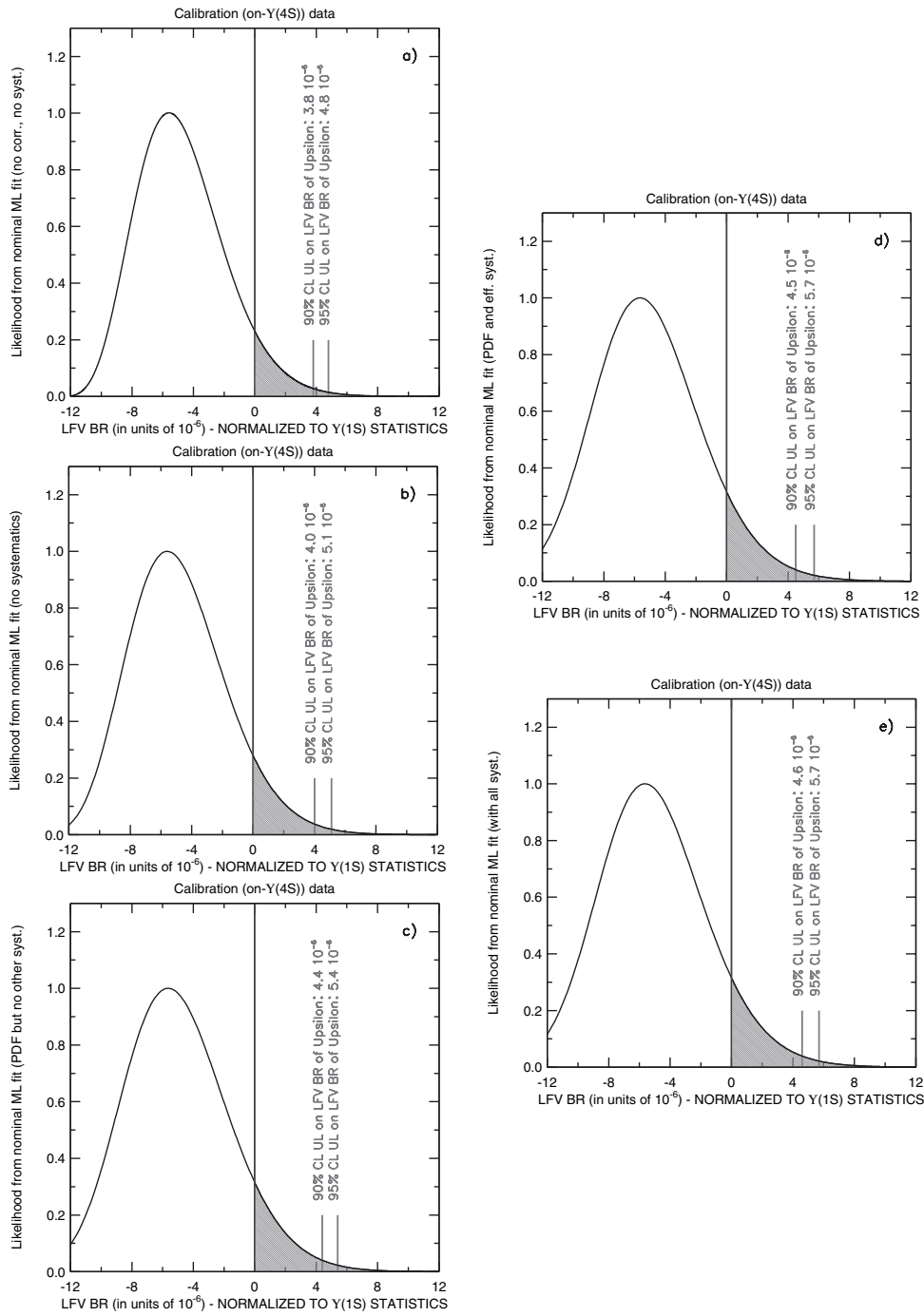


Figure 36: Likelihood profiles *vs.* LFV branching fraction for $\Upsilon(4S)$ calibration data.

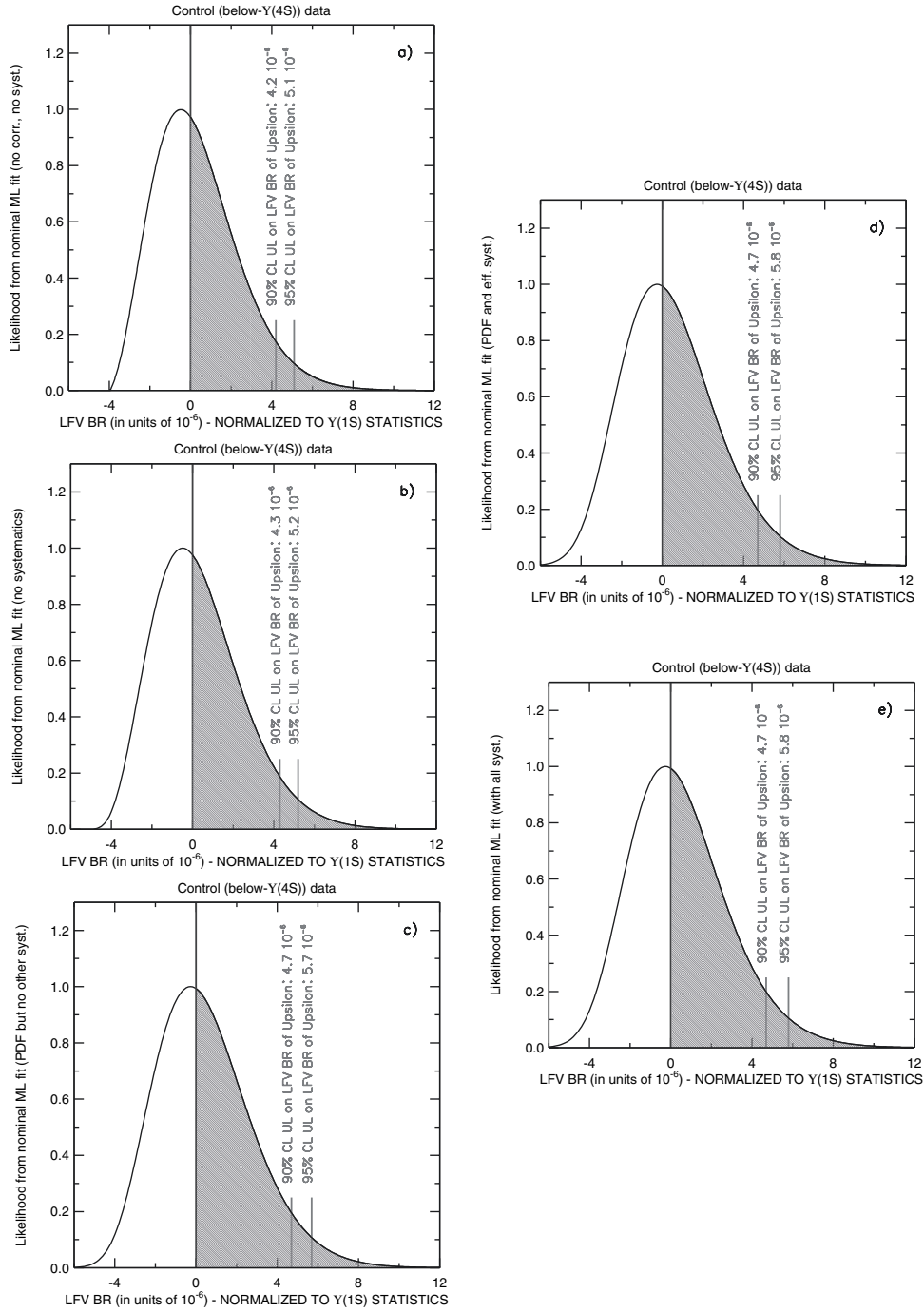


Figure 37: Likelihood profiles *vs.* LFV branching fraction for continuum control data.

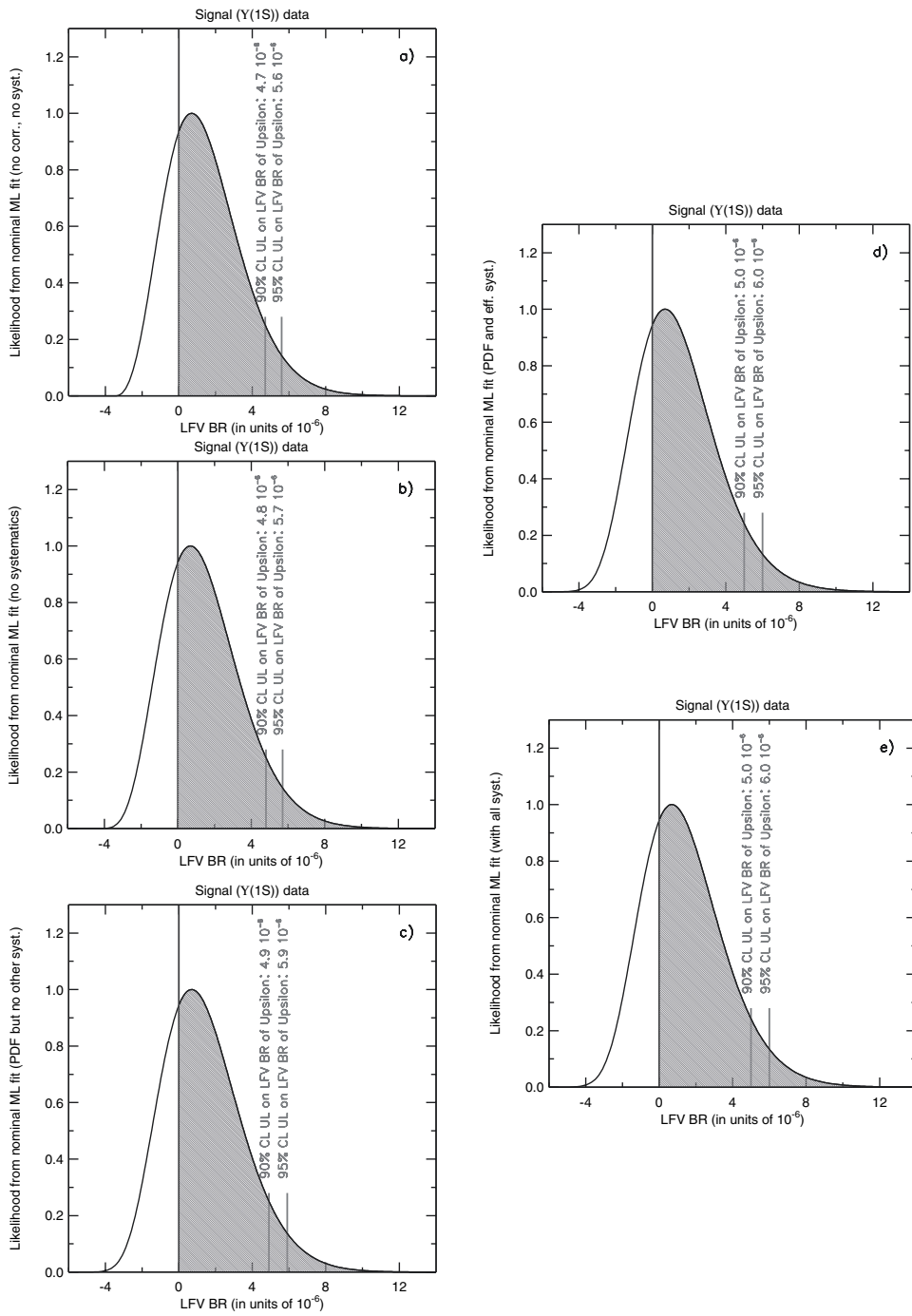


Figure 38: Likelihood profiles *vs.* LFV branching fraction for $\Upsilon(1S)$ signal data.

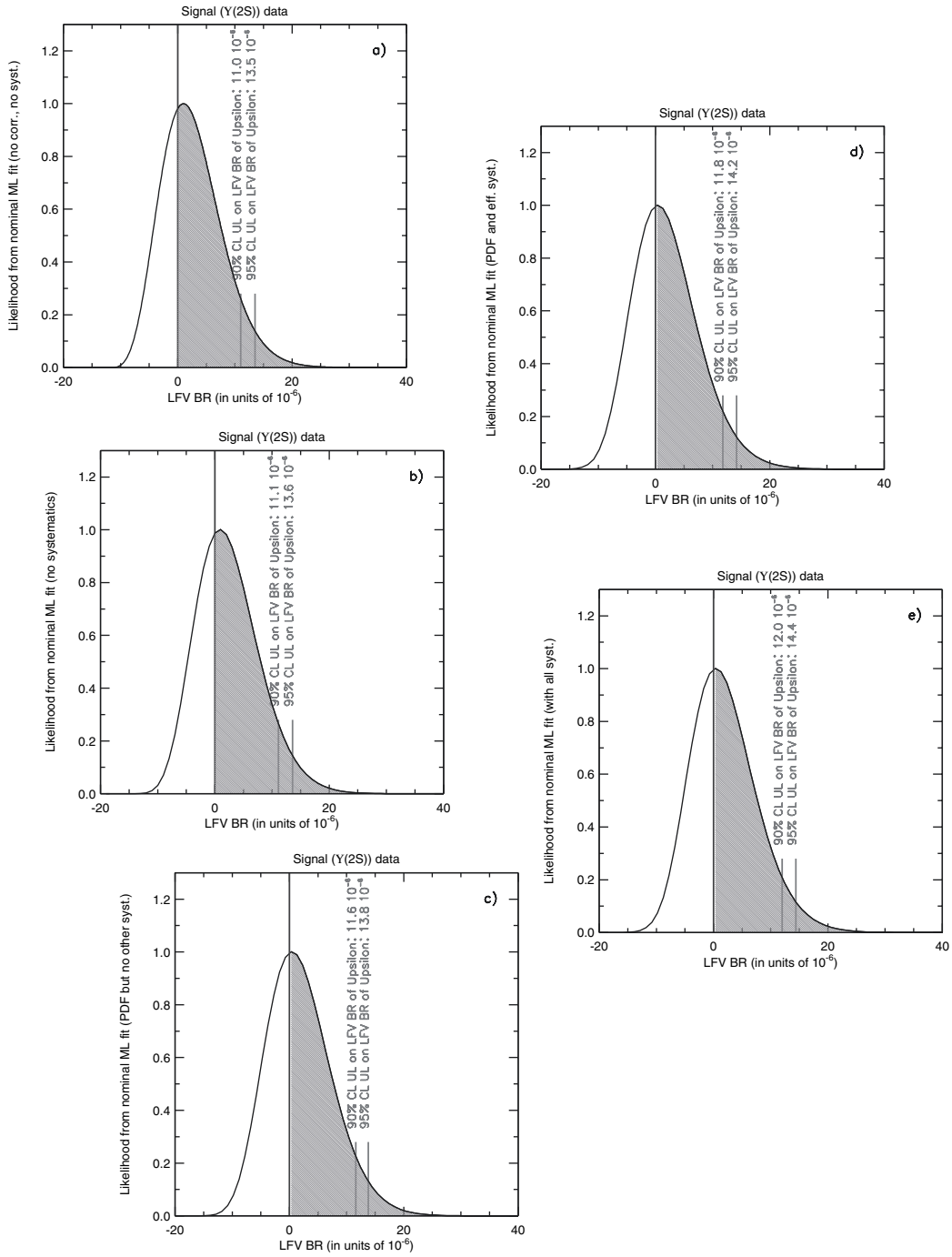


Figure 39: Likelihood profiles *vs.* LFV branching fraction for $\Upsilon(2S)$ signal data.

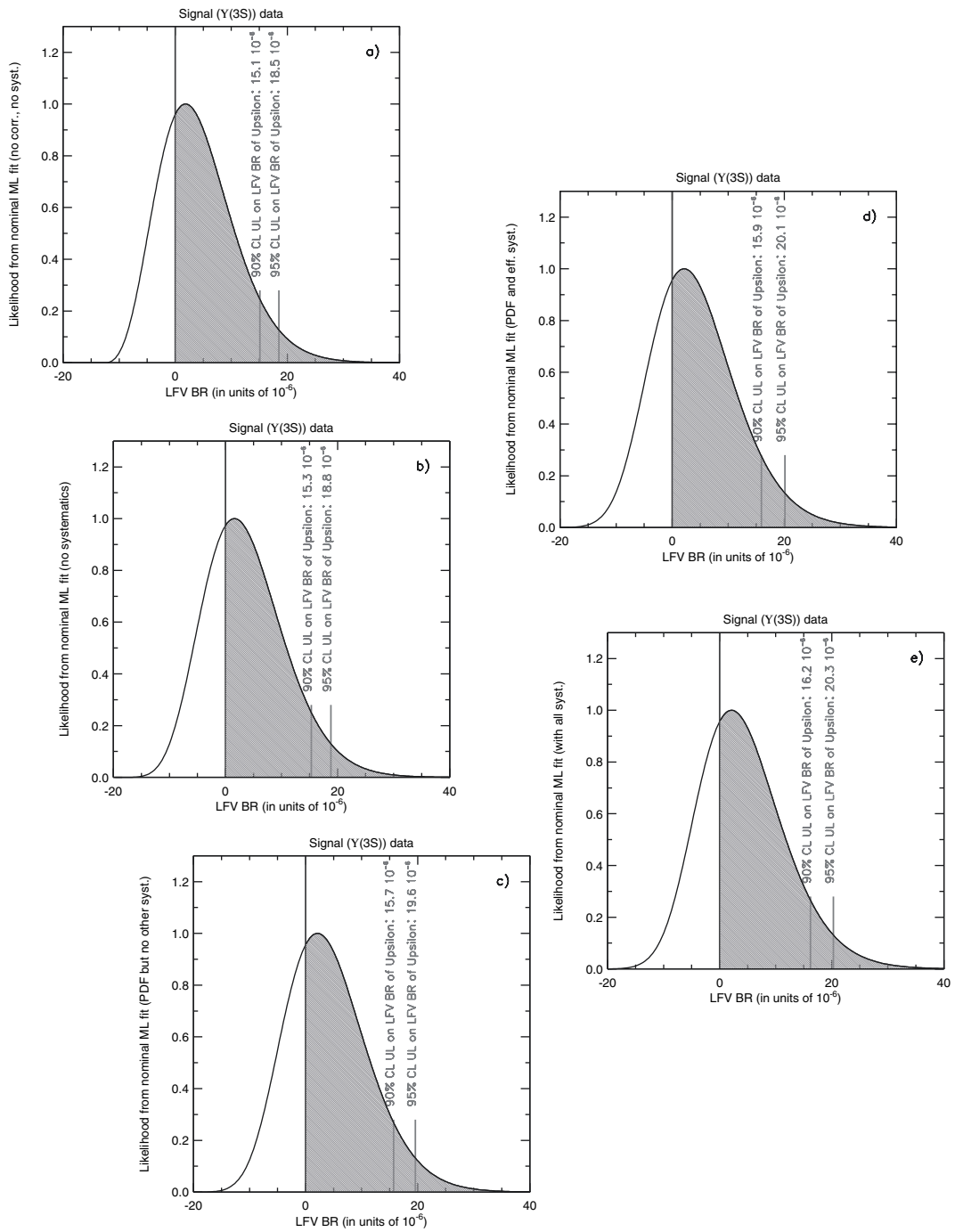


Figure 40: Likelihood profiles *vs.* LFV branching fraction for $\Upsilon(3S)$ signal data.

BIBLIOGRAPHY

- [1] M. Ablikim *et al.* (BES Coll.), Phys. Lett. B **598**, 172 (2004).
- [2] B. Aharmim *et al.* (SNO Coll.), Phys. Rev. C **72**, 055502 (2005).
- [3] M. Artuso *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **554**, 147 (2005).
- [4] M. Artuso *et al.* (CLEO Coll.), Phys. Rev. Lett. **94**, 032001 (2005).
- [5] R. Barlow, *Statistics*. New York: John Wiley & Sons (1989).
- [6] K. Berkelman, *A Personal History of CESR and CLEO*. Singapore: World Sci. (2004).
- [7] D. Besson *et al.* (CLEO Coll.), Phys. Rev. D **74**, 012003 (2006).
- [8] D. Black *et al.*, Phys. Rev. D **66**, 053002 (2002).
- [9] A. Bornheim *et al.* (CLEO Coll.), Phys. Rev. Lett. **93**, 241802 (2004).
- [10] D. Bortoletto *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **320**, 114 (1992).
- [11] J. Gaiser, PhD thesis, Stanford U., [Report No. SLAC-255, 1982 (unpublished)].
- [12] D. Griffiths, *Intro. to Elementary Particles*. New York: John Wiley & Sons (1987).
- [13] Y. Kubota *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **320**, 66 (1992).
- [14] W. Love *et al.* (CLEO Coll.), Phys. Rev. Lett. **101**, 201601 (2008).
- [15] S. Nussinov *et al.*, Phys. Rev. D **63**, 016003 (2000).
- [16] D. Perkins, *Intro. to High Energy Physics*. Menlo Park, Calif.: Addison-Wesley (1987).
- [17] V. Savinov, PhD thesis, U. of Minnesota (1996).
- [18] Z. Silagadze, Phys. Scr. **64**, 128 (2001).
- [19] G. Viehhauser, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 146 (2001).
- [20] W.-M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006).