SLAC-R-529
UC-410, UC-414

A Study of High Field Quantum Electrodynamics in the Collision of High Energy Electrons with a Terawatt Laser*

Glenn A. Horton-Smith

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-529
July 1998

Prepared for the Department of Energy
under contract number DE-AC03-76SF00515

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

*Ph.D. thesis, Stanford University, Stanford, CA 94309.

Abstract

An experiment is described which studied quantum electrodynamic interactions under conditions of extremely high fields, along with a review of the relevant theory. The high fields were created by an intense, tightly-focused pulse of laser light at green or infrared wavelengths, into which was sent an ultra-relativistic electron beam of 46.6-GeV energy. The relevant theory is that of an electron in an electromagnetic wave so intense that the electron's mass is effectively shifted by the transverse momentum imparted to it by the wave, and the electron encounters field strengths comparable to the Schwinger critical field strength of 511 kV per Compton wavelength. An electron in the intense wave may radiate a photon and balance 4-momentum by absorbing multiple photons from the laser, which can lead to real photons with energies above the kinematic limit for conventional Compton scattering. All particles have significant probability of scattering multiple times while in the focus of the laser, including the photons radiated by the electrons, which may convert into electron-positron pairs, again with absorption of multiple photons from the laser. This experiment was able to measure the rates and spectra of positrons, electrons, and photons emerging from the interaction region. Results from both experiment and theoretical simulations are presented and compared. The results from the electron and positron measurements are compatible with the accepted theory, within experimental uncertainties due mainly to the laser intensity measurement. The photon spectrum shows the correct shape, but the ratio of rates in the linear and two-absorbed-photon portions of the spectrum does not vary as expected with the laser intensity, suggesting a disagreement with the accepted theory, with a significance of roughly two standard deviations. A follow-up experiment would be in order.

Acknowledgments

Many people contributed to making this experiment a success. Everyone in the E-144 collaboration deserves collective and individual thanks for supporting an atmosphere of productivity, inventiveness, friendship, and mutual support. The experiment would not have been possible without the collaboration, nor would it have been possible without DOE contract DE-AC03-76SF-00515, which supports SLAC (and thus me) financially.

Both the collaboration as a whole and myself in particular benefited greatly from the insight, high standards, and calm voice of reason of my advisor, David Burke.

I wish to thank R. Clive Field and T. Koffas for much helpful dialog and useful information, particulary in discussing Cherenkov monitor issues.

I'm grateful for the many helpful suggestions and productive discussions with Christian Bula and Kostya Shmakov regarding ECAL reconstruction, and I also thank K. Shmakov for useful discussions regarding the effect of laser polarization.

I would also like to thank Peter Tennenbaum and Thomas Koffas for their excellent work with the beam optics, Robert Holtzapple for providing information on bunch length versus RTL compressor setting, Christian Bula for creating the fine E-144 data acquisition infrastructure, and all who have contributed to making this experiment a success.

I would like to thank the SLC Controls Software Group for providing needed assistance and for allowing the rather non-standard, user-written `E144DAQ` code to exist and to manipulate the control system. Nancy Spencer was particularly helpful, in terms of technical advice and support, and in granting the code special status outside of the normal controls software bureaucracy by putting it in the special "DWIA" (dies with its author) category. I feared that the converse (ADWI) might also apply: I am pleased that E-144 has been successfully completed, and neither the code nor myself had to be prematurely terminated.

I'd like to thank the beam diagnostic group for lending hardware, providing technical support, and sharing their knowledge. In particular, the help of Sandy Horton-Smith in commisioning and calibrating two LVDT position monitors and in measuring ground motion using a laser position transducer is much appreciated.

Sandy, my wife, also provided an immeasurably large amount of encouragement and emotional support, not to mention earning a relatively large amount of financial support for our family. When her position at SLAC was unexpectedly terminated, she demonstrated her great courage and vast skills by immediately obtaining a higher-paying job in Silicon Valley. She is a loving mother, a good provider, a true friend, and a full partner in my life. I couldn't have done this without her.

Conventions, units, and notation

Units

Generally, "natural units" where $\hbar = c = 1$ will be used, except when it is necessary to convert to "laboratory units," *e.g.*, for generating actual rate predictions. Laboratory units will be Rationalized MKSA. (But magnetic fields will be specified in both gauss and tesla, just to make everyone happy.)

The convention for electric charge when working in "natural units" is that the electron charge squared is $e^2/(4\pi) \cong 1/137$. This is like Rationalized MKSA with $\epsilon_0 = \mu_0 = 1$. This convention could be stated concisely as $Z_0 = 1$, where Z_0 is the impedance of free space.

Relativity

Conventions are the same as Peskin and Schroeder, Jackson, and Berestetskii *et al.*.

Four-dimensional tensor indices are denoted by Greek letters, taking the values 0,1,2,3. Three-dimensional indices are denoted by Roman letters, taking the values 1,2,3.

$k^\mu x_\mu = g^{\alpha\beta} k_\alpha x_\beta = k_0 x_0 - \sum_i k_i x_i = k_0 x_0 - \vec{k} \cdot \vec{x}$.

$g^{\alpha\beta} = g_{\alpha\beta} = \mathrm{diag}(1, -1, -1, -1)$.

The displacement 4-vector x^μ is "naturally raised," *i.e.*, for any given reference frame, $x^i = \hat{e}_i \cdot \vec{x}$, $x^\mu = (x^0, \vec{x})$, $x_\mu = (x^0, -\vec{x})$, $x^0 = x_0$.

The derivative operator $\partial_\mu \equiv \partial/\partial x^\mu = (\partial/\partial t, \vec{\nabla})$ is "naturally lowered."

The momentum 4-vector is naturally raised. The momentum operator on Schrodinger wavefunctions is $\hat{p}^\mu = (E, \vec{p}) = i(\partial/\partial t, -\vec{\nabla}) = i g^{\mu\nu} \partial_\nu$. This can be conveniently written $p^\mu = i\partial^\mu$, but some people find it confusing if indices aren't kept in their "natural" positions as much as possible.

$(kx) \equiv k^\mu x_\mu$ if k and x are 4-vectors. $(x)^2 \equiv (xx) \equiv x^\mu x_\mu$. To give the reader a fighting chance, I solmenly promise to never omit the parentheses around a scalar product of 4-vectors, and to never assign the same letter to both a vector and a scalar quantity.

$\not{A}, \not{k}, \not{p}$ denote $A^\mu \gamma_\mu, k^\mu \gamma_\mu, p^\mu \gamma_\mu$, where γ^μ are the Dirac matrices.

Contents

List of Tables

List of Figures

Chapter 1

Study of Nonlinear QED in Intense Laser-High Energy Electron Collisions

Synopsis

The E-144 experiment has measured the products of nonlinear Compton scattering and Breit-Wheeler pair production occurring in the collision of ultra-relativistic electrons with a tightly-focused, terawatt pulse of laser light. The light was intense enough to impart near-relativistic transverse momenta to any electrons and positrons within it, and the electric field strength in the electron's rest frame reached to within a factor of three of the Schwinger critical field $m^2/e \sim 10^{18}$ V/m. Understanding quantum electrodynamic (QED) interactions in such intense fields is important for future collider interaction point, positron source, and photon source design, and also relevant to certain problems in astrophysics and relativistic heavy ion collision experiments. Here, the theory, experiment, and previously-published results pertaining to the scattered electrons[1, 2, 3] and created positrons[4, 5, 6] are reviewed to establish a context, and new results regarding the non-linear Compton photons and two critical monitors of the electron spectrum are presented in that context. The prior results are found to be in agreement with theoretical calculations, within quoted uncertainties and using previous assumptions; however, the new results apparently contradict these assumptions and lead to an apparent overall irreconcilability of the electron, positron, and photon data with the accepted theory and model of the experiment. Possible systematic effects are discussed. It is concluded that the possible systematic effects may not be sufficient to explain the results, but that the significance of the discrepancy, if any, does not exceed two standard deviations. The experiment has made valuable contributions to the study of QED: in particular, the photon data unambiguously demonstrates the second-harmonic Compton scattering process at high field strength. A follow-up experiment would be in order.

1.1 Overview

The theory of the interaction of electrons with an intense electromagnetic plane wave was developed many years ago[7, 8, 9, 10, 11]. The theory suggests two different ways of characterizing a wave as "intense": (1) high root-mean-square (rms) field strength of the wave as seen by the electron, and (2) high rms vector potential of the wave in the Lorentz gauge, a measure of the magnitude of the integral of the field strength over time. When criterion 1 is satisfied, the electron undergoes high transverse acceleration, and when criterion 2 is satisfied, high transverse momentum is imparted to the electron by the wave. The two criteria are not mutually exclusive, but usually neither one

is satisfied. Only the low field, high transverse momentum limit of this theory had previously been studied in the interaction of low-energy electrons with light from terawatt lasers[12].

The E-144 experiment, completed recently at the Stanford Linear Accelerator Center, is the first to study electrons in an electromagnetic wave with high field strength *and* high potential.[1] The high potential results in significant "nonlinear" Compton scattering, in which electrons effectively scatter off the n-th harmonic of the incident wave. The high field strength results in measurable Breit-Wheeler production of electron-positron pairs by backscattered photons interacting with the intense wave. Using ω to represent a quantum of (the first harmonic of) the intense wave and γ to represent any other photon, the two reactions studied are

$$e + n\omega \rightarrow e + \gamma \tag{1.1}$$

$$\gamma + n\omega \rightarrow e^+ + e^-. \tag{1.2}$$

1.1.1 Theory

Quantities characterizing the strength of the electromagnetic wave

Two dimensionless, Lorentz-invariant parameters characterize the intensity of the wave. They are defined as follows:

$$\eta \equiv \frac{|e|}{m}\sqrt{\langle -(A^2)\rangle} \tag{1.3}$$

$$\Upsilon \equiv \frac{|e|}{m^3}\sqrt{\langle -(p^\mu F_{\mu\nu})^2\rangle}, \tag{1.4}$$

where A is the 4-vector potential in the Lorentz gauge, F is the usual electromagnetic field tensor, e and m are the electron charge and mass, p is the 4-momentum of the electron, $(p^2) = m^2$, and units in which $c = \hbar = 1$ are employed. The parameter Υ characterizes the wave's strength based on the field strength in the electron's rest frame, and η characterizes the strength of the wave based on its 4-vector potential.

The parameter Υ depends on the momentum of the electron and the field of the wave, while η is a characteristic of the wave alone. Since $E^\star_\nu = (1/m)p^\mu F_{\mu\nu}$ is the electric field in the electron's rest frame, Υ is the ratio of the rms electric field strength to the "critical field" in which an electron's energy would increase by its rest mass over a Compton wavelength:

$$\Upsilon = \frac{E^*_{\rm rms}}{E_{\rm crit}} \tag{1.5}$$

$$E_{\rm crit} \equiv \frac{m^2}{e} \cong \frac{(511\ {\rm kV})}{\lambda\!\!\!^{-}_c} \cong 1.32 \times 10^{18}\ {\rm V/m}, \tag{1.6}$$

where $\lambda\!\!\!^{-}_c \cong 3.86 \times 10^{-13}$ m is the Compton wavelength of the electron.

For a plane wave with wavevector k^ν, the 4-vector potential and field tensor of the wave are a function of a single scalar $\phi \equiv (kx)$, so that

$$F^{\mu\nu} = k^\mu A^\nu - k^\nu A^\mu \tag{1.7}$$

$$(p^\mu F_{\mu\nu})^2 = (kp)^2(A^2) \tag{1.8}$$

$$\Upsilon = \frac{(kp)}{m^2}\eta. \tag{1.9}$$

[1] Crystal channeling experiments[13] have studied electrons in periodic fields imparting large transverse acceleration and momentum, but the channeling of an electron by a fixed lattice is presumably different from the propagation of an electron in a wave.

Thus, for a given k^μ and p^μ, there is a direct proportionality between η and Υ. For a monochromatic plane wave such as $A^\mu(\phi) = A_1^\mu \cos\phi + A_2^\mu \sin\phi$ (where A_1^μ and A_2^μ are constant 4-vectors satisfying $(kA) = 0$) the quantity $\omega^* \equiv (kp)/m$ is the angular frequency of the wave in the electron's rest frame. Therefore, η may also be expressed as a ratio of electric field strengths: it is the ratio of the rms electric field strength to the field in which an electron's energy would increase by its rest mass over a distance $1/\omega$:

$$\eta = eE^*_{\rm rms}/m\omega^* = eE_{\rm rms}/m\omega. \tag{1.10}$$

Because η is a property of the wave alone, η may be evaluated using the above equation with electric field E and frequency ω measured in any convenient frame.

A "free" electron in an intense electromagnetic wave: the Volkov wavefunction and quasimomentum

The wavefunction of a spin-1/2 particle in an external electromagnetic field obeys Dirac's equation. In the case where the external field is a plane wave of arbitrary shape and waveform, the exact solution to the Dirac equation was found by Volkov[14]; a clear and accessible derivation is given in section 40 of Berestetskii *et al* [11]. The result is

$$\psi_{{\rm V};p}(x) \quad = \quad \left(1 + \frac{e\not{k}\not{A}_{\rm ext}}{2(kp)}\right) u e^{-i(px)+iS(\phi)}, \tag{1.11}$$

where u is a Dirac spinor satisfying $(\not{p} - m)u = 0$, and S is the part of the classical action due to the external field, given by

$$S(\phi) \quad \equiv \quad \int_{-\infty}^{\phi} d\phi' \left(\frac{e^2}{2(kp)}(A_{\rm ext}(\phi'))^2 - \frac{e(pA_{\rm ext}(\phi'))}{(kp)}\right). \tag{1.12}$$

As mentioned previously, the wave must be a plane wave, $A = A(\phi)$, where $\phi = (kx)$, and A must be expressed in the Lorentz gauge, $(\partial A) = (kA) = 0$. The term containing $\not{k}\not{A}$ acting on the spinor u represents the effect of the external field on the spin orientation of the particle, and S produces a "frequency modulation" of the electron wavefunction as regions of various potential in the external wave pass by.

In any region where A is zero, the Volkov wavefunction reduces to simply $\exp(-i(px))u$, the correct solution to the Dirac equation in the case of no external field. Thus, p^μ is the 4-momentum of the electron "outside of the wave," *e.g.*, for a pulsed plane wave of finite longitudinal extent, it is the particle's momentum before arrival of the wave and after its departure from the vicinity of the electron.

The average momentum of the electron in the intense electromagnetic wave, the so-called "quasi-momentum," is

$$\begin{aligned} q^\mu \equiv \langle \bar\psi_{\rm V} \hat p^\mu \psi_{\rm V} \rangle \quad &= \quad p^\mu - \langle S' \rangle k^\mu + i k^\mu \frac{e}{2(kp)} \bar u \not{k} \langle \not{A}' \rangle u \\ &= \quad p^\mu + \frac{m^2}{2(kp)}\eta^2 k^\mu + \left\langle 2e(pA_{\rm ext}(\phi)) + \frac{ie}{4} F_{\lambda\nu} \bar u \sigma^{\lambda\nu} u \right\rangle k^\mu \end{aligned} \tag{1.13}$$

If the time-averaged value of $A_{\rm ext}(\phi)$ is zero, then

$$q^\mu \quad = \quad p^\mu + \frac{m^2}{2(kp)}\eta^2 k^\mu. \tag{1.14}$$

The square of the quasi-momentum is called the "effective mass":

$$m_*^2 \equiv (q)^2 = (1+\eta^2)m^2. \tag{1.15}$$

The physical significance of the quasi-momentum q^μ and the effective mass m_* will be seen below, where it is shown that the kinematics of scattering involve q^μ rather than p^μ, and that a center-of-mass energy of at least $2m_*$ is needed to produce e^+e^- pairs in photon-photon interactions.

It is interesting to note that the quasi-momentum is simply the "external" momentum p^μ plus a scalar times the wavevector k^μ. One might define

$$\bar{n} \equiv \frac{m^2}{2(kp)}\eta^2 \tag{1.16}$$

whereupon one might ascribe to $\bar{n}$ a physical interpretation in terms of a number of photons. It would then be tempting to expect $\bar{n}$ to be quantized, or else to rewrite Eq. 1.14 as

$$q^\mu = p^\mu + n_? k^\mu, \tag{1.17}$$

where $n_?$ is an integer giving the number of photons "carried by" the electron, whose average value would be given by the semi-classically-derived value $\bar{n}$. (For the E-144 experiment, $\bar{n} \simeq 1.2\eta^2$ or $0.6\eta^2$, depending on the laser wavelength used.) Such a quantization of $\bar{n}$ would have a significant effect on the kinematics of pair production, for example.

In fact, as we shall shortly see, no such quantization of $\bar{n}$ (or $n_?$) is suggested or allowed by the standard theory of high-field quantum electrodynamics. This is not surprising since the external field is treated as an entirely classical object, not something composed of photons or otherwise quantized. Indeed, since the waveform is entirely arbitrary, there is no reason to associate k^μ with the momentum of a "photon" of the external wave: for example, even for a monochromatic wave, one could choose $A^\mu(\phi) = A_1^\mu \cos 2\phi + A_2^\mu \sin 2\phi$, thus changing the values of k and $\bar{n}$ by a factor of 2 for the same physical situation. It can be readily verified that this leaves all physical quantities such as $\eta, \Upsilon, q^\mu, m_*^2$, all scattering rates, and even the wavefunction $\psi_{V;p}(x)$ completely unchanged. Furthermore, "quantizing" the quasimomentum as in Eq. 1.17 results in a wavefunction that does not satisfy Dirac's equation: it is *not* the correct way to account for quantization of the external field.[2]

Prospects for experimentally discerning whether the effective mass shift is quantized or unquantized are discussed in a later section.

Finally, the normalization of the wavefunctions must be considered. In a given reference frame, the wavefunctions satisfy the orthogonality and normalization condition

$$\int \psi_{p'}^* \psi_p d^3x = \int \bar{\psi}_{p'} \gamma^0 \psi_p d^3x = (2\pi)^3 \delta(\vec{p'} - \vec{p}) \frac{\bar{u'}\gamma^0 u}{2E_p}, \tag{1.18}$$

where $E_p = \sqrt{m^2 + |\vec{p}|^2}$. If desired, the normalization of the spinors u can be chosen such that $\bar{u}^s \gamma^0 u^r = 2E_p \rho_p \delta_{rs}$, in which case the electron has probability density ρ_p as measured in the chosen

[2] An informal survey of theorists suggests that they break into two camps on how to correctly account for quantization of the external field: those who feel that the semi-classical approach is entirely correct even at $\bar{n} \lesssim 1$, and therefore have not attempted it; and those who feel a quantized treatment is necessary, but have not succeeded in deriving one.

reference frame when it is outside of the wave. In general, the current density of the electron wavepacket is given by

$$j^\mu = \bar{\psi}_p \gamma^\mu \psi_p = \frac{\rho_p}{E_p} \left[p^\mu - eA^\mu + \bar{n} k^\mu \right]. \tag{1.19}$$

Note that in many texts, the wavefunction is normalized to have an average spatial density of unity "inside the wave,"[3] whereas for a practical colliding-beam experiment, with detectors and beam sources outside of the wave, it would seem to be more reasonable to normalize to unit density *outside* of the wave. The convention adopted above is flexible enough to accommodate either normalization, and has the advantage of being Lorentz invariant (although not manifestly so), since ρ_p/E_p is an invariant scalar for a plane wave. Clearly, the choice of normalization has no effect on the physics. The significant point here is that the density of an electron wavepacket changes as it enters and exits the wave. However, this actually turns out to be a very small effect as viewed in the laboratory frame of this experiment, where $p^0 \gg k^0$.

Quantum processes in an intense electromagnetic wave: kinematics and dynamics

Although it is beyond the scope of this section to prove it, the scattering matrix element for an electron with initial momentum p to emit electromagnetic radiation $A'^\mu = \epsilon'^\mu e^{-i(k'x)} \sqrt{\rho_{k'}/(2E_{k'})}$ while in an intense electromagnetic wave (with the electron emerging with final momentum p') is given by [7, 8, 9]

$$\langle p' \,|S|\, p \rangle \;=\; -ie \int d^4x \, \bar{\psi}_{\mathrm{V};p'}(x) \gamma^\mu \psi_{\mathrm{V};p}(x) A'^*_\mu. \tag{1.20}$$

When written in terms of an invariant amplitude $\mathcal{M}$, this matrix element is

$$\begin{aligned} \langle p' \,|S|\, p \rangle \;=\; & \sum_n i\mathcal{M}_n(-p, p', k') \sqrt{\frac{\rho_p}{2E_p}} \sqrt{\frac{\rho_{p'}}{2E_{p'}}} \sqrt{\frac{\rho_{k'}}{2E_{k'}}} \\ & \times (2\pi)^4 \delta^{(4)}(q + nk - q' - k') \end{aligned} \tag{1.21}$$

The delta functions come from the integral of the product of oscillatory terms in the wavefunction. The result has been interpreted as expressing conservation of the quasi-momentum of the electron upon absorption from the wave of n photons with momentum k and emission of the single photon with momentum k' [10, 11]. Noting that the arguments of the delta functions may be rewritten as

$$q + nk - q' - k' \;=\; p + (n + \bar{n} - \bar{n}')k - p' - k', \tag{1.22}$$

one could also consider the delta function as expressing conservation of the "outside of the wave" momentum upon absorption of $n + \bar{n} - \bar{n}'$ "photons" from the field and emission of a single photon with momentum k', where $\bar{n}$ and $\bar{n}'$ are not required to be integers, according to the accepted theory. Arguably, neither interpretation involving photons in the external wave is particularly valid given that the external wave is never quantized. The fundamental point is that radiation is emitted at the n-th harmonic of the incident wave by an electron whose mass is effectively shifted by the classical action of the external wave, with a generalized Compton effect represented by the delta function in Eq. 1.21 arising as a consequence of the Lorentz invariance of the theory.

[3] This is the case in Berestetskii[11] section 101, but not in section 40.

One useful consequence of the kinematics of the non-linear Compton process is that the end-points of the scattered electron and photon spectra vary with n, such that part of the spectrum at a given n is unobscured by lower-n contributions. Defining the invariant

$$x \equiv (kk')/(kp), \tag{1.23}$$

and defining the center of mass (COM) frame for non-linear Compton scattering at harmonic n as the frame in which

$$\vec{q}_{\rm COM} + n\vec{k}_{\rm COM} = 0, \tag{1.24}$$

we have

$$x = \frac{1-\cos\theta}{2 + m_*^2/(n(kp))}, \tag{1.25}$$

where θ is the angle between the emitted photon and the incident wave in the COM frame, while in the lab frame we have

$$x \cong \frac{E_{k'}}{E_p}. \tag{1.26}$$

Thus, the maximum backscattered photon energy in the lab frame for a given n and η is

$$E_{k',\max} = \frac{E_p}{1 + m_*^2/2n(kp)} \tag{1.27}$$

$$= \frac{E_p}{1 + (1+\eta^2)/(nx_0)}, \tag{1.28}$$

where

$$x_0 \equiv 2(kp)/m^2. \tag{1.29}$$

For E-144, x_0 is of order unity, leading to good separation of the different orders n (cf Fig. 1.8). Discussion of specific values of parameters for this experiment will be deferred until after the description of the experimental arrangement below.

Another interesting feature of the kinematic constraint is that pairs may be produced by high-n processes when they are kinematically forbidden to do so by low-n processes, and the minimum value of n depends on η as well as the incident photon energy. For pair production, the kinematic constraint is $nk + k' = q + q'$, where q and q' are the final-state electron and positron momenta; thus, one must have $(q+q')^2 = (nk+k')^2 = 2n(kk') > m_*^2$.

Finally, the scattering rates: in terms of the matrix element $\mathcal{M}$ (which is very complicated and will not be given here), the differential probability (per unit volume and unit time) of radiation of a photon from the n-th harmonic is given by

$$dW_{{\rm Compton}(n)} = \left|\mathcal{M}_n(-p,p',k')\right|^2 \frac{\rho_i}{2E_p}\frac{d^3p'}{2E_{p'}(2\pi)^3}\frac{d^3k'}{2E_{k'}(2\pi)^3} \times (2\pi)^4\delta^{(4)}(q+nk-q'-k'). \tag{1.30}$$

Similarly, the probability for the cross-channel process of pair production by a photon in an intense electromagnetic wave is given by

$$dW_{{\rm pair}(n)} = \left|\mathcal{M}_n(p,p',-k')\right|^2 \frac{d^3p}{2E_p(2\pi)^3}\frac{d^3p'}{2E_{p'}(2\pi)^3}\frac{\rho_i}{2E_{k'}} \times (2\pi)^4\delta^{(4)}(k'+nk-q-q'). \tag{1.31}$$

The above expressions allow the scattering rate to be calculated, but they give no direct insight into how the rates vary with η and n. To gain such insight, the above expressions can be compared to the generalized expression for scattering rate involving arbitrary numbers of particles in the initial state:

$$\begin{aligned} dW_{i\to f} &= |\mathcal{M}|^2 \left(\prod_i \frac{\rho_i}{2E_i}\right)\left(\prod_f \frac{d^3p_f}{2E_f(2\pi)^3}\right) \\ &\quad \times (2\pi)^4\delta^{(4)}(\text{kinematic constraint}), \end{aligned} \tag{1.32}$$

where ρ_i is the number density of incident particles of type i, E denotes energy, and p momentum. Each of the four main factors on the right hand side is an invariant. In particular, note that ρ/E, equal to the density of a beam of particles of defined momentum times the reduced wavelength corresponding to the particle energy, is an invariant characterizing the intensity of the beam independently of any reference frame, having the same dimensionality as particles per unit area. The quantity η^2 is directly proportional to this invariant beam surface density, as can be seen by squaring and rewriting Eq. 1.10:

$$\begin{aligned} \eta^2 &= 4\pi(r_e/m)\omega^{-2}I \;\cong\; (6.93\times10^8\ \mathrm{W})^{-1}\lambda\!\!\!^{-2}I \\ &= 4\pi r_e \lambda\!\!\!^{-}_c \hbar c(\rho/E) \;\cong\; (1.37\times10^{-26}\ \mathrm{m}^2)\lambda\!\!\!^{-}\rho, \end{aligned} \tag{1.33}$$

where r_e is the classical radius of the electron, $\lambda\!\!\!^{-} \equiv 1/\omega$ is the reduced wavelength of the wave, I is the intensity of the electromagnetic radiation (*e.g.*, in units of $\mathrm{W/m^2}$), and ρ is the photon number density. Thus, if we were to expand perturbatively in a (quantized) external field, one expects from Eq. 1.32 that a process involving n photons from the intense wave will vary in rate as the n-th power of η^2. In fact, this is true for low η^2, while at high values of η^2 even the "linear" $n = 1$ Compton scattering process begins to show higher-order corrections (cf Fig. 1.11), *e.g.*, a term of order η^4 which can be interpreted as arising from interference with the process in which 2 photons are absorbed from the intense wave and one photon of the same momentum is re-radiated back into the wave.

The differential scattering rates have been explicitly derived in convenient form for several special cases: linearly polarized external wave with unpolarized initial and final states [9]; circularly polarized external wave with unpolarized initial and final states [10]; and both of the above with polarized photons [10, 15, 16][4]. For example, with circularly polarized incident photons the results can be expressed in closed form as follows:

$$\begin{aligned} \frac{dW}{dx} &= \sum_{n=1}^{\infty}\frac{dW_n}{dx} \\ &= \frac{\alpha m^2\rho_i}{16\pi E_p}\sum_{n=1}^{\infty}\left\{-4J_n^2(z)+\eta^2\left(2+\frac{x^2}{1-x}\right)\right. \\ &\quad \left.\times\left(J_{n-1}^2(z)+J_{n+1}^2(z)-2J_n^2(z)\right)\right\} \end{aligned} \tag{1.34}$$

[4] For an extensive discussion of experimental aspects of polarization, see [3].

where

$$z = 2n\frac{\eta}{\sqrt{1+\eta^2}}\sqrt{(u/u_n)(1-u/u_n)}, \quad (1.35)$$

$$u = (kk')/(kp') = x/(1-x), \quad (1.36)$$

$$u_n = 2n(kp)/m_*^2 = nx_0/(1+\eta^2), \quad (1.37)$$

and x and x_0 are as defined previously (Eq. 1.23 and 1.29). The expansion of Eq. 1.34 as a power series in η^2 gives the Klein-Nishina cross-section for the leading order of the $n = 1$ process. It is been shown [17] that the leading and next-to-leading orders can also be calculated using conventional Feynman diagram techniques for the $n = 1$ and $n = 2$ pair production and Compton scattering process, under certain assumptions.

Plots of the scattering rates and specific values for the kinematic edges for the E-144 experimental conditions are given in the section on the simulation, after the description of the experiment below.

1.1.2 Experiment

The experiment has three basic requirements: a controlled and suitably focused source of intense electromagnetic radiation; a controlled and suitably focused source of ultrarelativistic electrons; and appropriate detectors for the final products of the collision. These were provided by the University of Rochester Table Top Terawatt laser [18], the SLAC two-mile accelerator and Final Focus Test Beam[19, 20], and by the E-144 collaboration as a whole, respectively.

Organization of data-taking

A primary determinant of the pattern of data-taking was the linear accelerator schedule. E-144 was allocated approximately one week of time every 9 months from 1994 through 1996. This time was coordinated to be immediately before, after, or interleaved with FFTB beam time. In the April 1994 and September 1994 periods of data-taking, the interleaving approach was used; in later runs, a monolithic block of time was allocated. The chronology of data-taking is summarized in Table 1.1.

Data within each running period was organized by consecutive run file numbers, generally starting with a big, round number such as "12000". Run data files were categorized according to their purpose: scans to achieve beam overlap, collision data, dedicated pedestal measurement runs, and miscellaneous other categories such as "test". A different number of "events"[5] would be acquired in each run data file according to the type of run: up to ten thousand events for a collision data run, and typically only a few hundred for overlap scans. Overlap scans were alternated with data runs. Table 1.2 presents the total number of events acquired in runs of each time for each running period, as a quantitative measure of how much data was collected for each purpose.

The laser in brief

From Eq. 1.10, one sees that η may be made arbitrarily large by maintaining a given field strength while reducing the wave frequency to zero. However, considerable power is required in order to

[5] An "event" is any beam-pulse occurring in a specified pre-programmed pattern, as described in the data acquisition system subsection below: 3 "events" occurred in every 2 seconds.

Table 1.1: Summary of data-taking chronology.

run series	start date	days of data-taking	activities
3000	5-Apr-1994	2.1	checkout, Compton polarimetry
4000	13-Apr-1994	2.6	checkout, Compton polarimetry
5000	27-Apr-1994	3.0	commissioning
6000	10-May-1994	4.0	first high-power laser data
7000	5-Sep-1994	1.7	some non-linear Compton ($e-$) data
8000	12-Sep-1994	2.0	some non-linear Compton ($e-$) data
9000	17-Sep-1994	1.8	some non-linear Compton ($e-$) data
12000	26-Mar-1995	5.8	first XT scans; high-quality non-linear Compton ($e-$) data; commissioned CCD spectrometer; first hint of positrons
13000	8-Dec-1995	6.0	sought more positron and CCD data; frustrated by hidden obstruction in pipe
15000	2-Aug-1996	9.5	$n = 2, 3$ Cherenkov monitors commissioned; high-quality $e-$, $e+$, and CCD photon data

Table 1.2: Summary of events acquired in each run series.

run series	start date	days of data-taking	number of recorded events in runs of each type			
			overlap scans[a]	collision data	pedestal[b]	other
3000	5-Apr-1994	2.1	0	21692	8088	9869
4000	13-Apr-1994	2.6	0	38603	21352	16522
5000	27-Apr-1994	3.0	26256	27068	10994	31725
6000	10-May-1994	4.0	39123	141522	29643	13697
7000	5-Sep-1994	1.7	53476	49104	7657	53885
8000	12-Sep-1994	2.0	55329	206047	10002	8459
9000	17-Sep-1994	1.8	32999	121032	21385	4071
12000	26-Mar-1995	5.8	150274	191880	11045	39000
13000	8-Dec-1995	6.0	58641	199694	18966	74512
15000	2-Aug-1996	9.5	122978	292517	3631	107188

[a] Overlap tuning in the 3000 and 4000 runs was done manually.

[b] The "pedestal" column counts only events in dedicated pedestal runs; in all runs of other types, approximately one event in twenty had the electron beam suppressed to provide mid-run pedestal data.

obtain high η in any propagating field remotely resembling a plane wave. This may be seen from Eq. 1.33, reproduced below:

$$\eta^2 \quad = \quad 4\pi(r_e/m)\lambda\!\!\!^{-2} I \quad \cong \quad (6.93 \times 10^8 \text{ W})^{-1} \lambda\!\!\!^{-2} I \tag{1.38}$$

where r_e is the classical radius of the electron, $\lambda\!\!\!^{-} \equiv 1/\omega$ is the reduced wavelength of the wave, and I is the intensity of the electromagnetic radiation, *e.g.*, in units of W/m^2. For a practical wave of finite transverse extent, the intensity is given by the total power in the wave divided by the transverse area of the wave, which must be much larger than $\lambda\!\!\!^{-2}$ due to diffraction.

In fact, it is difficult in practice to obtain focal areas much smaller than $\simeq 1000\lambda\!\!\!^{-2}$ with paraxial focusing, which leads to the requirement for a terawatt of power to obtain $\eta \sim 1$, regardless of the wavelength used. In our experiment, this power was obtained with a terawatt laser using chirped pulse amplification by a slab amplifier[18], operating at a wavelength of 1054 nm (IR). A frequency-doubling crystal also allowed us to experiment with light at 527 nm (green). Focusing to an area as small as 50 μm^2 in IR and 30 μm^2 in green allowed us to obtain η as high as 0.5 in IR and 0.4 in green.

The laser system included diagnostics for the pulse energy, pulse length, and focal area, as described fully by Boege [5]. For this purpose the beam was recollimated and transported from the interaction point back to the laser room. The laser energy was measured both by a leakage monitor behind one of the mirrors before the transport line and behind a flat in the diagnostic line after the transport. The IR pulse-width was measured with a single shot autocorrelator in the diagnostic line, while a "typical" width of the green pulses was assessed occasionally using a streak camera. Early in the experiment, shot-by-shot measurements of the laser focal area were attempted by re-focusing the returned beam on a digitized video camera, after deliberate attenuation of the light by a factor of more than 10^6 using multiple reflections from glass plates plus filters. This provided some useful data, particularly for our first results comparing the measured electron spectrum with theory[1], but the statistical and systematic uncertainties in the area measurement were both of order 30%. Neither pulse length nor focal spot measurements were available in the most recently taken data, which were oriented towards the positron and photon measurements; instead, we relied on estimations of the focal properties based on observed rates in the $n = 1$, $n = 2$, and $n = 3$ regions of the electron spectrum[Chapter 5].

The electron beam in brief

The desire for $\Upsilon \sim 1$ and the relationship $\Upsilon = x_0\eta/2$ (Eqs. 1.9 and 1.29) mandate the use of high-energy electrons. With 46.6-GeV electrons from the Stanford Linear Accelerator, the value of x_0 is 0.82 for IR and 1.64 for green, allowing us to attain maximum Υ values of 0.2 and 0.3, respectively. No lower-energy electron beam would have sufficed.

As mentioned previously, the near-unity value of x_0 leads to good separation of the kinematic edges of the different orders n. Table 1.3 summarizes the kinematic edges for the specific conditions of E-144.

The electron beam had an invariant emittance of $< 3 \times 10^{-5}$ horizontally and $< 2 \times 10^{-6}$ vertically, and an energy spread typically $\delta E/E < 0.0005$. The E-144 interaction point was located in the dump line of the FFTB, between the last focusing element and the first of dipole safety magnet. At this point, the beam could be focused to a spot as small as 30 μm by 30 μm (root mean square) or as large as 200 μm horizontally and 100 μm vertically. Longitudinally, the electron

Table 1.3: Scattered electron kinematic edge energies for IR and green laser in the limit of $\eta \to 0$, for various numbers of absorbed photons n.

	$E_{p',\max}(n)$ at $\eta = 0$ [GeV]	
n	IR	Green
1	25.5	17.6
2	17.6	10.8
3	13.4	7.8
4	10.8	6.1

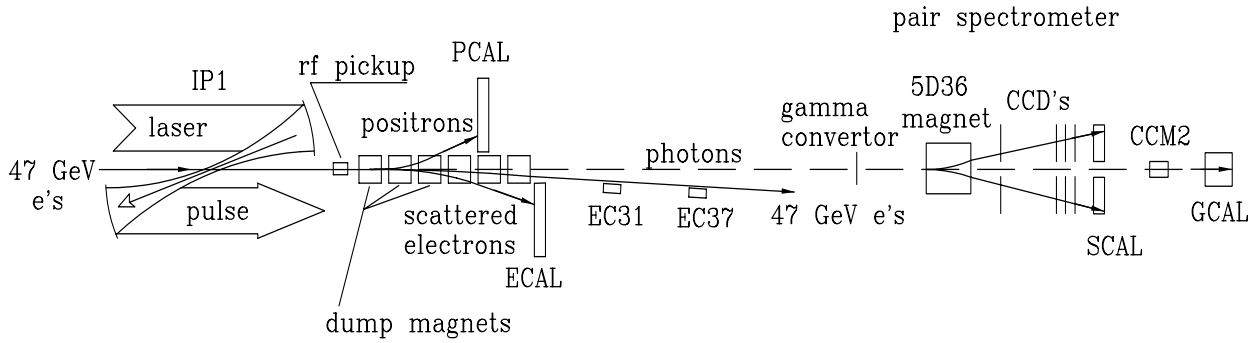

Figure 1.1: E-144 experimental arrangement.

pulse length could be adjusted from 0.5 to 1.0 mm (rms).[6]

The electron beam had a number of diagnostic tools to measure its properties and assure its quality, as described elsewhere [Chapter 3].

The detectors

Figure 1.1 shows the experimental arrangement. Electrons travel from left to right in the figure, interacting with the laser at interaction point IP1. The focusing of the laser by off-axis paraboloid mirrors near IP1 is shown. A radio frequency pickup after IP1 provided electron bunch timing information. A set of permanent magnets dispersed the charged particles according to momentum and separated the primary beam from the photons. Each of the three products of the interactions of interest, positrons, electrons, and photons, were measured by various detectors.

The entire positron spectrum and a selected portion of the electron spectrum were intercepted by silicon-tungsten calorimeters, called PCAL and ECAL. A diagram of ECAL is shown in Fig. 1.2. Each layer of ECAL consists of one radiation length of tungsten, 300 microns of active silicon, and the circuit board on which the silicon is mounted. The silicon is subdivided into rectangular "pads" arranged in 4 columns and 12 rows, each pad measuring 1.6 × 1.6 cm. The longitudinal layers are ganged into 4 segments, each containing approximately the same total energy deposition, and the $4 \times 12 \times 4 = 192$ channels are read out by 16-bit analog-to-digital converter (ADC) modules built for

[6] This was done by varying the strength of the compressor klystron in the transport from the SLC north damping ring to the linac [21].

a liquid argon calorimeter used in a Fermilab experiment [22]. PCAL is of identical construction, except for having 4 additional rows; its 128 channels[7] are also read out by the same modules.

The gain of the ECAL channels was set at approximately 13 ADC counts per GeV, so that with proper centering of the calorimeter, a single tower could accommodate a total energy of 10 TeV without ADC saturation. The calorimeters were calibrated *in situ* using a parasitic test beam[23]. The results of a calibration run for 13-GeV electrons are shown in Fig. 1.3. The resolution was found to be

$$\sigma^2_{E[\mathrm{GeV}]} = (0.19)^2(E[\mathrm{GeV}]) + (0.4)^2 + (0.05)^2(E[\mathrm{GeV}])^2. \tag{1.39}$$

The gain of the PCAL was set 32 times higher, since generally only one or zero positrons would hit the calorimeter, allowing better resolution for reconstructing the positron momentum.

Because of the wide dynamic range of the electron spectrum expected (cf Fig. 1.9), ECAL was mounted on a stage which could be moved vertically, allowing it to retreat from high flux regions as the laser intensity was raised. The stage was controlled by the SLC control system, but a direct interface between that control system and the E144 data acquisition system [Chapter 3] allowed the ECAL position to be scanned under experimenter control throughout a data acquisition run; extended runs of data at fixed position were also taken. Fig. 1.4 shows the regions of expected saturation and insufficient signal as a function of the laser pulse energy with all other parameters at "standard" values. Figures such as these were used as guides to the data-taking strategy. The final data set contains both ECAL position scans and fixed-ECAL runs.

The electron rates were also monitored by air or oxygen gas Cherenkov electron counters [24], two of which are shown labeled as EC31 and EC37; these two detectors were located in the $n = 1$ portion of the electron spectrum. Similar detectors called N2MO and N3MO (not shown) were located in front of ECAL, below the $n = 1$ kinematic edge, where they monitored electrons with energies in the range 12.2 to 13.2 GeV ($n = 2$) and 8.5 to 9.6 GeV ($n = 3$), respectively. These Cherenkov monitors were also calibrated *in situ* using the parasitic test beam [Chapter 4].

High-energy photons emitted from IP1 proceeded down their own beamline. Two collimators cleaned the photon beam of synchrotron and beam-scraping backgrounds. A movable Cherenkov counter (CCM1) could be inserted into this beamline, or the photons could be allowed to continue to a thin aluminum or nitrocellulose foil, in which a small fraction would convert into electron-positron pairs. A second dipole magnet (labeled 5D36 in Fig. 1.1) provided horizontal kick of 100 to 250 MeV/c to these charged particles, after which they passed through a particle tracking device consisting of multiple CCD planes. Individual particle tracks could then be reconstructed from the hits in the CCDs, allowing the momentum of each particle to be determined, and thus the gamma spectrum was inferred. (The CCD system and analysis are described in detail in separate papers [Chapter 6, Chapter 7].)

For the data presented in this paper, the CCD spectrometer was used in "single-arm" mode, in which no attempt was made to reconstruct pairs by matching electrons with their positron partners. The primary advantage of this mode of operation is that there is no need to significantly limit the number of tracks per event. Event rates of 100 tracks/event were easily accommodated. No attempt was made to use the CCDs in the front plane of the spectrometer in this mode, since the high number of hits led to significant ambiguity in the projection from the back planes to the front. The momentum spectrum of "single-arm" tracks is the convolution of the gamma spectrum

[7] In early runs, PCAL was ganged to have 2 longitudinal segments; in later runs, the same longitudinal segmentation as ECAL was used, and only the 8 rows covering the kinematically-accessible range of positron energies were read out.

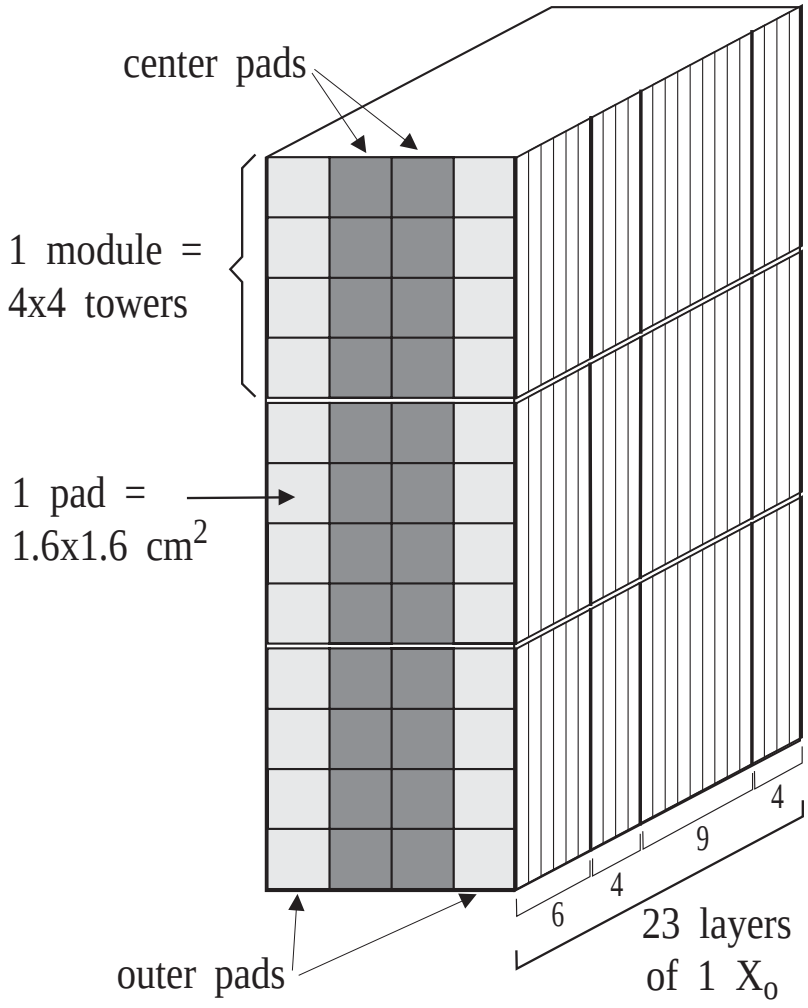

Figure 1.2: The electron calorimeter (ECAL).

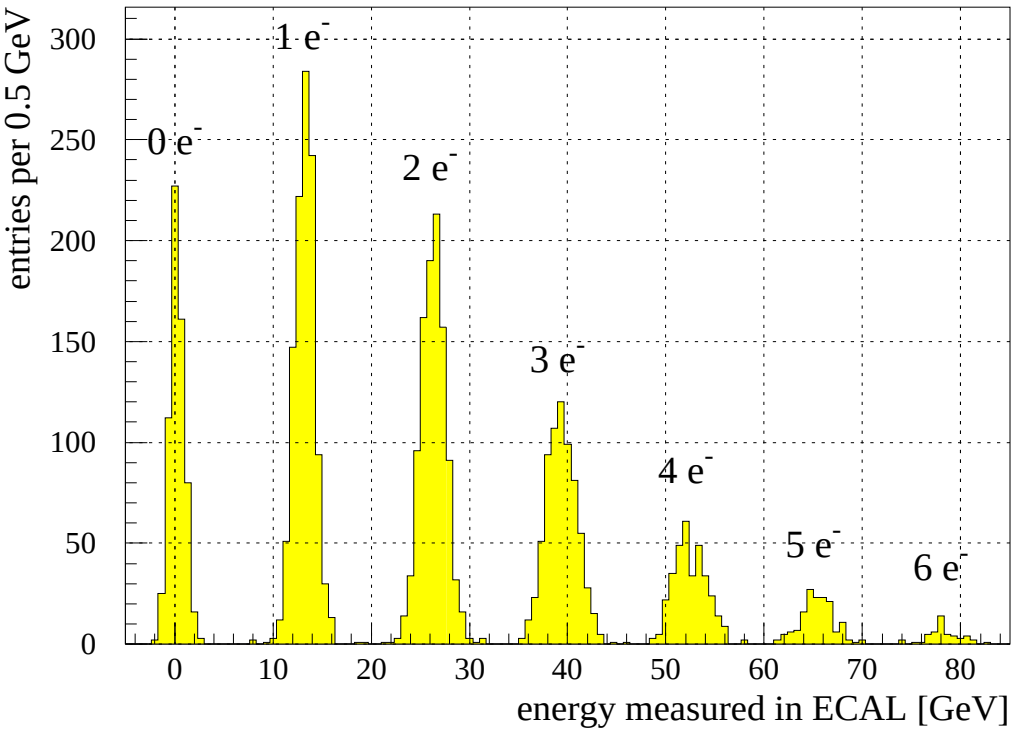

Figure 1.3: The response of the ECAL to 13-GeV incident electrons. The peaks due to the simultaneous arrival of up to six electrons are clearly distinguished.

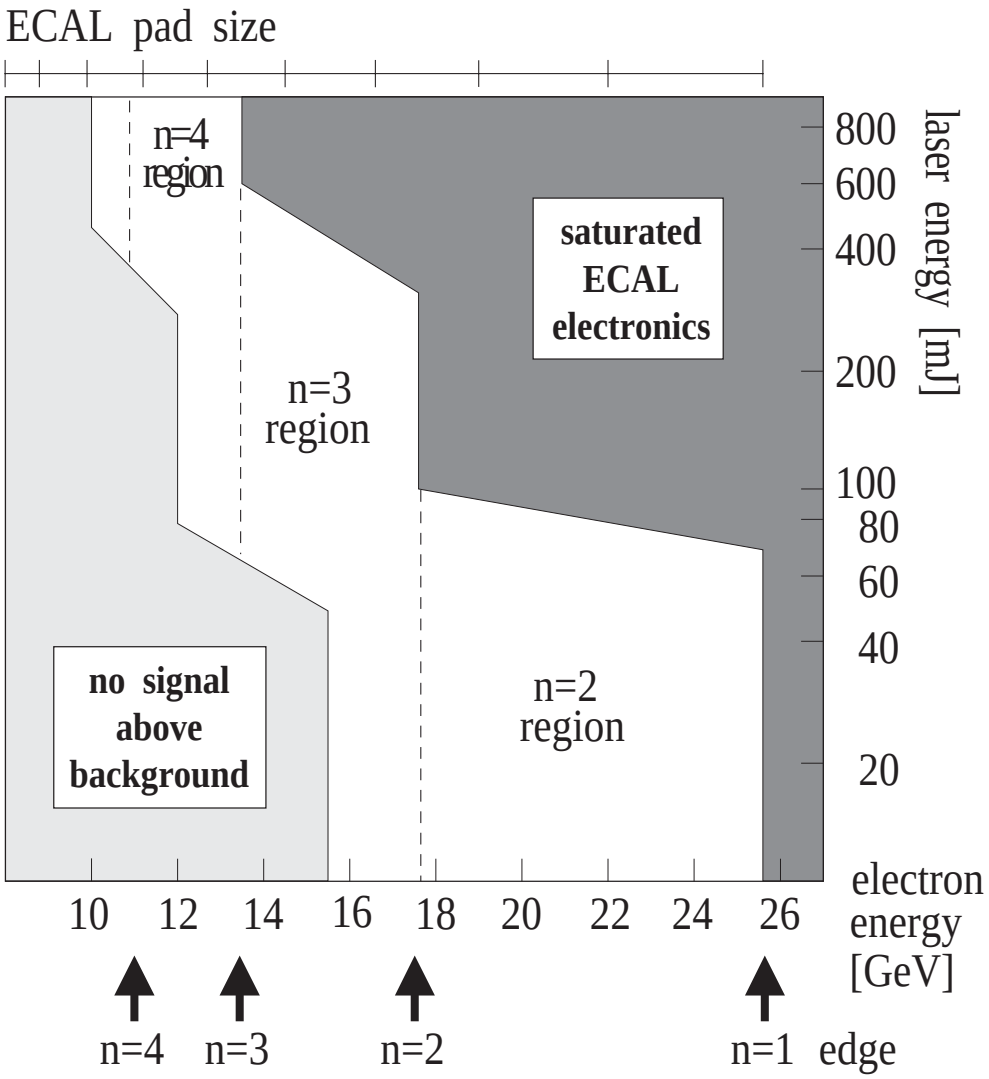

Figure 1.4: Portion of recoil electron spectrum accessible to the ECAL for different particle energies and laser energies for IR pulses. The acceptance of a single ECAL tower is shown at the top.

with the Bethe-Heitler pair spectrum. This spectrum is easily calculated, and the kinematic limits and relative scales of the $n = 1$ and $n = 2$ processes are clearly observable even in the convolved data.

The total number of backscattered photons, dominated by the linear Compton scattering process, was measured at the end of the beamline by Cherenkov monitor CCM2 and silicon-tungsten calorimeter GCAL.

Not shown in Fig. 1.1 are various devices which provided non-invasive measurements of laser pulse energy (described in [5, 2, 6, 18]) and electron beam position, angle, total charge, and energy on every event [Chapter 3].

Beam synchronization and overlap

The fine timing of the laser pulse was set by adjusting an optical delay line with sub-picosecond resolution. A ringing cavity in the electron beam line provided a reference signal which was used to compare the phase of the laser signal to that of the electron bunch. Final timing was established by observing the scattering rate as a function of optical delay. Such a timing curve is given in Fig. 1.5 and has an rms width (standard deviation) $\sigma = 4.3$ ps; the spread can be attributed to the electron beam length which was $\sigma_e \simeq 3$ ps plus relative timing jitter of about the same magnitude. A detailed analysis of the fluctuations in the collision rate [25] shows that the timing jitter between the laser pulses and the arrival of the electron bunch was typically of order $\sigma_j \simeq 2$ ps to 3 ps.

The final optics for the laser at IP1 were contained in an evacuated box, which was placed on three motorized, elliptical cams which could position the focus in the horizontal (x) and vertical (y) directions and could also introduce a roll in the xy plane. This capability was used to scan the laser focus with respect to the electron beam.

To obtain initial alignment of the electron and laser beams in the transverse plane, a fluorescent flag was lowered into the path of the beam and viewed remotely. By moving the IP1 box, it was possible to bring the electron beam and laser images into overlap. Final adjustment was made by monitoring the forward-photon rate as a function of transverse (x-y) position of the IP box. While the vertical overlap (y) was unambiguous, the overlap in the horizontal plane (x) depended on the relative timing of the two beams as indicated in Fig. 1.6(a). Thus it was necessary to carry out a raster scan in both the x-position of the IP box and in timing delay. This is shown in Fig. 1.6(b) where the linear Compton rate is plotted as a function of Δx and Δt. The correlation between the two offsets,

$$\Delta x \left(\frac{1}{\sin\alpha} + \frac{1}{\tan\alpha} \right) = c\Delta t, \tag{1.40}$$

is clearly evident. In Fig. 1.6(c) the nonlinear rate (specifically, $n = 2$) is plotted for the same raster scan. Here a large signal was obtained only when the electrons crossed through the laser beam in the focal region. By scanning both the position where the electrons intersected the laser path and the time at which the laser pulse crossed the focal waist, the xt scan identified the location where η was highest, which was an essential consideration for the experiment. The xt scan data were also used to separate unwanted backgrounds from linear ($n = 1$) processes from the desired $n > 1$ signal, based on their different shapes in the xt plane. Such xt scans were performed frequently during the run to assure correct spatial and temporal overlap. Each xt scan was preceded by a y scan to ensure that the electrons and photons were in the same plane, and a simple t scan to verify the scan range and $n = 1$ collision stability.

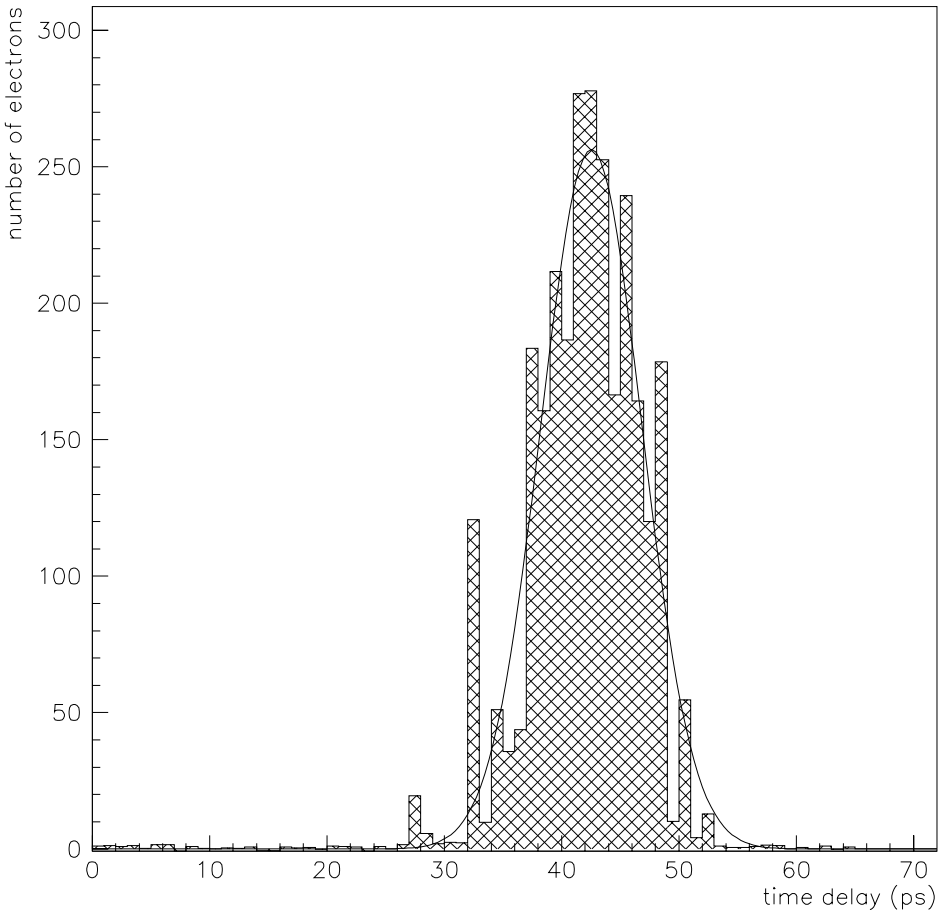

Figure 1.5: A "timing curve" showing the number of electrons scattered into the top row of the electron calorimeter as a function of delay of the optical pulse. The standard deviation of the curve is $\sigma = 4.3$ ps.

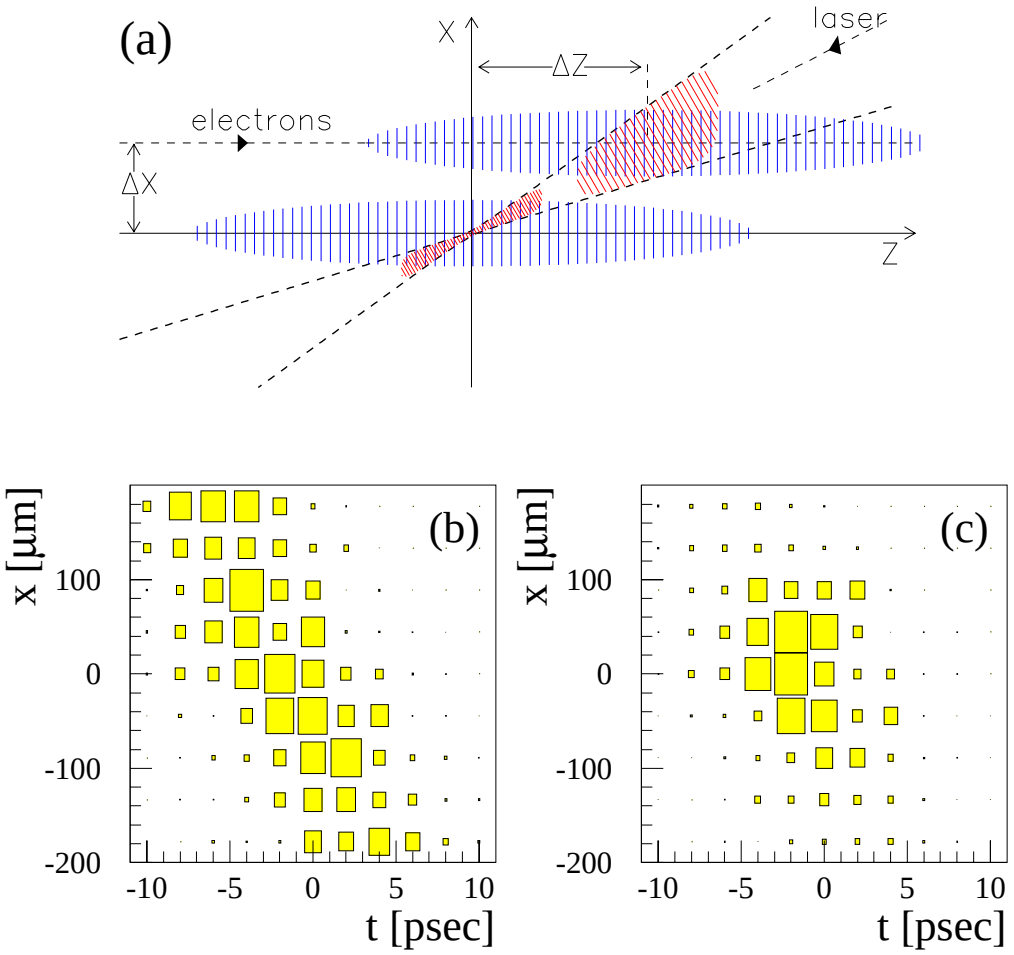

Figure 1.6: (a) The crossing of the laser pulse and electron beam in the x-t plane; two possible collisions are shown, each giving approximately the same linear Compton scattering rate but drastically different non-linear Compton rates. (b) Linear Compton event rate as a function of transverse beam displacement and relative timing. (c) As above but for nonlinear events.

The data acquisition system (DAQ)

The data acquisition system (DAQ) collected data from the detectors as well as information on the laser system and the electron beam. Although the accelerator provided electrons at 10 or 30 Hz, the DAQ recorded data less frequently. Physics events were recorded at the 0.5 Hz repetition rate of the laser. Every beam crossing when the laser fired constituted an event to be recorded in this experiment. In addition, events were collected on the electron beam pulses 400 msec and 200 msec prior to each laser shot, in order to measure electron-beam-related backgrounds in the detectors. A third set of events, obtained by dumping the electron beam far upstream in the linac, was used to measure the pedestal mean value and noise in the detectors and readout electronics during data-taking. These events occurred at a rate of 0.05 Hz, and one third of them coincided with laser shots.

The moderate event rate and data volume of the experiment allowed a low-cost solution for the data acquisition system, which is shown schematically in Fig. 1.7. The system was based on personal computers with Intel i486 microprocessors and Industry Standard Architecture (ISA) motherboards, connected by a local Ethernet. The communication between the computers was established using the standard TCP/IP and UDP protocols. The main part of the DAQ system consisted of one back-end computer and several front-end computers. The back-end computer controlled and synchronized the DAQ system and provided the interface to the user, while the front-end computers collected data from the detectors and diagnostic equipment and responded to command messages received from the back-end computer. A standard interface between the programs running on the back-end computer and the front-end computers allowed for a modular and very flexible DAQ. The third type of personal computer shown in Fig. 1.7, the display computers, provided detailed online monitoring of the collected data.

The readout cycles were controlled by the trigger logic, which received triggers from the SLAC control system and distributed them to the readout electronics and/or front-end computers. Once a trigger signal passed through the logic, any further triggers were blocked until the logic was reset by a 'READY' signal from the back-end computer. Upon receiving a trigger signal, each front-end computer collected its assigned set of data and sent it over the Ethernet link to the back-end computer, where the data were assembled, analyzed and stored to disk. As soon as the back-end computer finished processing the event, it reset the trigger logic and broadcast the full event information to the display computers. The display computers, running unsynchronized to the readout cycle, allowed online monitoring of the experiment or individual front-end equipments, as well as CPU-intensive data processing. One display computer was set up in the accelerator control room, guiding operators in tuning the electron beam for low background in the detectors.

An RS-232 link between one of the front-end computers and the master computer of the SLAC control system made it possible to bring additional experimental parameters into the data stream as well as to control the position of the IP box and the ECAL. The ability to move these devices by the DAQ program was essential in performing scans where the scattering rate was measured as a function of these positions.

Summary

The experiment studies the interaction of 46.6-GeV electrons with a terawatt laser under conditions in which η and Υ approach, but do not exceed, unity. Under these conditions, the following were measured on every trigger:

- Beam and laser properties

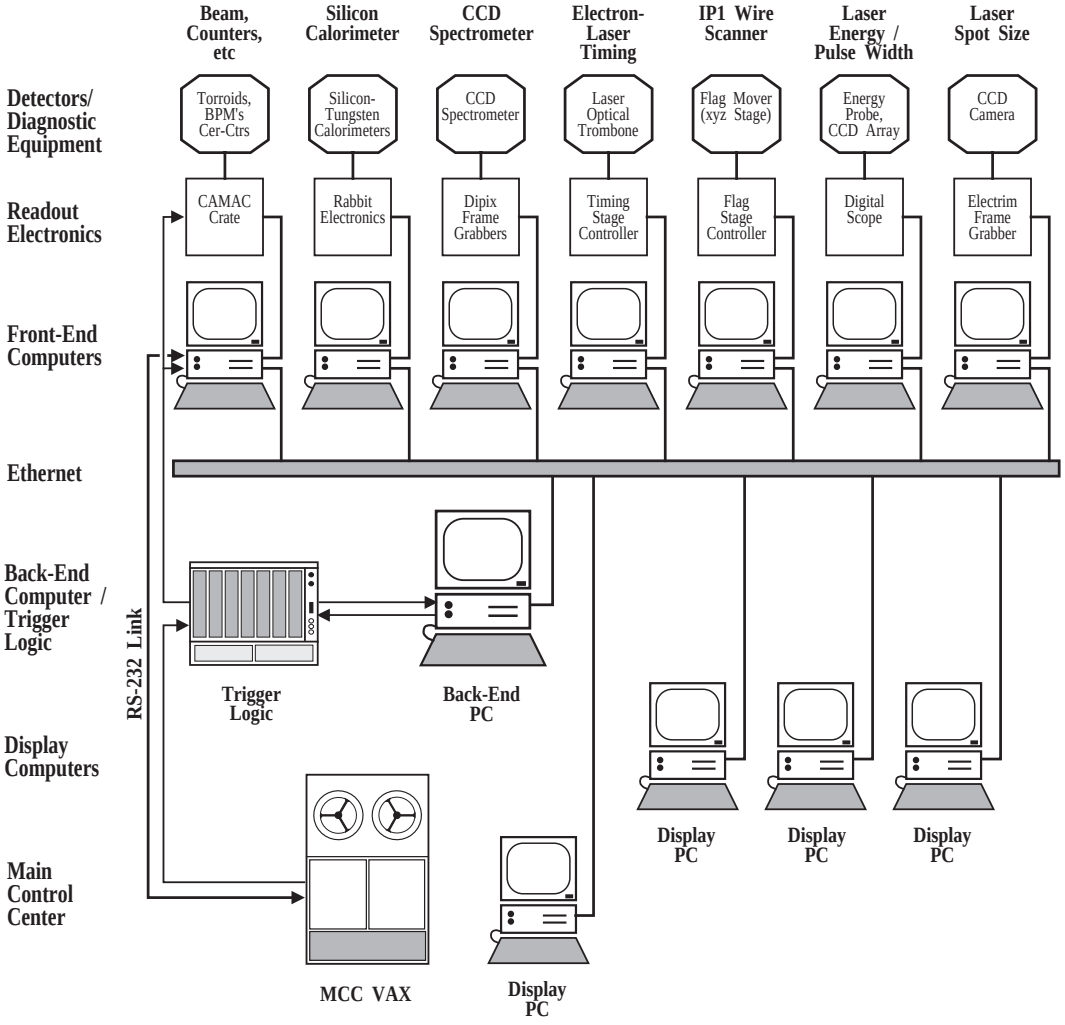

Figure 1.7: Hardware components of the data acquisition system.

Table 1.4: An overview of the precision and accuracy of E-144's measurements

measurement	uncertainty in measurement	
	source	magnitude
η, Υ	systematic	$\pm$ 15%
N_γ	systematic	$\pm$ 15%
N_{e-}	systematic	$\pm$10%
N_{e+}	statistical	$\sqrt{175+66/4}/(175-66) = \pm 13\%$
$e-$ spectrum	systematic	$\pm$10% in each 1- to 2-GeV/c-wide row
$e+$ spectrum	statistical	$\pm$20% in each 2-GeV/c-wide bin
γ spectrum	statistical	$\pm$10% in each 0.4-GeV/c-wide bin ($n = 2$)
		$\pm$1% in each 0.4-GeV/c-wide bin ($n = 1$)

- Total number of backscattered photons
- Total number of scattered electrons within the ECAL acceptance
- Total number of positrons
- Electron spectrum shape
- Photon spectrum shape
- Positron spectrum shape

Table 1.4 gives an overview of the systematic and statistical uncertainties in each measurement. Due to the nonlinear nature of the interaction, the focal intensity measurement is the most serious limitation of the experiment.

1.1.3 Simulation

As already described in detail, our predictions are based on the theory of an electron propagating in an infinite plane wave. Figure 1.8 is a plot of the $n = 1$ and $n = 2$ recoil electron spectra after propagation through a very short region of the green laser at various (constant) intensities; the $n = 2$ spectrum is almost invisible on the linear scale except at the highest intensity. There is a noticeable shift in the kinematic edges caused by the predicted increase in the "effective mass" of the electron in the wave, but this shift is not as noticeable in the spectrum of electron colliding with a realistic laser focus.

In order to simulate the interaction of electrons with the focussed laser, we numerically integrate over time and position in the laser focus, using the infinite-plane-wave theory locally at every point.[8] Thus, the total scattering rate gets contributions from scattering at all the different intensities encountered. The simulation also allows for scattered electrons to re-scatter any number of times as they traverse the laser. In Fig. 1.9 is plotted the results of the simulation of the complete interaction with a Gaussian laser focus. A "close-up" view of the $n = 1$ edge is provided at the right of the figure. There is a subtle change in shape with increasing intensity, signifying the mass shift effect, which can be discerned in the simulation but which our calorimeter was not suited to detect due to the coarseness of its momentum resolution. Note that in this figure, "$\eta_{\rm foc}$" refers to the maximum value of η achieved at the focus.

[8] Actually, there are two, independently-written simulations for the experiment, which give equivalent predictions: one, a binned numerical integration [26], and the other, a true Monte-Carlo simulation[Chapter 2].

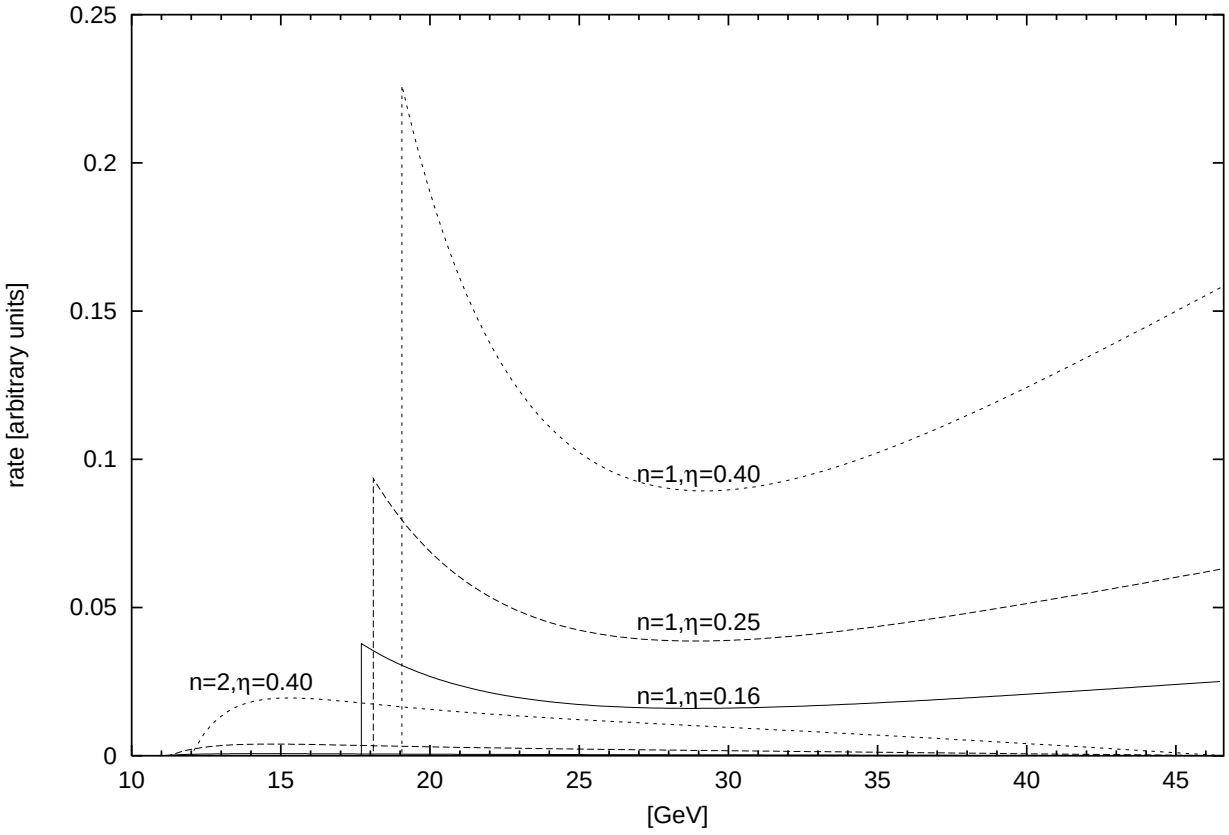

Figure 1.8: The $n = 1$ and $n = 2$ recoil electron spectra after propagation through a very short region of the green laser at various (constant) intensities.

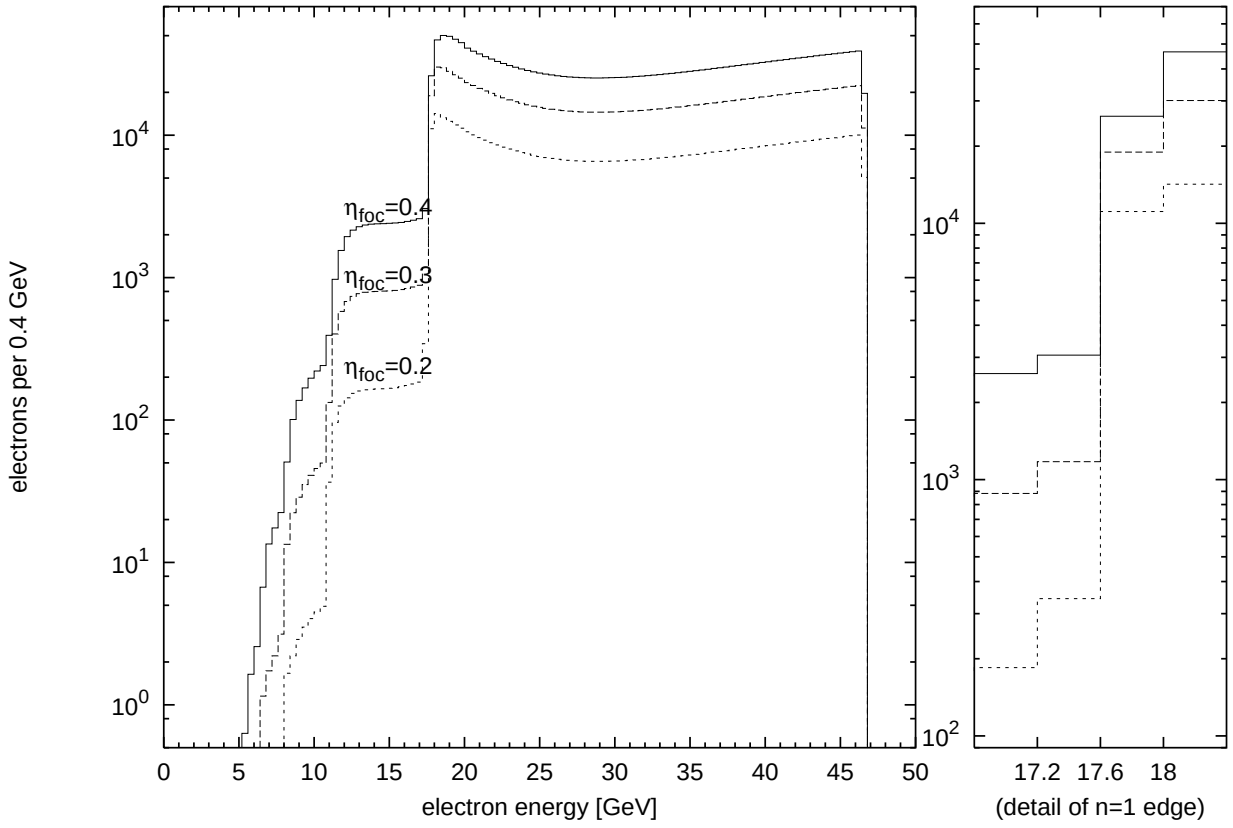

Figure 1.9: Results of the simulation of the complete interaction, assuming a Gaussian laser focus; a "close-up" view of the $n = 1$ edge is provided at right.

The contributions of various processes can be separated in the simulation, though not in the experiment, as shown in the simulated electron spectrum in Fig. 1.10. Note that the simulation below was done for the infrared laser, in contrast to those above which were for the green laser. The curve marked "$n = 1$ plural scattering" shows the contribution of repeated scattering of the same electron by the $n = 1$ ("linear") Compton process.

At sufficiently high intensity, the "linear" and non-linear Compton scattering process rates diverge from their leading-order (in incident wave intensity) behavior. In Fig. 1.11, the simulated total number of scatters from each order is plotted as a function of the maximum value of η achieved at the focus. Note that the maximum value of η actually achieved by E-144 with the green laser is 0.4.

Simulation of pair production in the collision was handled by the same mechanism as multiple scattering of the electrons: photons radiated by the Compton process were propagated through the laser and allowed to interact to form pairs, the components of which could then Compton scatter: *i.e.*, a complete particle shower simulation was implemented for the two processes of Eqs. 1.1 1.2. Figure 1.12 shows calculations comparing the positron rate predicted by the complete theory with all non-leading-order-in-incident-wave-intensity effects (*e.g.*, mass-shift effect) with a prediction using only the leading-order effects. It is apparent that E-144 explores a region that is non-linear, but is sensitive only to the leading-order contributions.

As mentioned previously, the CCD tracking spectrometer was used in "single-arm" mode, such that only the spectrum of the separate components of the pairs is known. The simulated single-arm spectrum is plotted in Fig. 1.13.

These predictions, based on the accepted theory and the measured or estimated properties of the laser focus, are compared to the observations in order to judge the overall consistency of the theory and experimental model with the data.

1.1.4 Aspects of the theory which may be investigated with this experiment

The experiment has the capability to detect the presence of particles scattered at different orders n, to measure their overall rate, to resolve the detailed shape of their spectra to some extent, and to correlate these measurements with an estimate of the focal intensity parameter η^2.

The presence of particles with the appropriate signature is sufficient to demonstrate the existence of the reactions in Eqs. 1.1 and 1.2. In particular, the detection of any positrons created by the laser at IP1 demonstrates pair production (Eq. 1.2), and the detection of any tracks in the pair spectrometer above the $n = 1$ edge demonstrates non-linear Compton scattering (Eq. 1.1 with $n > 0$). The experiment is also capable of detecting electrons in the $n = 2$, $n = 3$, and $n = 4$ regions of the spectrum, and while electrons can also reach such energies through multiple scattering, detection of more electrons than expected from multiple $n = 1$ scattering is also evidence for non-linear Compton scattering.

The absolute rate can be compared to theory to verify the magnitude of the matrix elements $\mathcal{M}$ to some degree, and the variation of rate with η^2 can be compared to the expected leading-order behavior. The electron and photon rates in regions dominated by n^{th}-harmonic Compton scattering are expected to vary proportionally to η^{2n}, while the positrons, being produced primarily by an $n = 4$ pair creation process fed by photons created by an $n = 1$ Compton scattering process, are expected to vary approximately as η^{10}.

Due to the transverse size of the shower in ECAL, the shape of the electron spectrum cannot be resolved sufficiently to allow discernment of the subtle mass-shift effect illustrated in Fig. 1.9.

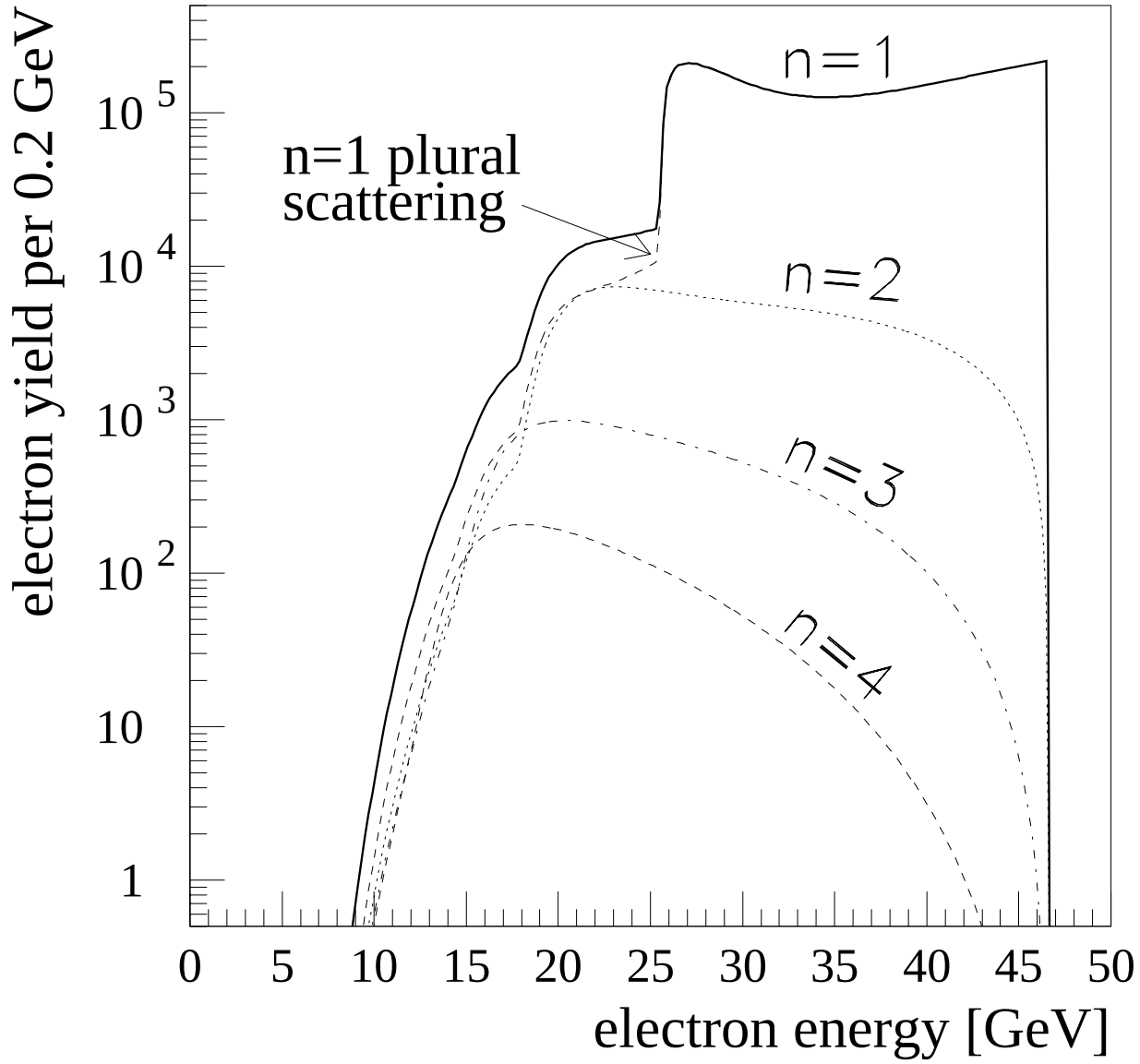

Figure 1.10: Simulated electron spectrum (solid line) and contributions of various processes separately plotted (broken lines). Note that this simulation was done for the infrared laser, in contrast to those above which were for the green laser. The curve marked "$n = 1$ plural scattering" shows the contribution of repeated scattering of a given electron by the $n = 1$ ("linear") Compton process.

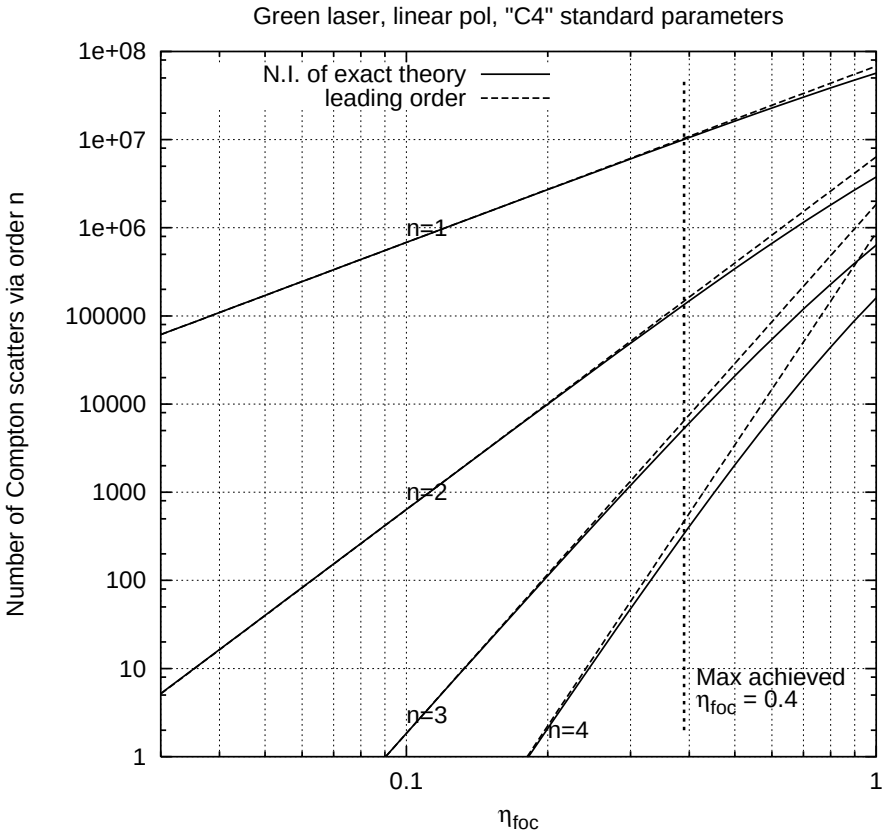

Figure 1.11: Total number of scatters from each order plotted as a function of the maximum value of η achieved at the focus. This simulation was done for the green laser.

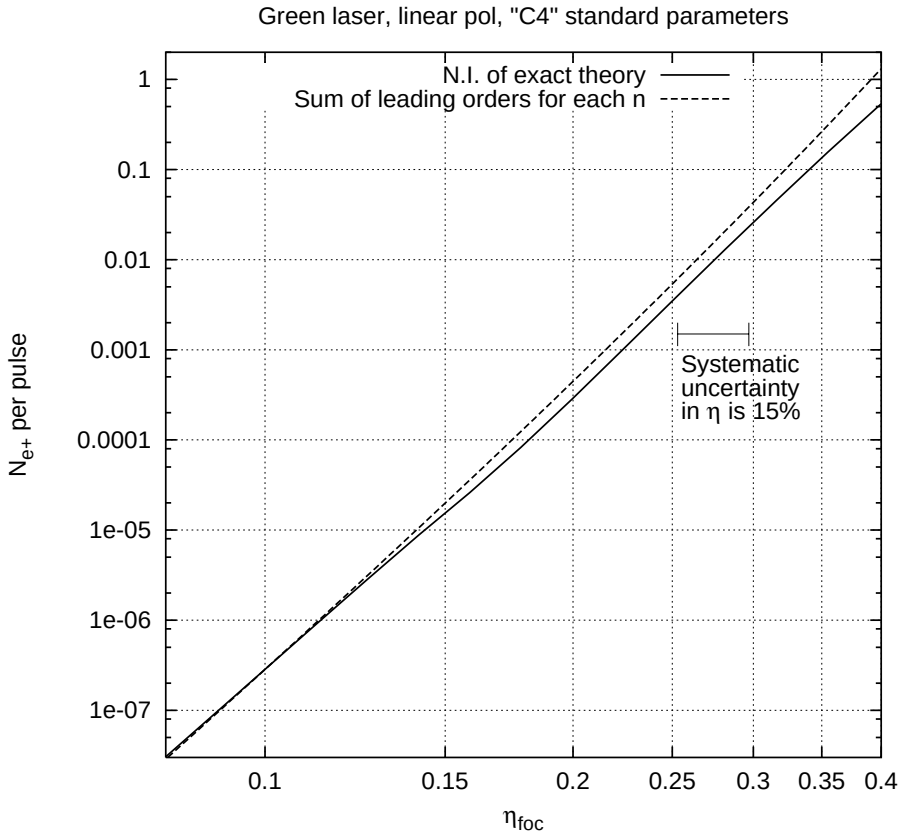

Figure 1.12: Total number of positrons plotted as a function of the maximum value of η achieved at the focus, along with a prediction using only the leading order contribution. This simulation was done for the green laser.

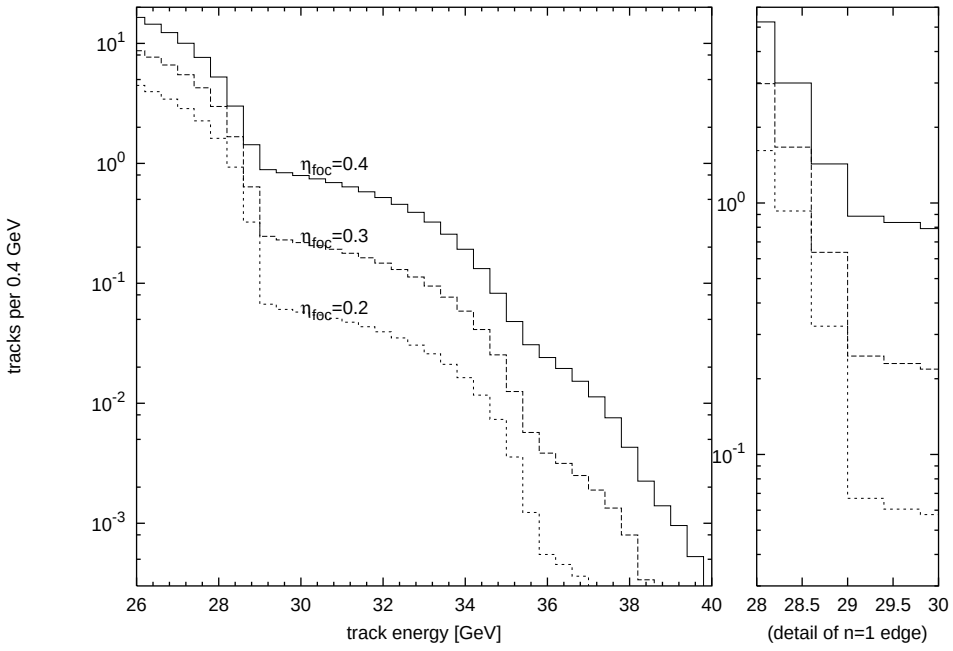

Figure 1.13: Simulated single-arm track spectrum, for the green laser. A "close-up" view of the n=1 edge is provided at the right.

The CCD spectrometer has no such problems, but as illustrated in Fig. 1.13, the single-arm track spectrum shows almost no η-dependent changes in shape, other than the η^{2n} dependence of regions corresponding to different orders n. Therefore, the mass-shift effect can not be detected. The detailed shape of the spectra does provide some additional check on the form of $|\mathcal{M}|^2$.

1.1.5 Relevance to beam physics and other fields

Any physics experiment using particle beams includes one or more particle sources, acceleration and/or damping systems, focusing systems, and one or more interaction points; this experiment is most relevant to the first and last of these. At the source end of beam physics, there is interest in exploiting Breit-Wheeler pair production (Eq. 1.2) as a polarized, low-emittance positron source[27], and also interest in using Compton scattering (Eq. 1.1) to provide high brightness sources of high energy photons for photon-photon colliders[28, 29]: E-144's positron and photon production results have direct relevance to these topics. At the interaction point end, it is generally thought that in next-generation linear collider designs (such as NLC[30]) it is desirable to have $\Upsilon \sim 0.3$ for reasonable luminosity while maintaining low pair creation backgrounds; this experiment provides results in exactly that high field regime. Critical fields of this strength and higher are also expected in the collisions of ultra-relativistic heavy ions.

It is currently infeasible to create experimental conditions wherein a potential drop of 511 kV exists over a Compton wavelength without using an ultra-relativistic particle beam. Therefore, terrestrial high-Υ experiments are limited to particle physics at present. However, it is believed that such fields may occur in astrophysical contexts. For example, the magnetic field at the surface of a neutron star is believed to exceed the Schwinger critical field. In such an environment, a

particle need not be highly relativistic to experience static, crossed magnetic and electric fields which both exceed the critical field. It is difficult to find even in nature any conditions exactly as probed by this experiment, *i.e.*, a charged particle encountering a coherent plane wave with non-zero frequency and near-critical field strength, particularly in the region where $\bar{n} \sim 1$, but this does not diminish the importance of the experiment. Studying $\Upsilon \sim 0.3$ can only improve our understanding of electrodynamic interactions under naturally-occurring and human-created conditions.

1.2 Review of previously-published experimental results

1.2.1 Electrons [1, 2, 3]

In Fig. 1.14 is shown the observed electron calorimeter data separated according to η and plotted versus momentum (closed circles). The predicted results for the same conditions are also plotted (open circles). These results are from data taken with the green, linearly-polarized laser[3], in the $n = 3$ and $n = 4$ region of the spectrum. The data show clear evidence for the presence of electrons scattered to these energies at rates significantly in excess of those expected from multiple $n = 1$ scattering alone; furthermore, the rates agree at multiple values of η, to within a factor of 2 (note the log scale).

We also have data for circularly-polarized green and infrared laser[1] and for the linearly-polarized infrared laser.

1.2.2 Positrons [4, 6]

In Fig. 1.15.a is plotted the number of observed positrons for laser-on triggers and for laser-off pulses, the latter being scaled to the number of laser-on pulses. The spectrum of laser-associated positrons, obtained by subtracting the laser-off from the laser-on distribution, is plotted in Fig. 1.15.b, along with the expected momentum spectrum from the model calculation. Parts (c) and (d) of the figure show similar plots using only events which satisfy the requirement $\eta > 0.216$.

Figure 1.16 shows the observed positron yield as a function of the intensity. In this analysis, the intensity of the focus has been inferred using data from Cherenkov electron counter data on the electron $n = 1$, $n = 2$, and $n = 3$ rates. The positron yield is expressed in terms of the number of positrons per high energy photon. The anomalously high yield at low η is believed to be due to a slight residual background, which is normalized by a low number of photons at low η. The solid line in the figure shows the simulated positron rates for Compton scattering (Eq. 1.1) followed by light-on-light pair production (Eq. refEq:pairprod). The dashed line in the lower right corner is an estimate of the pair production rates from virtual photons associated with the beam electrons interacting with the intense wave, *i.e.*, trident production or nonlinear (sub-threshold) Bethe-Heitler pair production with an electron target: it is clearly a small contribution. The horizontal error bars on each data point indicate the uncertainty in the η measurement, and the vertical error bars are purely statistical.

The data presented above was taken with the green, linearly-polarized laser and a 46.6-GeV electron beam. The presence of positrons is unmistakable, and the variation with η is as expected; furthermore, the absolute rate is consistent with the theory, within a factor of $e^{\pm 0.15 \times 10} \simeq 5^{\pm 1}$ uncertainty, assuming a 15% uncertainty in η and given the fifth-order dependence of the positron rate on η^2.

We also have data taken with the same laser and 49.2-GeV electron beam[6].

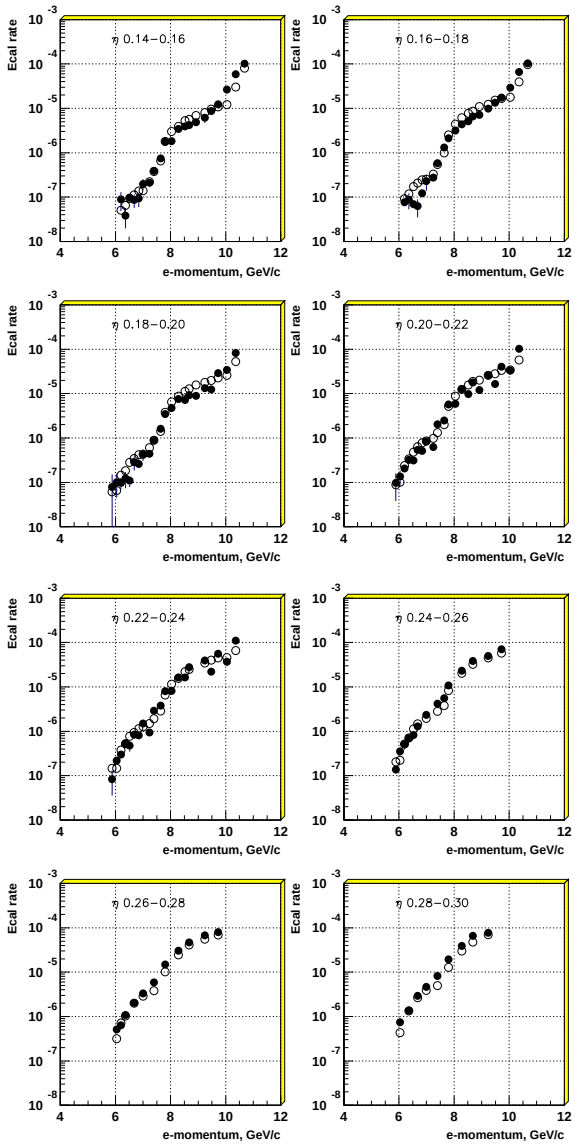

Figure 1.14: Electron calorimeter data separated according to η and plotted versus momentum (closed circles) and predicted results for the same conditions plotted (open circles). (From [3].)

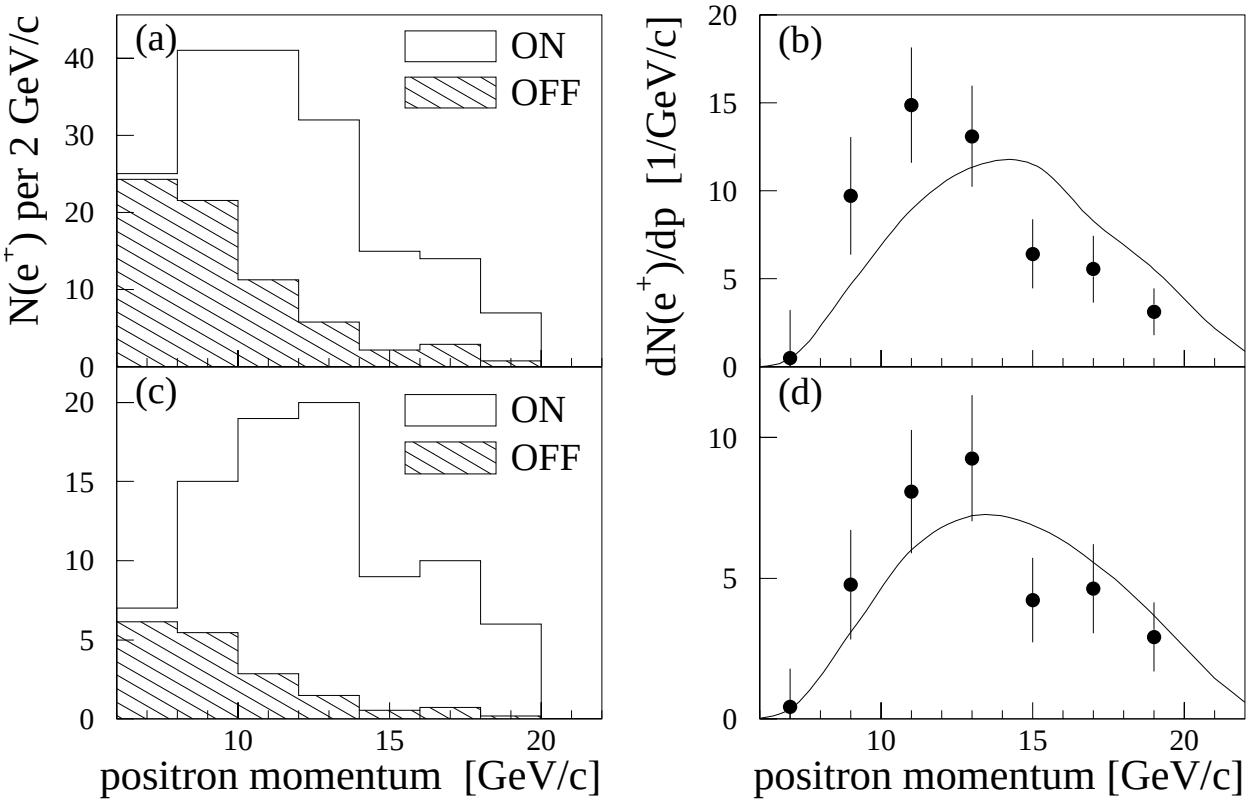

Figure 1.15: (a) Observed number of positron candidates vs. momentum for laser-on pulses and for laser-off pulses scaled to the number of laser-on pulses. (b) Observed spectrum of signal positrons obtained by subtracting the laser-off from the laser-on distribution. The curve shows the expected momentum spectrum from the model calculation. (c) and (d) are the same as (a) and (b) but with the requirement that $\eta > 0.216$. (From [4].)

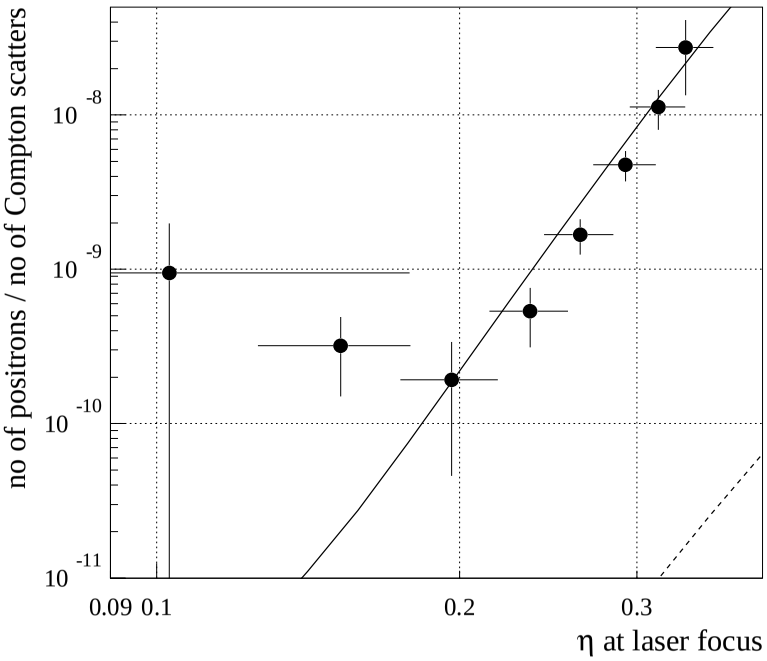

Figure 1.16: Observed dependence of the positron yield on the laser field-strength parameter η, with yield normalized to the total number of Compton scatters. The solid line is the prediction based on the numerical integration of the two-step Breit-Wheeler process. The dashed line represents the simulation for a one-step trident process. (From [4].)

1.2.3 Laser focal property estimation

For data in which measurements were made for every event of laser focal area, pulse length, and total energy, these parameters were used directly in the simulation. The peak intensity at the focus was calculated as

$$I_{\rm foc} = \frac{E_{\rm pulse}}{A_{\rm eff}\tau_{\rm eff}}, \tag{1.41}$$

where $E_{\rm pulse}$ is the total energy of the laser pulse, $A_{\rm eff}$ is the effective area of the focus, equal to $2\pi\sigma^2$ for a Gaussian focus with rms spot size σ, and $\tau_{\rm eff}$ is the effective pulse length, equal to $\sqrt{2\pi}\sigma_t$ for a Gaussian distribution with rms width σ_t. The important quantity $\eta_{\rm foc}^2$ was computed from Eq. 1.38. Note that A, τ, and $\eta_{\rm foc}$ all have significant and independent effects on the scattering rate, and in general should all be specified.

Another variable parameter affecting the scattering rates is the spatial and temporal overlap of the two beams. It is believed that the temporal jitter was the most significant contribution to overlap fluctuations. Since the electron beam is much larger than the laser focus and most of the rate is dominated by scattering at the focus, the main effect of overlap variation is a change in all rates by a common factor. For a given set of laser focal parameters A, τ, and $\eta_{\rm foc}$ and the separately-measured electron beam parameters, the total number of forward photons expected in the case of perfect overlap may be calculated. This number is dominated by ordinary ($n = 1$) Compton scattering. The overlap factor v is defined as the ratio of the observed number of photons to the predicted number for perfect overlap:

$$v \equiv \frac{N_{\gamma,\rm observed}}{N_{\gamma,\rm MC}} \simeq e^{-\frac{1}{2}\Delta t^2/\Sigma_t^2}, \tag{1.42}$$

where $N_{\gamma,\rm MC}$ is the prediction for perfect overlap, Δt is the timing offset, and Σ_t is given by adding the electron and laser rms pulse lengths in quadrature.

For the latest set of data, the electron Cherenkov monitor data on scattering rates in the $n = 1$, $n = 2$, and $n = 3$ portions of the spectrum were used to find an area A, pulse length τ, and relative overlap v consistent with the simulation and the observed Cherenkov monitor and laser energy data. This method provides a tighter check on the consistency of the positron rate data with the electron rates, and therefore a better overall test of the theory than could be done using only the laser parameter measurements. As an illustration of the technique, consider the case where the area and pulse energy are known, but not the pulse length or overlap. In this case, the rate in each detector may be written as

$$N_1 = k_1(E,A)v\tau^{-1} \tag{1.43}$$
$$N_2 = k_2(E,A)v\tau^{-2} \tag{1.44}$$
$$N_3 = k_3(E,A)v\tau^{-3}, \tag{1.45}$$

where k_i are coefficients found from the simulation, and the powers in τ reflect the leading-order (η^{2n}) behavior of the theory. Upon taking the logarithm, this is seen to be a simple linear model, easily fitted. It turns out that using this simple approach with a fixed, "typical" value of A leads to fits that are reasonably good, and that allowing A to vary over its physical range does not improve matters substantially.[9] Given a measurement of the $n = 2$ and $n = 3$ electron yield with 10%

[9] In the focal parameter fits used for the published positron results [4], the area and pulse length were varied together in a fixed relationship $A[\mu{\rm m}^2] = 20 \times \tau[{\rm psec}]$.

Table 1.5: Comparison of original calibration with cross-calibration results. Quoted errors include best-guess systematic errors, not including the suspected large calibration shift effect.

	calibration constant [e-/(GeV/c)/count]	
detector	original	cross-calibrated
N2MO	61	17 ± 2
N3MO	1.6	$0.65 \pm .07$

	calibration constant [photons/count]		
detector	original	cross-calibrated	adjusted for mutual consistency
EC31	289	179 ± 20	185
EC37	8874	6045 ± 600	5900

precision, the corresponding positron rates theoretically can be predicted with a precision of 45%, and the $n = 2$ photon yield theoretically can be predicted with 7% precision.

Unfortunately, it was immediately apparent that the values of τ predicted by this technique were approximately a factor of 2 smaller than the pulse lengths known to be typical of the laser, and significant bias in the residuals of the fits were also evident. It was concluded that the calibration of the Cherenkov monitors must somehow be incorrect, and an iterative technique for varying the calibration constants for best overall consistency of the data was developed [6]. Ultimately, the three Cherenkov monitor signals were scaled down by factors of 1.3, 8.0, and 8.0, respectively. This eliminated the bias in the residuals, and also brought the pulse length into line with *a priori* expectations.[10] As will be seen shortly, these values also give near-perfect agreement of the positron rate data with the simulation. Note, however, that different scaling factors could be chosen which would equally well eliminate the bias in the residuals, yet lead to different predictions for the overlap and pulse length. This consideration motivated the "cross-calibration" of the Cherenkov monitors with the electron calorimeter presented below in the section on "new results" and in [Chapter 4].

More details on the Cherenkov monitor focal parameter fits are given in [6] and [Chapter 5].

1.2.4 Summary of previously-published results

Within the uncertainties listed in Table 1.4, all measurements reviewed above are consistent with the accepted theory. Due to the nonlinear nature of the interaction, the focal intensity measurement was the dominant limitation in testing the electron data versus the theory. The positron results are internally consistent with the electron data, providing a better test of the theory than if only externally measured laser intensities were used.

1.3 Summary of new results

1.3.1 Cross-calibration of the electron Cherenkov monitors

As mentioned previously, the original calibration constants for the Cherenkov electron detectors as given by their creator [31] led to inconsistencies between the model predictions and the actually

[10] Actually, the average area and pulse length obtained using this scaling were significantly larger than the "standard" area and pulse length quoted prior to the beginning of the data analysis, but it was subsequently explained that the "standard" parameters are actually "best case running" parameters.

reported number of electrons observed by the detectors. This disagreement led us to distrust the original calibration constants, and to rescale those constants by large factors (up to eight) for purposes of estimating the laser focal parameters[6]. Subsequently, several sets of bremsstrahlung and Compton data was reanalyzed to obtain a cross-calibration of these detectors with the electron calorimeter (ECAL).

The results of the cross-calibration analysis of data from the period of data-taking encompassing most of the positron and photon data are summarized in Table 1.5. The new calibration is within a factor of four of the original, unscaled calibration. However, some less-representative data sets, taken under conditions of unusual steering, gave calibration constants higher by a factor of two, thus even closer to the original calibration. The steering was not radical enough to change the response of the detectors by this degree, according to the designer's best understanding of the detectors. In general, it is completely unknown why the calibration of these detectors changed from the original calibration or why the calibration apparently changed over time and/or with beam steering conditions.

Using the calibration given in Table 1.5, consistent values for the focal parameters A and τ and the overlap v may be found without additional scaling. However, the resulting values of η (calculated from observed area and fitted A and τ) are higher than previously estimated, by approximately 30%. This is roughly twice the estimated systematic error on η. At first glance, one would therefore expect this recalibration to change the interpretation of the positron results from agreement within a factor of $\sim 5^{\pm 1}$ to an observed deficit of a factor of $\sim 5^{-2\pm 1}$.

Using either the original calibration or the calibration constants from the "less representative" data set would suggest an even greater deficit, and there is no evidence to suggest that the electron flux could have been less than estimated using the constants in Table 1.5. Furthermore, data from the few positron-producing runs in which ECAL was positioned in the $n = 3$ region is consistent with the $n = 3$ Cherenkov monitor using the new calibration constants as shown, as one would expect given the cross-calibration, and is not consistent with the scaling previously applied. Unfortunately, the bulk of the positron data comes from runs where the $n = 3$ rate exceeded the saturation level of ECAL, and so ECAL was kept well below the $n = 3$ kinematic edge; much of this data consists of ECAL scans, in which ECAL had no fixed position even in the $n = 4$ spectrum and often had no signal at all. In the absence of any understanding of the Cherenkov monitors' calibration fluctuations and with no cross-check on their calibration for the bulk of the positron data, it must be admitted that the larger scale-down factors applied previously could indeed be more correct.

In conclusion, the absolute scale of the positron rate should only be said to agree with the theory to within a factor of $\sim 10^{+1}_{-2}$. Even with this uncertainty, the positron results do demonstrate the creation of pairs, presumably by the two processes (Eq. 1.1) and (Eq. 1.2), and the results show a rate dependence proportional to $\eta^{10(\pm 2)}$ as expected.

1.3.2 Photon spectrum

Tracks in the CCD tracking spectrometer were reconstructed, and backgrounds estimated[Chapter 7]. Focal properties were estimated on every pulse, as described previously and elsewhere[Chapter 5]. Using this data, the photon spectrum according to our simulation was calculated and redistributed according to the Bethe-Heitler spectrum, using the same formulae as used by EGS4[32]. The single-arm track spectrum thus calculated can be compared directly to the observed data, with an overall normalization to account for the net system efficiency as the only unknown.

Figure 1.17 shows the track spectrum integrated over all events (histogram) and the simulation

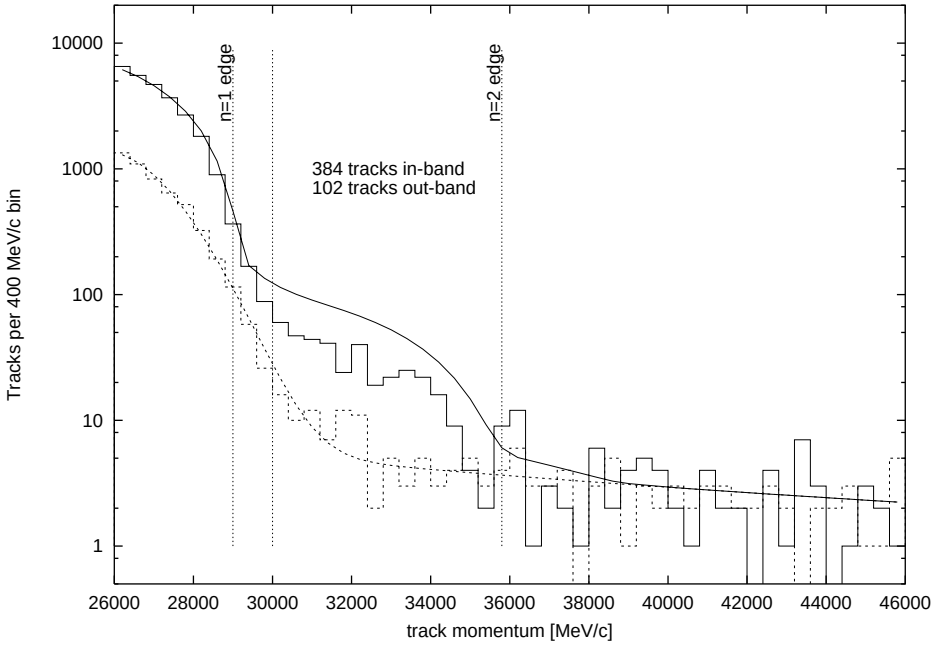

Figure 1.17: The track spectrum integrated over all events (solid histogram), simulation with fitted background (solid curve), and measured and fitted background (dashed curve and histogram). In the $n = 2$ dominated region between 30 GeV/c and 36 GeV/c, there are only 102 tracks in the "out-band" region of the CCDs used to estimate backgrounds, while there are 384 tracks in the "in-band" (signal) region of the CCDs.

with fitted background and normalization (solid curve). The background estimate is also shown (histogram and curve with short dashes). Many features of the data match the theory very well: the $n = 1$ kinematic endpoint at 29.0 GeV/c momentum is clearly seen, there are hundreds of tracks above background in the $n = 2$ region, and the $n = 2$ kinematic endpoint at 35.6 GeV/c is obvious. However, there is a clear deficit in the number of tracks in the $n = 2$ region relative to the number in the $n = 1$ region. The deficit is approximately a factor of two.

Note that the simulation was done for each event using focal parameter data from the Cherenkov monitor fits[Chapter 5] using the "new" calibration constants. In both the "new" focal parameter estimates and the "old" (scaled) estimates, a fixed relationship between focal area and pulse length was maintained: $A = (20\ \mu\mathrm{m}^2/\mathrm{ps})\tau$. Because of the crossing angle of the colliding beams, the laser focal area actually determines the electron's path length. Thus, as area and pulse length are increased at fixed energy, the drop in rate due to decreased intensity is partly offset by the increase in path length.

Instead of using the Cherenkov-monitor-derived laser intensity information to calculate the theoretical spectra, it is possible to introduce laser pulse length and/or area as another free parameter. With focal intensity thus adjustable, the observed spectra in the CCDs can be fit quite well; however, it is apparent from Fig. 1.17 that a significantly lower intensity must be required in order to fit the data. In order to study the relationship between CCD-derived laser intensities and Cherenkov-monitor-derived intensities, the data sample is split into five approximately-equal subsets according to the Cherenkov-monitor-based η measurement. Figure 1.18 plots the CCD-based

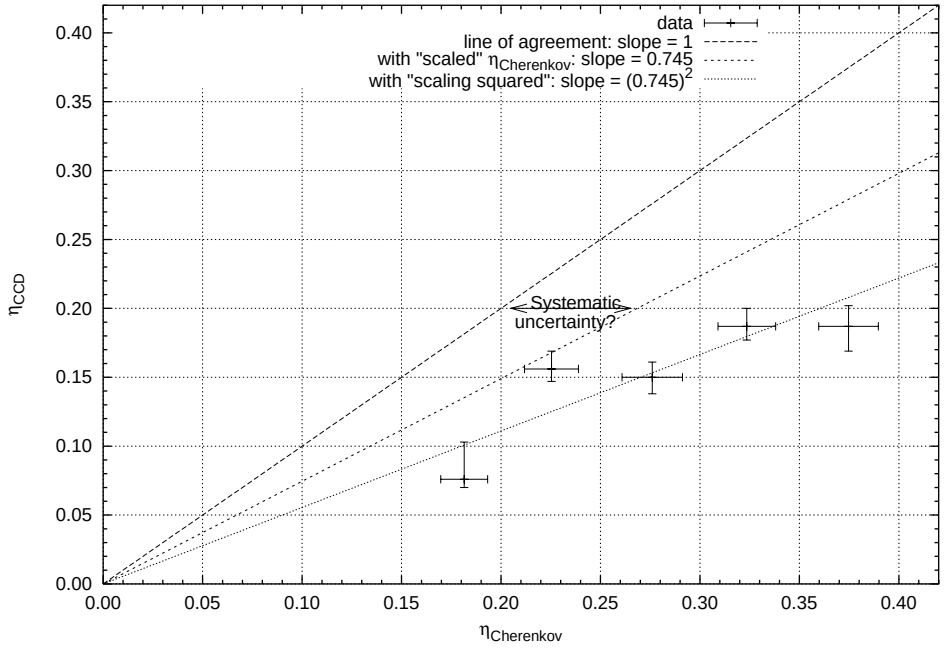

Figure 1.18: CCD-based η fit result versus the mean value of Cherenkov-fit-based η in each data subset: vertical error bars are from the CCD data fit, and horizontal error bars are the rms of the distribution of Cherenkov-fit η.

η fit result versus the mean value of Cherenkov-fit-based η in each data subset. In fitting the CCD track spectra, the relationship $A = (20\ \mu\mathrm{m}^2/\mathrm{ps})\tau$ has again been maintained, for consistency.

It is immediately obvious in Fig. 1.18 that the CCD data is consistent with a lower values of η overall, compared to the electron data. The values of η from the CCD fits are almost a factor of two different from the Cherenkov monitor values. (*N.B.* if laser energy were the unmeasured, free parameter, then only a factor of $\sqrt{2}$ change in η would be required to effect a change in the $n = 2$ to $n = 1$ ratio, but a larger change is required when energy is known and the geometry is adjusted.) If the theory is correct and the Cherenkov monitors are properly calibrated, then the two spectra should be consistent with the same value of η; this expected agreement is represented by a line with slope 1 on the graph. Also plotted is a line with slope 0.745, to represent the result that would be obtained if the "old," scaled Cherenkov monitor η fits were used. As suggested in the previous section, this difference may be adopted as a new estimate of the systematic uncertainty in η, effectively doubling the estimate on the relative systematic error in η from $\pm 15\%$ to $\pm 30\%$, *i.e.*, a factor of $e^{\pm 0.3}$. Even with the larger systematic error estimate, the significance of the difference in the spectrum is two sigma, as indicated by the line with slope $(0.745)^2$.

It should be noted that the data set for the CCD analysis is significantly smaller than the data set used for the positron analysis, and largely does not overlap with the positron data. This is partly because much of the CCD data-taking was devoted to experimenting with different converter targets, and ultimately only the thickest target data yielded sufficient rate for the single-arm analysis. It so happens that all of the thick target data was taken at approximately the same laser energy. In contrast, the positron data set contains a variety of laser energies. One possible explanation for the lack of correlation seen in Fig. 1.18 is that the statistical fluctuations in the

Cherenkov monitors are much larger than believed, and therefore lead to a wide spread in the focal parameters reconstructed for each event. Note that the residuals of the Cherenkov monitor fits are consistent with the assumed fluctuations, so the additional variance would have to come from a *correlated* fluctuation in the two $n = 2$ and $n = 3$ electron detectors, or possibly large *correlated* fluctuations in the $n = 1$ monitors EC31 and EC37 which are used to cancel overlap variations in the other two detectors. Such a purely instrumental variation in the Cherenkov-monitor-based η would not be expected to correlate with the CCDs, and separating events based on the Cherenkov monitors would indeed lead to five essentially randomly-chosen subsets each characterized by the same true value of η. According to this argument, the positron data shows the expected behavior because the variety of laser energies provides a physical variation in η driven by a well-monitored variable. The CCD analysis was performed for all runs for which the thick target data was available, but unfortunately, the positron analysis was performed only for certain selected subsets of the data, which has little overlap with the CCD data. Therefore, this hypothesis could not be tested by replotting Fig. 1.16 using the data sample and event selection used for Fig. 1.18.

Another possible source of systematic error is misbehavior of the track reconstruction algorithm in the presence of the high track densities experienced in the $n = 1$ portion of the track spectrum. This issue has been carefully studied by comparing the performance of the algorithm on Monte-Carlo data sets with its performance on real data. No evidence for significant misbehavior is seen under any conditions similar to those encountered in the real data. For example, it is shown in [Chapter 7] that the total number of reconstructed tracks in the $n = 1$ region tracks the $n = 1$ rate as monitored by the Cherenkov monitors precisely, arguing against any saturation effect in the densely-hit portion of the CCDs.

In conclusion, the overall ratio of $n = 2$ to $n = 1$ tracks produced by converted backscattered photons is a factor of two lower than the theoretical prediction using the cross-calibrated Cherenkov monitor, a disagreement at the "two sigma" level.

Even with these uncertainties, the CCD spectrometer results do provide clear evidence of the presence of backscattered photons with energies exceeding the $n = 1$ Compton kinematic limit, unambiguously demonstrating the nonlinear Compton scattering process (Eq. 1.1).

The CCD system, track reconstruction algorithm, and track spectrum are discussed in detail elsewhere [Chapter 6, Chapter 7, Chapter 8].

1.4 Conclusion

A colliding beam experiment involving non-linear processes is interesting in that the results are affected more significantly and in a more fundamental way by the intensity and shape of the beams than is the case in more familiar scattering experiments. Over the history of this experiment, the degree of sophistication increased in the approaches to addressing the novel challenges of the non-linear colliding-beam experiment. In particular, the initial reliance on error-prone remote monitoring of the laser focus was replaced by reliance on fixed-momentum particle counters (the Cherenkov monitors) to estimate the moments of the intensity distribution; the scale factors of these devices were initially freely adjusted, but later fixed by a cross-calibration with the electron calorimeter; and, the estimates of systematic errors and the overall interpretation of the results has evolved accordingly.

In the preceding sections, previously-published results were summarized, and new data presented. The new data presented consists of the analysis of the data from the CCD tracking

spectrometer, giving new information on the forward photon spectrum, plus a re-analysis of the data from the electron Cherenkov monitors intended to resolve an overall scale uncertainty in the Cherenkov monitor calibration. In Table 1.4, one may see that the CCD data has the best momentum resolution and highest precision of any of the detectors in the experiment. When one improves the precision and resolution of an experiment and addresses a systematic uncertainty in a direct way, one generally expects the uncertainties in the experiment to diminish and the significance of the results to improve. To some extent, the opposite has happened here: previously unstated or under-emphasized sources of systematic errors are emphasized in the presentation of the newest results, and should be considered when interpreting the previous results.

One reason for this broadening of systematic uncertainties is that the careful study of the Cherenkov monitor data, intended to resolve the overall uncertainty in their scales, has suggested the possibility of unmonitored changes in the Cherenkov monitors' gains. Even with the larger uncertainties, the $n = 2$ photon rates are two standard deviations displaced from predictions consistent with the calibrated Cherenkov monitors.

It would be extremely desirable to repeat the experiment, preferably with better control over and monitoring of the laser focal spot and pulse length, more redundant electron and photon monitors, and a larger data sample containing a deliberately-chosen range of focal intensities, with all detectors fully functional for the entire data set. Building on past experience, it should be possible to design, test, and agree upon data analysis techniques and standards by which to judge the results prior to the actual data-taking, thus eliminating any question of bias for or against the currently-accepted theory.

Overall, the experiment was quite successful. In particular, the observation of tracks created by photons above the $n = 1$ kinematic edge, which could not arise through multiple $n = 1$ scattering, demonstrates unambiguously the second-harmonic Compton scattering process,

$$e + 2\omega \rightarrow e + \gamma, \tag{1.46}$$

at field strengths up to one third of the critical field strength.

Chapter 2

Simulation of the electron beam/laser beam collision

Synopsis

The Monte-Carlo simulation of the collision of the electron and laser beams is described, and its results are compared with results obtained by numerical integration.

2.1 Goals and specifications

The basic goal of the simulation is to estimate the energy spectra and rates of photons, electrons, and positrons produced and/or scattered at the interaction point(s) of a practical high-field collision experiment, using the differential scattering probabilities per unit volume per kinematic parameters, $\frac{d^2}{dud\phi}W_{\text{Compton}}$ and $\frac{d^2}{dud\phi}W_{\text{Breit-Wheeler}}$, where Compton and Breit-Wheeler denote the processes of interest:

$$\text{Compton:} \quad e^- + n\omega \quad \rightarrow e^- + \gamma \tag{2.1}$$

$$\text{Breit-Wheeler:} \quad \gamma + n\omega \quad \rightarrow e^- + e^+. \tag{2.2}$$

The simulation could easily be expanded to include other coherent processes, *e.g.*,

$$\text{Trident:} \quad e^- + n\omega \quad \rightarrow e^- + e^- + e^+ \tag{2.3}$$

$$\text{Photon-photon scattering:} \quad \gamma + n\omega \quad \rightarrow \gamma + \gamma. \tag{2.4}$$

The simulation must take into account the possibility of multiple, sequential interactions of the above two types, which add incoherently to the sum. The simulation must also take into account practical features of the experiment such as the geometry of the colliding beams (*e.g.*, bunch shape, relative placement in time and space, crossing angle), and provide accurate predictions for many different-order processes occurring in the same collision.

2.1.1 Collision geometry

The crossing of the beams in the $\hat{z} \times \hat{x}$ planes of two coordinate systems is shown in Fig. 2.1. The distribution of the electrons $\rho_{e0}(\vec{x}_0)$ can most easily be described in the coordinate system on the left, while the shape of the laser field $\rho_l(\vec{x}, t)$ can be most easily described in the coordinate system on the right.

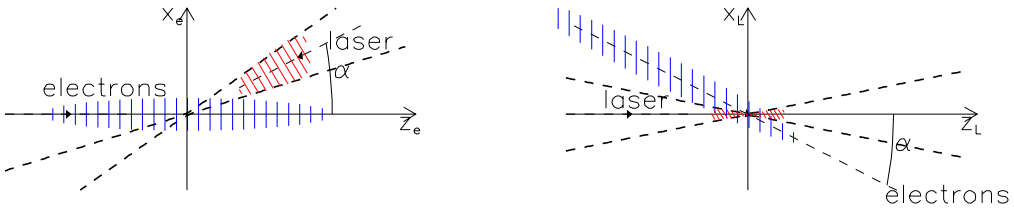

Figure 2.1: The collision geometry, top view, in electron-aligned (left) and laser-aligned (right) coordinate systems.

In the "laser-aligned" coordinate system, the density of photons at point $\vec{x}_l = (x_l, y_l, z_l)$ and time t is given by

$$\rho_l(x_l, y_l, z_l, t) = \frac{N_l \exp(-\frac{1}{2}\frac{(z_l - t)^2}{\sigma_{lz}^2} - \frac{1}{2}\frac{(x_l/\sigma_{lx})^2 + (y_l/\sigma_{ly})^2}{1+(z_l/z_R)^2})}{(2\pi)^{3/2}\sigma_{lx}\sigma_{ly}\sigma_{lz}(1 + (z_l/z_R)^2)} \tag{2.5}$$

where σ_{lx} and σ_{ly} are the RMS widths of the spot at the focus, σ_{lz} is the RMS bunch length, and z_R is the effective Rayleigh length of the focus, known as $\beta_\star$ to accelerator physicists. Table 2.1 summarizes few useful definitions and relations taken from Meyerhoffer's technical notes on Gaussian, non-diffraction limited laser focusing [33, 34] and adapted to my notation. Note $t = 0$ is implicitly defined to be the time that the centroid of the laser pulse crosses the focal plane.

In the "electron-aligned" coordinates, a Gaussian electron bunch can be described simply by

$$\rho_{e0}(\vec{x}_e) = \frac{N_e\sqrt{\det(\underline{\underline{V}}_{eb}{}^{-1})}}{(2\pi)^{3/2}} \exp(-\frac{1}{2}(\vec{x}_e - \vec{x}_{e0})^T \underline{\underline{V}}_{eb}{}^{-1}(\vec{x}_e - \vec{x}_{e0})) \tag{2.6}$$

where in the simplest case of a beam with no skew, the covariance matrix $\underline{\underline{V}}_{eb}$ is simply

$$\underline{\underline{V}}_{eb}{}^{-1} = \begin{bmatrix} 1/\sigma_{ex}^2 & 0 & 0 \\ 0 & 1/\sigma_{ey}^2 & 0 \\ 0 & 0 & 1/\sigma_{ez}^2 \end{bmatrix} \tag{2.7}$$

The vector $\vec{x}_{e0}$ is the location of the centroid of the electron bunch at time $t = 0$, when the centroid of the laser pulse crosses the focal plane.

2.1.2 Multiple scattering

In a high-field collision experiment, incident or created particles may scatter multiple times in succession within the field. Examples of possible scattering sequences are depicted in Fig. 2.2. The top diagram shows an electron radiating a photon; the second diagram shows the same electron

Table 2.1: Some relations between various optics parameters

$$
\begin{aligned}
\lambda &\equiv \text{wavelength of the laser} \\
w_0 &\equiv 1/e^2 \text{ radius of intensity at focusing optic} \\
\sigma_{lx}, \sigma_{ly} &\equiv \text{RMS width } (1/\sqrt{e} \text{ radius) of intensity at focus} \\
A_{\text{eff}} &\equiv 2\pi\sigma_{lx}\sigma_{ly} \\
\sigma_0 &\equiv \text{diffraction-limited RMS width of intensity at focus} \\
A_0 &\equiv 2\pi\sigma_0^2 \\
\alpha &\equiv A_{\text{eff}}/A_0 \\
z_{R0} &\equiv \text{diffraction-limited Rayleigh length} \\
z_R &\equiv \text{effective Rayleigh length as used in equation 2.5} \\
f &\equiv \text{focal length of the focusing optic} \\
f_\# &\equiv f/(2w_0) \\
A_0 &= \frac{2}{\pi}(f_\#\lambda)^2 \\
z_{R0} &= \frac{4}{\pi}f_\#^2\lambda \;=\; \sqrt{\frac{8}{\pi}A_0}\, f_\# \\
z_R &= \sqrt{\alpha}\, z_{R0} \;=\; \sqrt{\frac{8}{\pi}A_{\text{eff}}}\, f_\# \;=\; 4\sigma_l f_\#
\end{aligned}
$$

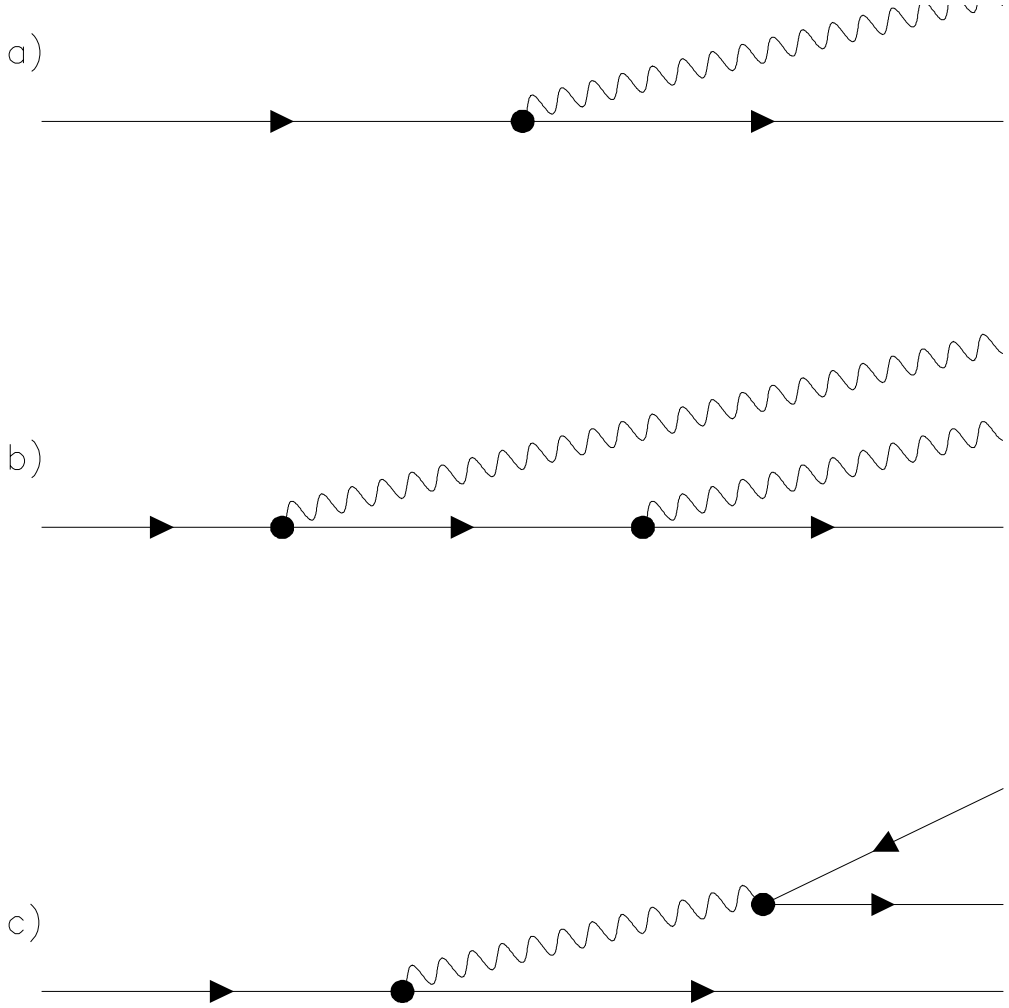

Figure 2.2: Schematic depiction of three scattering sequences. These are not intended to be Feynman diagrams: all internal particles are on-shell, and therefore the final state particle probabilities add incoherently (see text).

undergoing two Compton-like interactions in succession; and the third diagram shows a radiated photon converting to a pair.

In such sequences of scattering events, it is assumed that all internal particles are on their mass shell, so that the number of deBroglie wavelengths in the interval between successive scattering events is essentially a uniform random number. This means that the spectra from different events add incoherently. (Mathematically speaking, the cross-term vanishes when the expectation value is taken of the square of the sum of two amplitudes whose relative phase is uniformly random.) In particular, the two gammas in Fig. 2.2.b do not interfere with each other, nor do the two electrons in Fig. 2.2.c. Furthermore, the size of the laser focus is assumed to be large compared to the coherence length of all processes, so the pair produced by the two-step process depicted in Fig. 2.2.c *does not* interfere with the trident process (whose Feynman diagram admittedly looks identical to the graphical depiction of the incoherent process), nor can it interfere with anything else, due to the random phase factor introduced by the propagation of the on-shell photon created by the initial Compton scattering event. (If there were more than one channel by which a particular photon could produce a pair, then those channels would interfere; however, pair production by a photon would still be depicted as a single process in an incoherent scattering sequence diagram such as Fig. 2.2.) Interference requires coherence, and processes involving rescattering of on-shell particles are by nature incoherent.

There is an easy way of distinguishing between incoherent and coherent scattering processes: the total probability of a coherent process increases linearly with the time elapsed during which the conditions that support the process are maintained (*i.e.*, with the time the electron spends in the electromagnetic wave, in the case of Compton scattering), while the probability of an incoherent process increases non-linearly with that time (*i.e.*, the probability of an electron scattering twice increases as the square of the time spent in a laser field, etc.). The trident scattering rate, being expressed in terms of a differential probability per differential 4-volume, does not include the two-step, incoherent pair production process.

Making the distinction between coherent and incoherent processes is important, since our theory need only provide the probabilities for coherent processes, while the Monte-Carlo simulation need only sum probabilities, not amplitudes, in order to generate realistic predictions. Coherence effects can (and must) be absorbed into the probabilities $\frac{d^2}{du d\phi} W_{\mathrm{Compton}}$ and $\frac{d^2}{du d\phi} W_{\mathrm{BreitWheeler}}$.

In order to express mathematically the integrals implied by Fig. 2.2, it is helpful to express the probabilities in terms of initial and final values of $x \equiv 2(kp)/m^2$, instead of the parameter u. The parameter x can be called the "energy parameter". For Compton scattering, the following relations can be derived from the definition $u \equiv (kk')/(kp')$:

$$\begin{aligned} u &= \frac{x_{f,\gamma}}{x_{f,e^-}} \\ x_{i,e^-} &= x_{f,\gamma} + x_{f,e^-} \\ x_{f,e^-} &= \frac{1}{1+u} x_{i,e^-} \\ x_{f,\gamma} &= \frac{u}{1+u} x_{i,e^-} \\ \frac{du}{1+u^2} &= \frac{dx_{f,e^-}}{x_{i,e^-}}. \end{aligned} \tag{2.8}$$

For pair production, $u \equiv \frac{(kl)^2}{4(kp)(kp')}$ gives the relations

$$\begin{aligned} u &= \frac{x_{i,\gamma}^2}{4x_{f,e^-}x_{f,e^+}} \\ x_{i,\gamma} &= x_{f,e^-} + x_{f,e^+} \\ x_{f,e^\pm} &= x_{i,\gamma}\frac{1 \pm \sqrt{1-1/u}}{2} \\ \frac{du}{4u\sqrt{u(u-1)}} &= \pm\frac{dx_{f,e^\pm}}{x_{i,\gamma}}. \end{aligned} \tag{2.9}$$

It is satisfying that the final-state phase-space factor, which is messy and unintuitive when expressed in terms of u, always takes the form dx_f/x_i when the differential probabilities are expressed in terms of a final-state energy parameter.

The differential probability $\frac{d}{dx}W_{\rm Compton}$ for an electron with energy parameter x_i to scatter at time t and position $\vec{x}$ into a final state with energy parameter in the range x_f to $x_f + dx_f$ and azimuthal angle ϕ in the range ϕ_f to $\phi_f + d\phi_f$ will be denoted by $w_C(t, \vec{x}; x_i, x_f, \phi_f)$; a photon with energy parameter $x_i - x_f$ will be emitted in this process. Similarly, $w_B(t, \vec{x}; x_i, x_f, \phi_f)$ denotes the differential probability $\frac{d}{dx}W_{\rm BreitWheeler}$ for a photon with energy parameter x_i to create a pair at time t and position $\vec{x}$ consisting of an electron with energy parameter in the range x_f to $x_f + dx_f$ and azimuthal angle ϕ in the range ϕ_f to $\phi_f + d\phi_f$ along with its positron counterpart.[1]

Using this notation, it is relatively easy to write the integrals corresponding to the incoherent scattering sequences depicted in Fig. 2.2. (These integrals may seem somewhat unnecessarily detailed, but for a complete and quantitative description, there is no substitute for explicitly "writing it all out.") The number of electrons scattered into the range $x_1 \le x < x_2$ from a "pencil beam" with initial kinematic parameter $x_0 > x_2$, initial velocity $\vec{v}_0$, "$t = 0$"-projected position $\vec{x}_0$, and unit density is given by

$$\begin{aligned} &n_{e^-}(\vec{x}_0; x_1, x_2) = \\ &\quad \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt\ w_C(t, \vec{x}_0 + t\vec{v}_0; x_0, x, \phi) \\ &+ \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt \int_x^{x_0} dx' \int d\phi' \int_{-\infty}^{t} dt' \\ &\qquad w_C(t', \vec{x}_0 + t'\vec{v}_0; x_0, x', \phi') w_C(t, \vec{x}_0 + t'\vec{v}_0 + (t-t')\vec{v'}; x', x, \phi) \\ &+ \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt \int_0^{x_0 - x} dx' \int d\phi' \int_{-\infty}^{t} dt' \\ &\qquad w_C(t', \vec{x}_0 + t'\vec{v}_0; x0, x', \phi') w_B(t, \vec{x}_0 + t'\vec{v}_0 + (t-t')\vec{v'}; x_0 - x', x, \phi) \\ &+ \ \ldots . \end{aligned} \tag{2.10}$$

The integrand in the first term represents a single Compton scattering event which takes the electron directly from x_0 to some value x in the range $x_1 < x < x_2$, as in Fig. 2.2.a. The second term represents the case where the initial electron scatters from x_0 to x' at time t', and subsequently

[1] Two further notes on this paragraph: first, I apologize for using x for the energy parameter and $\vec{x}$ for position, but hopefully the arrow over $\vec{x}$ will help avoid too much confusion; second, the idea of simultaneously specifying both the space-time point at which an interaction occurs and the initial and final 4-momentum of all particles might seem to be raise problem with the uncertainty principle, but in fact this is *mathematically* how the integral is done to calculate rates from cross-sections, although the *physically* the 4-momentum and 4-position are conjugate observables.

scatters from x' to a value of x in the given range, as in Fig. 2.2.b. The third term represents the case where the gamma created when the initial electron scatters from x_0 to x' at time t', which has kinematic parameter $x'_\gamma = x_0 - x'$, subsequently creates a pair with an electron having x in the given range, as in Fig. 2.2.c. The dots indicate all possible such scattering sequences should be considered. The corresponding terms for evaluating the gamma spectrum are

$$\begin{aligned} n_\gamma(x_1,x_2) &= \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt\ w_C(t,\vec{x}_0+t\vec{v}_0;x0,x0-x,\phi) \\ &+ \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt \int_{x_0-x}^{x_0} dx' \int d\phi' \int_{-\infty}^{t} dt' \\ &\quad w_C(t',\vec{x}_0+t'\vec{v}_0;x0,x',\phi') w_C(t,\vec{x}_0+t'\vec{v}_0+(t-t')\vec{v'};x',x0-x,\phi) \\ &+ \ldots, \end{aligned} \tag{2.11}$$

and for the positron spectrum,

$$\begin{aligned} n_{e^+}(x_1,x_2) &= \int_{x_1}^{x_2} dx \int d\phi \int_{-\infty}^{+\infty} dt \int_{0}^{x_0} dx' \int d\phi' \int_{-\infty}^{t} dt' \\ &\quad w_C(t',\vec{x}_0+t'\vec{v}_0;x0,x',\phi') \\ &\quad w_B(t,\vec{x}_0+t'\vec{v}_0+(t-t')\vec{v'};x_0-x',x0-x'-x,\phi) \\ &+ \ldots. \end{aligned} \tag{2.12}$$

The total number of particles of type i with energy parameter in the range $[x_1, x_2]$ produced by the interaction of the laser with a beam having phase space density $d^6n_0/(d^3x d^3v) \equiv \rho_0^{(6)}(\vec{x},\vec{v})$ is given by

$$N_i[\rho_0^{(6)}](x_1,x_2) = \int d^3x \int d^3v\ \rho_0^{(6)}(\vec{x},\vec{v}) n_i(x_1,x_2). \tag{2.13}$$

Equations 2.10 through 2.13 explicitly show the integration over the initial positions and trajectories of the incident particles, which is of critical significance to the simulation of any practical experiment due to the great sensitivity to these parameters in the nonlinear interaction. The probabilities at a given time t are evaluated at the position of the incident particle at that time assuming it follows a straight-line trajectory between scattering events.

In short, the goal of the simulation is to evaluate the integrals 2.10 through 2.13, including the integrals implied by the dots, to within some accuracy which may be specified independently for each integral.

2.2 Implementation

Two separate programs were developed to evaluate the integrals 2.10 through 2.13. One program was written by Christian Bula to use conventional numerical integration techniques; it is described elsewhere.[26] The second simulation, described in this chapter, uses a Monte Carlo approach, designed along much the same principles as well-known shower simulation packages such as `EGS`,[32] with a few special adaptations to deal with issues such as the large dynamic range of the rates of many simultaneously-present "interesting" processes. Bula imaginatively named his numeric integration code "`NUMINT`"; the code described here has the equally imaginative name "`MCSCAT`".

In Bula's `NUMINT`, it is assumed that all particles follow parallel trajectories in the lab frame, with negligable deflection by the scattering process; due to its structure, the `MCSCAT` code is easily

able to include the divergence of the incident beam and the transverse kick from each scattering event. In fact, the divergence of the ultra-relativistic electron beam in the E144 experiment is ~ 10 μrad or less, and the scattering angles in the lab frame are of the same order, so that the effect on the calculated rate from scattering within the first interaction region (IP1) is in fact insignificant; such angular divergence effects would be more significant for experiments involving a small target at a second interaction point several meters away.

Both `NUMINT` and `MCSCAT` currently ignore the azimuthal angle ϕ: it has no meaning to `NUMINT` given that the polar angle is always zero, while `MCSCAT` currently uses a uniform random distribution in ϕ, which is precisely correct for circular polarization and less correct for linear polarization of the laser, but it is *insignificantly* incorrect for E144 where polar angle effect is insignificant. A proper azimuthal angle distribution could be implemented in `MCSCAT` for linear polarization of the laser, with some significant investment of effort. It does not appear to be necessary to do so for application of the code to the E144 experiment, the prototypical "high-field collision" experiment for our purposes.

The initial electron bunch for the E144 experiment consisted of $\sim 10^{10}$ particles, of which $\sim 10^6$ scattered via $n = 1$ processes, $\sim 10^4$ scattered via $n = 2$ processes, $\sim 10^2$ via $n = 3$, etc. Since each higher-n process covers more final-state space than the previous-n process, each n is individually interesting, even if the next lower-n process is much more likely. Because of this wide dynamic range of interesting processes, it was necessary to implement a particle weighting scheme in the Monte-Carlo code, such that the initial electron bunch could be represented by $\sim 10^4$ macro-particles of $\sim 10^6$ electrons each, while each $n = 1$-produced macro-particle might represent $\sim 10^4$ electrons and an $n = 2$-produced macro-particle might represent $\sim 10^2$ electrons, and so-on. This is likely to be advisable when simulating any nonlinear interaction phenomenon by Monte-Carlo means, *e.g.*, pair production in the focus of a future linear collider: the $n = 5$ process might be 10^4 times less likely than the $n = 3$ process, but $n = 5$-scattered particles might hit a part of the detector infinitely more sensitive than what the $n = 3$-scattered particles might hit.

2.2.1 Details of the Monte-Carlo integrator and event generator: `MCSCAT`

`MCSCAT` is written in C++, a programming language that supports many different styles of code modularization, data encapsulation, and other object-oriented techniques.[2] The data and functions related to the interaction are organized into two classes of objects: `Particle` and `Field`. This allows for easy expansion of the code to handle different particles and/or different types of fields in the future.

Each `Particle` object contains variables specifying that particle's position, velocity, energy, weight, and some other tracking information; there is also some "static" data associated with the `Particle` class that is the same for all particles, *e.g.*, the relative weights to assign successive

[2] A note about C++: generally, programming languages are designed around only one paradigm: for example, FORTRAN provides a type of data localization through COMMON blocks and by keeping a subroutine's local variables intact between calls, which is totally different from Pascal's emphasis on reentrancy and dynamically-allocated memory. After seven years of experimentation with C++, I have concluded that, for any given application, most of the "many different styles" C++ supports are about as useful and error-resistant as those supported by other languages, a few turn out to be more trouble than they are worth (or worse), and often there is one approach that is truly golden, much better than what could be done easily with any other language. Unfortunately, the best and worst approaches can switch places from problem to problem. The programming paradigm for `MCSCAT` did not turn out to be "golden," but it does clarify a few difficult parts of the code, it provides a high degree of protection against stupid things like type conflicts in parameter lists; the only real trouble it has caused is a perception of inaccessibility among those who don't know the language.

The `Particle` Class

What `Particle` knows (static member data)

- macro-particle weight for each generation
- min. weight cutoff
- min. energy cutoff

What `Particle` can do (static member functions)

- initialize static member data

What each `particle` knows (member data)

- type of particle (e^-, e^+, γ)
- position, velocity
- energy
- weight
- generation (for weight selection)
- ancestry (for diagnosing scattering sequences)

What each `particle` can do (member functions)

- create, initialize, and delete itself
- choose random position, direction of motion, and energy given description of the beam ellipse
- travel through a given `field` (*i.e.*, the `traverse()` routine — see Fig. 2.6)

Figure 2.3: Definition of the `Particle` class.

The `Field` Class

What each `field` can do (member functions)

- create, initialize, and delete itself
- suggest a good time step for `Particle::traverse()`
- calculate the total probability for a given macro-`particle` at a given time to scatter (or split) into child macro-`particle`(s) with a given weight (*i.e.*, the `scatt_prob_0()` function)
- generate the scattering event, causing new `particle`'s to be created (*i.e.*, the `scatter()` function)

Figure 2.4: Definition of the `Field` class.

generations. Functions associated with the `Particle` class include several pertaining to particle initialization, plus the very significant `traverse()` subroutine. (Subroutines and functions are equivalent terms in C++, and are conventionally indicated with parentheses after the function name.) When the `traverse()` subroutine is called for a given particle, with arguments that tell it what `Field` object is being traversed and over what time period, it takes appropriate actions to propagate the particle through the field, generating Monte-Carlo scattering events as it goes, and recursing into `traverse()` for any secondary particles created by interactions, ultimately returning after modifying the time argument and reducing the weight of the original particle as appropriate to account for any child-particles it may have spawned. When stepping the particle through the field, `traverse()` takes care never to step the particle over a distance larger than `max_dt`, in order to avoid stepping into a qualitatively different region of the field, a behavior analogous to the way in which EGS truncates a step if it would take the particle into a different medium. The definition of the `Particle` class is shown in Fig. 2.3.

The `Field` class is summarized in Fig. 2.4. The most important functions in this class are `scatt_prob_0()`, which gives an estimate on the overall probability of scattering of a particle in the field, and `scatter()`, which generates a Monte-Carlo scattering event. (How these two routines are used by `traverse()` will be described below.) The `Field` class declares all of its functions to be "virtual", which essentially means that the class keeps static pointers to its functions, which derived classes can define differently if desired. Multiple classes can be derived from the `Field` class, each with its own additional data and function members and each defining its own specific implementation of the "virtual" functions associated with the `Field` base class. In the current version of the code, only a `Laser` class is derived from `Field` (Fig. 2.5). If it was desired to extend the code to handle some new type of interaction, another class could also be derived from `Field` if desired, perhaps a `PlasmaLens` class, in which case a `PlasmaLens` object or a `Laser` object could be passed as the first argument of `Particle::traverse()`. The `traverse()` function, without needing to know what kind of field was passed to it, would still be able to call the appropriate `scatt_prob_0()` and `scatter()` functions in such a way that they would be able to access the appropriate object's internal data.

The basic steps taken by `traverse()` are as follows:

1. Call the `scatt_prob_0()` routine for the field in question to get a quick estimate of the total

The `Laser` Class

Inherits data and function members from

- the `Field` class (*i.e.*, `Laser` is a kind of `Field`)

What each `laser` knows (member data)

- properties: polarization, peak intensity, Rayleigh length, focal area, pulse length, wavelength, crossing angle, positron weight, extra weighting factor to apply to macro-particles created by non-linear processes
- internal "convenience variables" to speed calculations
- all necessary lookup tables

What each `laser` can do (member functions)

- create, initialize, and delete itself
- initialize internal variables and load lookup tables according to desired properties
- laser-specific implementations of `scatt_prob_0()` and `scatter()`, incorporating generalized Compton scattering and Breit-Wheeler pair production

Figure 2.5: Definition of the `Laser` class.

interaction probability per unit time. This "quick estimate" should either be an exact value for the total interaction probability (*e.g.*, from a look-up table, for speed), or an upper limit which can be corrected by Monte-Carlo event rejection in a later step. (Note: more than one field in a region could be supported by calling the `scatt_prob_0()` routine for each field and making a random decision which field to use based on the relative probabilities. This would require passing a list or array of `Field` objects to `traverse`, which would be easy to do.) Assign the value returned by `scatt_prob_0()` to the variable `p0`.

2. Generate an exponential random number with mean $1/$`p0`, and assign this value to `dt`. This is done using `dt` $= -\log(1 -$ `dran11()`$)/$`p0`, where `dran11()` returns a uniform random number in the range $[0, 1)$.
3. If `dt > max_dt`, then replace `dt` with `max_dt`, and set a flag to indicate that `dt` has been clipped.
4. Advance the particle position by $\vec{\mathtt{v}} \cdot$ `dt`, and advance the time clock by `dt`.
5. If `dt` was clipped in step 3, go back to step 1 now.
6. Since the step was not truncated, call the `scatter()` routine for the appropriate field. The `scatter()` routine has ultimate responsibility for deciding whether or not an interaction occurs, and if so, in what way. (*E.g.*, what value of n for the multiphoton Compton process, and what the scattering angles will be, etc.) The weight of macro-particles created or scattered by this process may be less than the weight of the original macro-particle, so when the

interaction destroys particles (*e.g.*, the photon in a pair-creation process), `scatter()` decrements the weight of the original macro-particle by the weight of the new macro-particles, and when the interaction scatters particles without destroying them (*e.g.*, the electron in a Compton process), `scatter()` creates a new macro-particle of the same type as the original macro-particle and decrements the weight of the original macro-particle by the weight of the new macro-particles.

7. For each new particle created, if any, call the `traverse()` routine (recursively) to propagate the new particle from the time and place of its creation to the specified ending time.

8. Repeat the above steps until the particle reaches the specified ending time.

A flow chart of the algorithm described above is shown in Fig. 2.6.

Steps 1 through 5 generate a random step `dt`, establishing a known probability distribution for the `scatter()` routine to be called. The `scatter()` routine can tailor that probability distribution by refusing to allow a scattering event in some cases.

It is worth contrasting the behavior of `traverse()` with a less efficient but better known technique. For simple simulations where the computational cost of calculating `p0` is small, one can take steps `dt` such that $\texttt{dt} \cdot \text{p0} \ll 1$, generate a uniform random number u_i in the range $[0, 1)$ at each step i, and proceed to the scattering routine in the (rare) event that the random number is smaller than the interaction probability $\texttt{dt} \cdot \text{p0}$. In the limit where `dt` goes to zero and `p0` is constant over the region, this is precisely the same as generating an exponential random number; generating a single random number is simply much more efficient. If `p0` varies, it is necessary that `dt` be small enough that the change in `p0` over the interval be small. Limiting `dt` to be less than `max_dt` accomplishes this. In fact, if `max_dt` is made small enough, the `traverse()` algorithm reduces to the less efficient "very small `dt`" technique. However, with the `traverse()` algorithm, it is only necessary that `max_dt` be small compared to the size scale of fluctuations in `p0`, so `max_dt` does *not* have to be small compared to $1/\texttt{p0}$, which is a good thing since `p0` can be extremely large in some locations in the interaction region of experiments involving non-linear QED processes. That this approach gives the correct answers can be demonstrated by considering the probability for an exponential random variable with mean $1/\texttt{p0}$ to exceed `max_dt`, which is $\exp(-\texttt{max\_dt} \cdot \texttt{p0})$. Therefore, the probability for `traverse()` to propagate the particle over M steps without scattering will be $\exp(-\texttt{max\_dt} \cdot \texttt{p0[step 0]}) \cdot \exp(-\texttt{max\_dt} \cdot \texttt{p0[step 1]}) \cdots \exp(-\texttt{max\_dt} \cdot \texttt{p0[step M-1]})$, which is a good approximation to $\exp(-\int dt\, p(t))$ if `max_dt` is small compared only to the size scale of variations in `p0`. This is the correct result.

Having sketched the broad outlines of the code, the following subsections are devoted to the intricacies of calculating the nonlinear Compton and Breit-Wheeler cross-sections, generating lookup tables, using these lookup tables in the `Laser` class's `scatt_prob_0()` and `scatter()` functions, and collecting the results in some useful way.

2.2.2 Calculation of scattering probabilities

Four functions are used to calculate scattering probabilities: `mcompton()`, `mbreitwheeler()`, `mcomptonlin()`, and `mbreitwheelerlin()`. Each accepted 4 parameters, the first two specifying the scattering to occur by number of photons absorbed and the final-state kinematic parameter, and the second two specifying the conditions of the interaction by η and $x_0 \equiv 2(kp)/m^2$, where p is the 4-momentum of the electron or photon in the initial state. The final state kinematic parameter

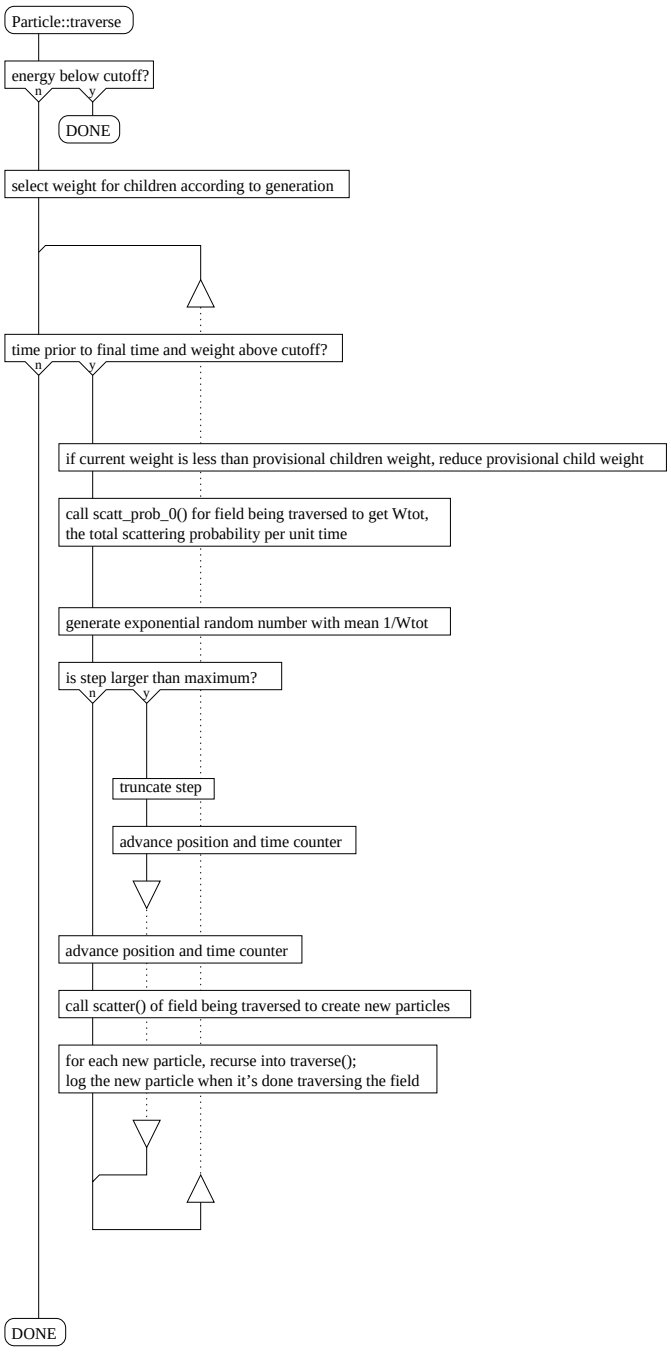

Figure 2.6: Flow chart showing the implementation of the `traverse()` routine.

was expressed in terms of x', defined as

$$x' \equiv \frac{u}{u_{\max,n}} \qquad \text{for Compton,} \tag{2.14}$$

$$x' \equiv \frac{u-1}{u_{\max,n}-1} \qquad \text{for Breit-Wheeler.} \tag{2.15}$$

Note $0 \leq x' \leq 1$.

The functions return a normalized, invariant, dimensionless number, which can be converted to the probability per unit time of scattering into a final state between x' and $x' + dx'$ by multiplying the return value by the quantity $e^2 m^2 N/(8\pi q_0)$, where e is the electron charge, m is the electron mass, N is the incident particle density in the reference frame of interest, and q_0 is the incident particle total energy, including the transverse kinetic energy in the case of an electron, *i.e.*, it is the energy component of the quasi-momentum in the case of a charged particle, and simply the energy for a neutral particle.

The definitions of x' given above were originally chosen to improve the performance of the event generation algorithm (described below) by flattening the differential probability over the range of x'. It turns out that $x \equiv (kk')/(kp)$ would have been a superior choice, where p is the incident particle 4-momentum and k' is the final state 4-momentum of a chosen particle. Not only would using x instead of x' result in a flatter probability spectrum, improving the efficiency of the random event generator, the return values of the normalized, invariant functions would be more physically-meaningful, since as discussed previously, dx would represent a constant amount of the final-state phase space at all x: the normalized, invariant differential scattering functions would reflect only the matrix element (dynamics), and kinematics would be entirely contained in x and the normalization factor. However, the performance obtained by using x' is adequate, and it was deemed not worth the effort to re-code these routines and the code that calls them.

2.2.3 Generation and use of probability look-up tables

The functions `mcompton()`, `mbreitwheeler()`, `mcomptonlin()`, and `mbreitwheelerlin()` are not used directly in the `Laser` class's `scatt_prob_0()` and `scatter()` functions. Instead, a pre-processor is used to generate look-up tables for the following:

1. $W_{tot}(\eta, x_0)$, the total scattering rate as a function of η and x_0;
2. $P_n(n'|\eta, x_0)$, the probability of an event having a number of absorbed photons $n \geq n'$, given η and x_0; and,
3. $F_{x'}^{-1}(y|n, \eta, x_0)$, the inverse probability distribution functions in x' given n, η and x_0.

Again, this approach is similar to that used in EGS4[32], which also uses a combination of a pre-processor to generate look-up tables and a separate Monte Carlo program that uses the tables to generate events quickly.

The use of the first of these tabulated functions, $W_{tot}(\eta, x_0)$, should be obvious: this is exactly the value `scatt_prob_0()` is supposed to return. All `scatt_prob_0()` need do is calculate η and x_0 for the particle's current time and position within the laser pulse, and look up the appropriate value in the table.

The second two tables are used by `scatter()`, the event generator. Since `scatt_prob_0()` returns the exact probability of event occurrence, there is no need for `scatter()` to correct the

probability by the rejection method, and all it has to do is chose the values of n and x' for the event. Note that the values of η and x_0 have already been calculated by `scatt_prob_0()` at the time `scatter()` is called.

Note that $P_n(n'|\eta, x_0) - P_n(n'+1|\eta, x_0) = p_n(n'|\eta, x_0)$, the conditional probability of an event with exactly n' absorbed photons. Therefore, finding a value of n is straight-forward: `scatter()` obtains a uniform random number r in the range $[0, 1)$, then finds the highest value of n' for which $P_n(n'|\eta, x_0) > r$. The probability of an event with n' photons will then be exactly $p_n(n'|\eta, x_0)$, as desired.

The use of the final table by `scatter()` is perhaps the least obvious to the uninitiated, so a brief review of the direct generation method for continuous random variables is in order.

First, the probability distribution function is defined as [35]

$$F_{x'}(x''|n, \eta, x_0) \equiv P(x' \leq x''|n, \eta, x_0), \tag{2.16}$$

the probability that the random variable x' will be less than or equal to x'', given n, η, and x_0. Thus, there is a 50% probability of x' having a value lower than the value $x''_{0.5}$ at which $F_{x'}(x''_{0.5}|n, \eta, x_0) = 0.5$, a 10% chance of finding $x''_{0.7} < x' \leq x''_{0.8}$, where $F_{x'}(x''_{0.7}|n, \eta, x_0) = 0.7$ and $F_{x'}(x''_{0.8}|n, \eta, x_0) = 0.8$, etc. Therefore, random values of x' can be generated with the correct distribution from uniform random numbers y in the range $[0, 1)$ by finding the value x''_y for each r at which $F_{x'}(x''_r|n, \eta, x_0) = y$: these x''_y are the random values of x' desired. Since the probability distribution function must be inverted to find these values, it is efficient to have the inverse function tabulated in the first place.

Thus, after generating n, the function `scatter()` obtains a uniform random number y in the range $[0, 1)$, then looks up the value $x' = F^{-1}_{x'}(y|n, \eta, x_0)$.

Having generated n and x', a random value for the azimuthal angle ϕ is generated, currently from a uniform distribution. Then `scatter()` creates new particles on the stack, initializing them with momentum and other properties as strictly determined by η, x_0, x', and n.

2.2.4 Weighting

It is neither economical nor necessary to track every individual electron in a bunch of 10^{10} electrons. In the simulation, every "particle" is really a "macro-particle," a single entity representing a large number of actual particles. The "weight" of a macro-particle is the number of actual particles it represents. In this experiment, it so happens that no single weight is equally good for all particles. For given experimental conditions, the laser might scatter 10^7 electrons by the $n = 1$ process, 10^5 by the $n = 2$ process, 10 electrons by the $n = 4$ process, and have an average pair production probability of 0.001 per pulse, whereas the incident bunch might consist of 10^{10} particles. Since results of comparable fractional accuracy are desired for each process, and since the macro-particles in the incident bunch simply serve to sample the beam shape, it is desirable to use less-weighty macro-particles for the higher-order processes.

Therefore, the weight of new macro-particles is determined according to their generation and the order n. The weight of new macro-particles is always less than or equal to the weight of the parent macro-particles. Scattering events thus appear like "fissioning" events in the simulation: for example, Compton scattering in the case where the initial particle has a weight of 1000 and the daughter particles have weights of 100 is treated as the splitting of the 1000-particle macro-electron into a 100-particle macro-photon, a 100-particle recoil macro-electron, and a 900-particle macro-electron having all the same properties as the original macro-electron. The probability of

such an event is 10 times higher than the probability for a single electron to Compton scatter. Renormalizing the look-up tables to account for this is straight-forward.

2.2.5 Initialization, accumulation of results, error estimation, and output

A driver routine accepts input specifying the situation to be simulated in a standard format which is based on a C++ implementation of the Fortran "namelist" capability. This input data includes description of the laser and electron beam parameters as well as information specifying what weighting to use, what output to generate, and where to write output.

The default behavior is to accumulate histograms of the final energies of each type of particle: electron, photon, and positron. The histogram bins contain the sum of the weights of all macro-particles of the given type and energy. In order to estimate errors, a second set of histograms is accumulated with the sum of the squares of the weights of the macro-particles. From these two histograms, the average weight and effective number of macro-particles can be calculated for each bin, from which the likely statistical error can be estimated.

All output is written to human-readable text files, for ease of processing.

2.3 Tests and comparisons to validate results

The predictions of MCSCAT have been carefully compared with those of NUMINT, and found to be in agreement. Each of the laser focal parameters (area, pulse length, and pulse energy) were scanned while the others were held at their "standard" values, and the total number of scatters by each process at each n were compared, as well as the total count of electrons in three different momentum ranges (chosen to correspond roughly with the Cherenkov monitor acceptances [Chapter 4]), and the width and peak momentum of the positron distribution.

The agreement between the two completely different simulations suggests that each is working correctly and that each gives valid results. In the final analysis, the numerical intergration was quicker and easier to use, and the special capabilities of the Monte Carlo were not needed, so the look-up tables generated by NUMINT were generally used for comparing experiment to theory.

Chapter 3

Electron beam diagnostics and controls, and more

Synopsis

The electron beam layout and diagnostic instrumentation are described in detail. Some important detectors not necessarily or exclusively used for beam diagnosis but read out by the same data acquisition subsystem are also described.

3.1 Beamline

The E-144 experiment took place in the Final Focus Test Beam (FFTB)[19, 20], the central beamline at the end of the two-mile-long Stanford linear accelerator. The *raison d'être* of this beam facility is to demonstrate the demagnification of the linear accelerator beam by a factor of 300 or more, for purposes of Next Linear Collider (NLC) accelerator research and development. However, this primary purpose was largely of no significance to E-144. At the inception of the experiment, the FFTB was the only soon-to-be-available beamline capable of accepting the highest energy electrons the accelerator could produce.[1] This and E-144's relevance to different aspects of NLC design motivated the choice of FFTB for E-144. The decision to conduct E-144 at FFTB was made prior to the construction of the FFTB, and allowances were made in the FFTB design to make it as compatible with E-144 as reasonably achievable.

The layout and use of the beamline is described in detail below.

3.1.1 Layout

The FFTB has a complex arrangement of first- and second-order beam optics to bring the beam to a tight focus in the FFTB's primary focal region. The minimum attainable transverse spot size is proportional to the square root of the longitudinal length of the focus, β^*, and so the beamline is designed to make β^* very small in the focal region. (Specific numbers will be given later.) This translates into large angular divergence of the beam after the focal region. Five quadrupoles after the focal region refocus the beam to achieve manageable angular divergence and β^*. The E-144 experimental region[2] is downstream of these final quadrupoles. From a beam optics viewpoint, the

[1] A few years later, the A-line 46.6-GeV upgrade extended another beamline's maximum energy to the required levels, but other aspects of the A-line would have been less suitable than E-144.

[2] Insensitively referred to as "the dump line" in most FFTB documentation.

E-144 region is essentially a 45-m-long drift space with some dipole bending magnets. Fig. 3.1 is a block diagram of the FFTB beam optic system.

3.1.2 Beam optics

The FFTB apertures and optics are designed for a linear-collider style "flat" electron beam, meaning that the vertical emittance (ϵ_y) is an order of magnitude smaller than the horizontal emittance (ϵ_x): generally, $\epsilon_y \lesssim 3 \times 10^{-11}$ m-radian and $\epsilon_x \lesssim 3 \times 10^{-10}$ m-radian. In the absence of higher-order aberrations, the spot size is given by $\sigma = \sqrt{\beta\epsilon}$, where β_y is as small as 0.1 mm at the FFTB focal point when seeking the smallest spots, and typically 10 or 100 mm during FFTB alignment; β_x is typically 10 mm.

Smaller spots at the FFTB focal point translate into larger spots at the E-144 interaction point (IP1) and larger spots in the final doublet quadrupoles. For low-n studies, a large electron beam at IP1 was desired to reduce the electron beam density and thus lower the rate, extending the range of laser intensity that could be scanned. The FFTB "small spot" configuration would seem to be ideal at first glance, but this configuration tends to have high backgrounds due to the large beam sizes upstream, particularly in the final doublet. In order to minimize backgrounds, the match into the final quadrupoles was adjusted to keep the beam size in the quadrupoles smaller than would be the case during normal FFTB operation; effectively, the beam waist was moved upstream from the focal point to be nearly inside the doublet itself; the "dump line" quadrupoles were also adjusted to minimize the divergence of the beam at IP1 and to further increase the beam size at IP1 while keeping backgrounds under control.[3] Spot sizes as large as 100 μm vertically and 200 μm horizontally were achieved in this way. For later studies of high-n processes, including positron production, a high rate was desired, and therefore the smallest possible spot at IP1. For this purpose, a more conventional high-β^* configuration, such as used for FFTB alignment, was adapted, with the optical match adjusted to move the primary waist downstream towards the "dump line" quadrupoles, forming a secondary waist at IP1.[4] The modest value of β^* at the FFTB focal point kept beam sizes comfortable throughout the beamline, resulting in low backgrounds. Spot sizes as small as 30 μ in each dimension were achieved in this way.

3.1.3 Beam size measurements: transverse and longitudinal

Wires are used to measure the transverse beam size in the FFTB[36]. Most of these are "flip type" wires, in which a horizontal, vertical, or diagonal wire is inserted in the beamline and the beam is scanned across it to measure the profile of the beam along a given dimension. A microcomputer in the SLC control system adjusts one or more dipole steering corrector(s) to deflect the beam while acquiring data from beam position monitors and various scattered particle detectors such as fast ion chambers and photomultiplier tubes attached to scintillators or gas Cherenkov monitors. "Step type" wires actually move across the beam, and detector signals are analyzed as function of wire position. A set of three wires were mounted on a movable fork in the IP1 box. Their position was controlled by an E-144 data acquisition computer, and they could be used either as "flip type" wires, scanned by the SLC control system, or as "step type" wires, scanned by the E-144 data acquisition system.

[3] The work of Peter Tennenbaum in devising this solution is gratefully acknowledged.

[4] Thomas Koffas developed this solution, with help from Jim Spencer.

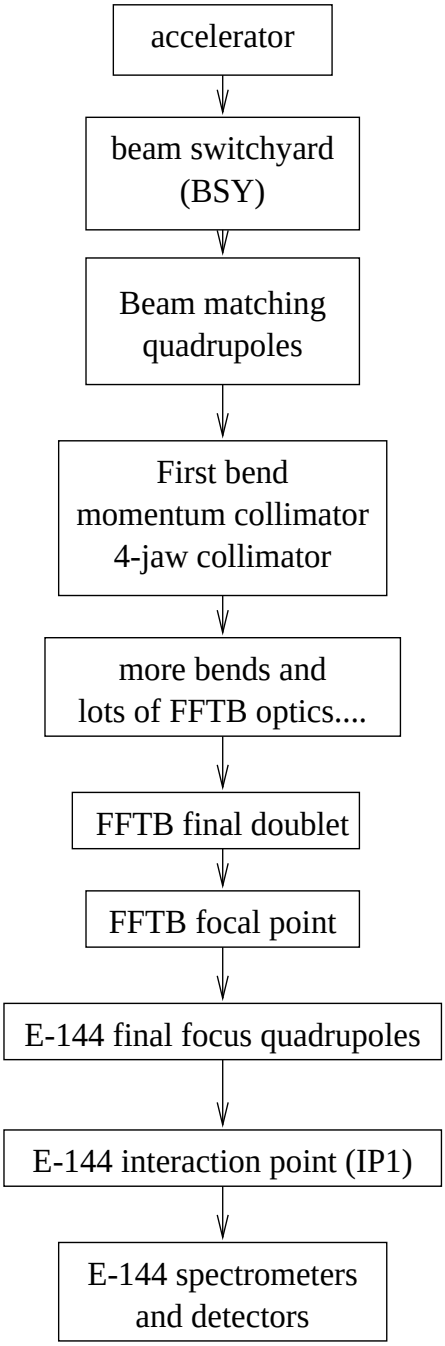

Figure 3.1: Block diagram of the FFTB beamline.

The energy spread of the beam is measured using a scanning wire near the beam dump, where the correlation between electron position and energy is high due to the action of the hard dipole bends upstream. A phosphor screen mounted at the entrance to the dump also provided an instantaneous, visual representation of the beam energy and energy spread.

The longitudinal bunch length was estimated from the setting of the compressor klystron in the north-damping-ring-to-linac (NRTL) transport line, making use of the relationship between compressor setting and bunch length as determined by Holtzapple[21].

3.1.4 Synchrotron light management

Synchrotron light is emitted by any charged particle moving in a magnetic field, and can become a significant contribution to background noise when the total energy radiated is large and the energy of the photons exceeds the lower threshold of the detector in question. The largest single generators of synchrotron light from the primary beam are the E-144 primary spectrometer dipole magnets (also known as "dump line safety magnets" or B06A-F), each of which is 94.4 cm long and bends the 46.6-GeV beam downward by 2.72 mr, equivalent to a transverse kick of 127 MeV/c, or a bending radius of 347 m. Softer bends, labeled B04A, B05A, B05B, and B04B in Fig. 3.1, are each 1 m long, and bend the beam 515 μr, 64 μr, 64 μr, and 429 μr, respectively. The critical synchrotron photon energy is[37, 38]

$$\hbar\omega_c = \frac{3\gamma^3\hbar c}{2R} \tag{3.1}$$

$$\simeq (2.22\text{keV})\ (E[\text{GeV}])^3/(R[\text{m}]), \tag{3.2}$$

where E is the energy of the beam and R is the bending radius. Specific values of $\hbar\omega_c$ are 650 keV for the hard bends, 120 keV for the 515 μr soft bend, 100 keV for the 429 μr soft bend, and 20 keV for the 65 μr very soft bends. The average energy per photon is $8/15\sqrt{3} \simeq 0.308$ times the critical energy. The mean number of photons emitted per radian of bend while traversing a magnet is

$$\frac{dN_\gamma}{d\theta} = \frac{5}{2\sqrt{3}}\alpha\gamma \tag{3.3}$$

$$\simeq 20.6\ (E[\text{GeV}]), \tag{3.4}$$

where θ is the bending angle and p_k is the transverse kick of the magnet: about 1 photon per milliradian per electron for the 46.6-GeV electron beam. Since the electron beam typically contains 7×10^9 electrons per pulse, a very large flux of hard x-rays and soft γ-rays is generated in the hard bends. The arrangement of very soft and soft bends around IP1 keeps the synchrotron light spectrum relatively soft in the forward photon line, reducing synchrotron light backgrounds greatly. The upstream soft bends also helped separate the E-144 photons from photons generated by interactions upstream, including synchrotron radiation from the strong final doublet quadrupoles and bremsstrahlung by beam tails.

Two fixed tungsten collimators in the forward photon line further reduced backgrounds[39]. The first collimator (COL1) had a square aperture 5.6 mm on each side, and was located 23 m downstream of IP1; the second collimator (COL2) had a square aperture 9.0 mm on each side, and was located 33 m downstream of IP1, half a meter upstream of the converter foil. This dual collimator system was arranged so that both direct and singly-scattered x-rays would have a clear path from COL1 to a point at least as far downstream as the inside wall of a 0.5-inch vacuum pipe.

This narrow downstream pipe, and the lead stacked around it, served to protect the CCDs against particles backscattered from or produced at the electron or photon beam dumps. Four scintillators arranged around the downstream pipe provided information regarding beam loss, aiding alignment and steering of the photon beam. Fig. 3.2 shows the arrangement of collimators, scintillators, and beam pipes, and the trajectories of photons from various sources.

3.1.5 Ground motion

The SLAC two-mile accelerator rests on stable Miocene- and Eocene-era sandstone[40, 41], which allows it to maintain good alignment of all components in all seasons under normal conditions, as well as to withstand local, occasionally severe seismic activity. However, the FFTB is located in the space originally allocated to the "C-line," an area once used for smaller-scale experiments than those conducted in the massive A and B end-stations, and later used as a storage area and parking lot. Beamline components in this area were not anchored to bedrock. Laser-based measurements of the position of a marker on the pavement made prior to construction of the FFTB showed diurnal variations of up to a millimeter [20], which would be disasterous for FFTB aligment. To address this problem, piers were sunk to the bedrock to secure all critical components of the FFTB beamline: this included all beam position monitors (BPMs) and all magnets except the "dump line" quadrupoles and permanent dipole safety magnets, plus the entire FFTB focal point, but *not* IP1. Additionally, the concrete apron (former parking lot) was "flayed" using a saw, separating the concrete into independent islands resting on the soil or rubble below, expanding and hopefully contracting more-or-less in place. The permanent magnets and IP1 were fastened to the concrete floor inside the FFTB tunnel.

The upstream end of the first permanent bend magnet came within inches of a structure mounted on the pier for BPM 6130, just downstream of IP1. In order to gauge the motion of objects fastened to the floor relative to those on fixed piers, a Heidenhein precision position transducer (model MT60k) was borrowed from the SLC beam diagnostics group and arranged to measure relative displacement of the two structures in either the vertical (Y) or horizontal transverse (X) dimension. Fig. 3.3 shows the history of positions measured over several days in January 1995. Each data point is the average of 100 measurements taken successively; the rms fluctuation of the measurements was 50 nm when the transducer was mounted horizontally and 20 nm when mounted vertically, so the mean has a statistical precision of better than 5 nm. The resolution of the transducer is 50 nm, and the systematic accuracy is claimed to be better than the resolution over the entire range of the transducer. Data was taken first with the transducer vertically positioned, then horizontally. The abrupt breaks and return to zero in the data are due to restarting of the data acquisition system. The ground positions clearly do not follow a diurnal cycle. Noting that the data was taken during a time of sunny weather after a long period of rain, one speculation is that the pavement motion is governed by hydrologic activity between the bedrock and cement.

3.1.6 Movers

There are two basic types of positioning devices known to the SLC control system: ordinary stepping motor devices, in which position is linearly related to number of steps taken, and "FFTB magnet movers," [42] specially designed and commissioned for the FFTB. In addition, there are a number of movable things specific to the E-144 experiment which were manually controlled or controlled by the E-144 data acquisition system.

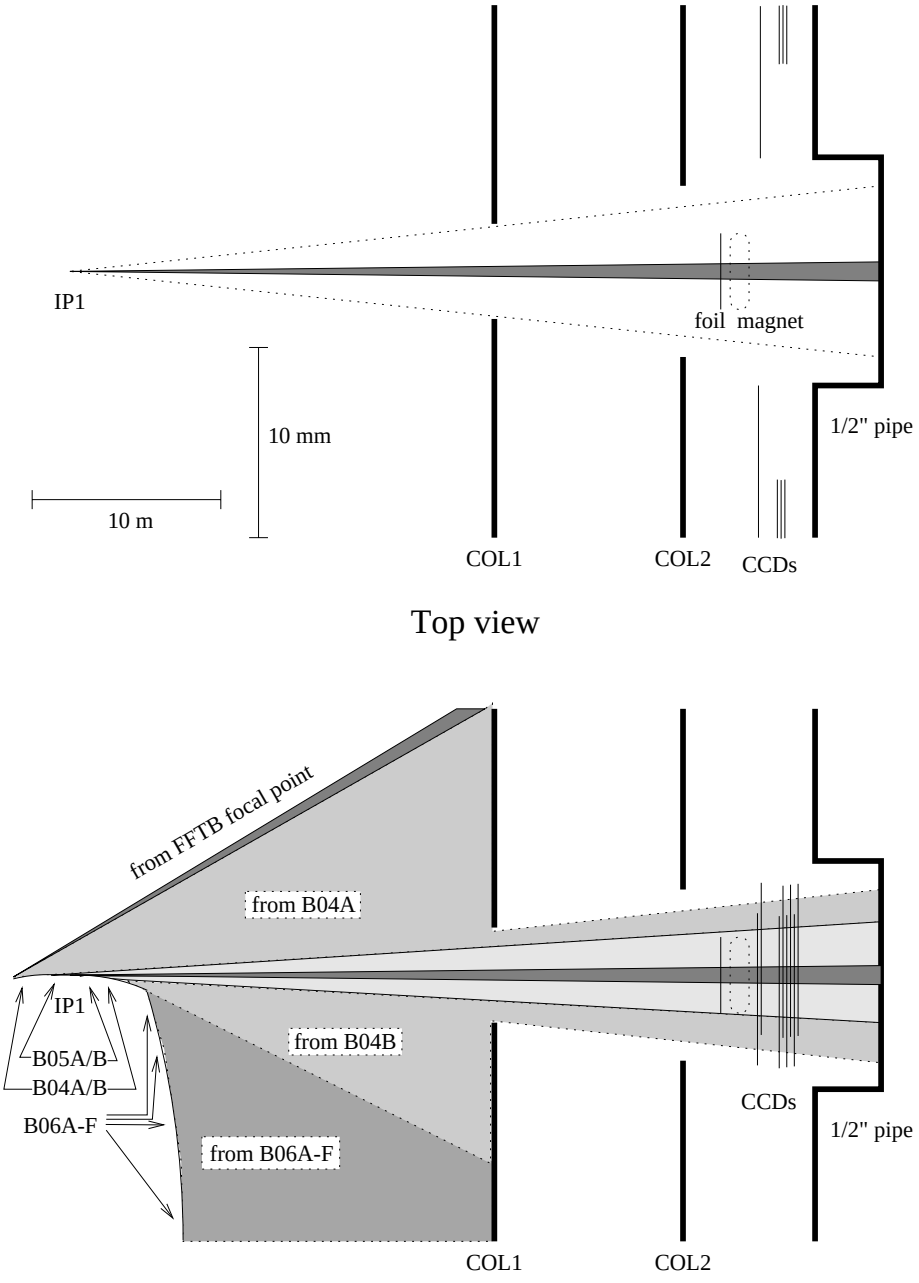

Figure 3.2: Arrangement of collimators and other objects in the forward photon line, and the trajectories of photons from various sources.

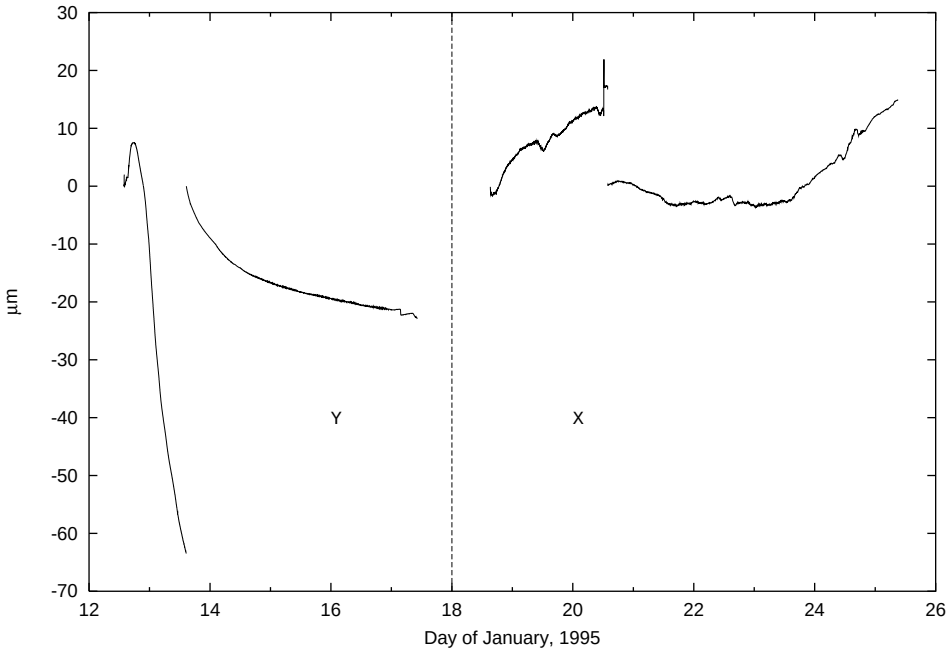

Figure 3.3: History of B06A positions measured relative to BPM 6130 over several days in January 1995.

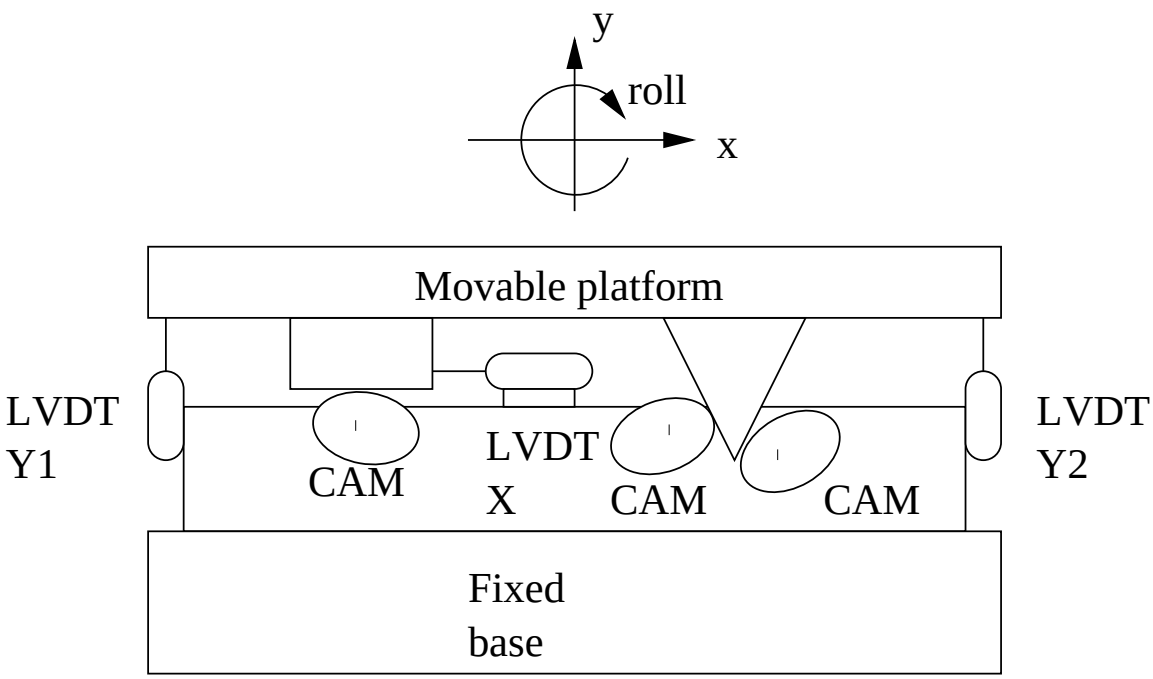

Figure 3.4: Conceptual drawing of FFTB magnet mover (not to scale).

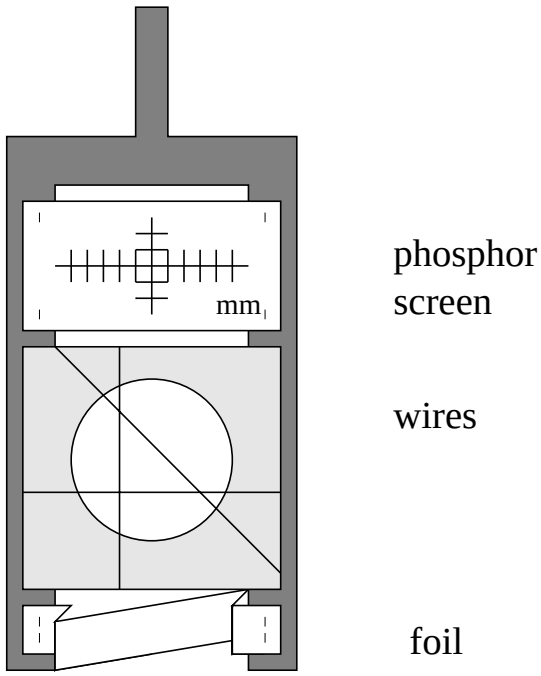

Figure 3.5: The IP1 "flag," a multi-purpose beam diagnostic device.

Every quadrupole and sextupole magnet on the FFTB beamline (with the exception of the last four of the five "dump line" quadrupoles) is mounted on a FFTB magnet mover. Each FFTB magnet mover consists of three elliptical cams supporting a platform as shown in Fig. 3.4, allowing the magnets to be moved in the vertical and horizontal transverse dimensions and rolled about the longitudinal axis. The actual position of the magnet is measured using three "linear voltage displacement transducers" (LVDTs). The complicated transformation between desired magnet position and cam rotation is managed by specialized software in the SLC control system [42].

The IP1 box is mounted on a FFTB magnet mover, allowing the laser focus to be scanned transversely across the electron beam. This was critical in maintaining the overlap of the two beams, due to the ground motion described above, and was also useful in performing xt raster scans, which demonstrate the nonlinear nature of the $n > 1$ signal and help identify $n = 1$ contamination of the detectors intended to monitor the $n > 1$ signal.

The electron calorimeter (ECAL) was mounted on a conventional stepping motor stage controlled by the SLC control system. This was used to perform ECAL scans and to position the ECAL below saturation for fixed-ECAL-position runs.

The control of these two devices from the E-144 data acquisition system was accomplished by the E-144-to-SLC interface described below.

A remotely-positionable "flag" inside the IP1 box contained a number of diagnostic devices on a single fork, including a thin aluminum foil used for generation of a stable bremsstrahlung beam used during photon beam steering and detector calibration, an array of three wires used to measure beam sizes via scans as described above, and a phosphor-coated screen used to view the beam profile and rough position relative to the laser. (See Fig. 3.5.) A video camera allowed the flag to be viewed remotely. The coarse vertical position of the flag was adjustable by a manually-controlled DC motor. The fine position of the flag could be adjusted either manually or under direct control of

the E-144 data acquisition via three stepping motors arranged to position the flag in the transverse and longitudinal dimensions. The longitudinal position of the wire array was adjusted such that a wire could fully block the alignment laser, indicating it was near the focal point. (The terawatt laser would have melted the wire.) Proper longitudinal positioning was necessary to ensure the vertically-oriented wire (used to measure the horizontal position and profile) would measure the correct relative offset between the laser and electron beams.

Other movable things in the E-144 experimental area include the Cherenkov monitor CCM1, which could be remotely inserted into or removed from the beamline, various filters and radiators in each Cherenkov monitor, and the CCDs and the photon converter foil as described elsewhere. Although not located on the FFTB beamline, it is also worth mentioning that the timing of the laser relative to the electron beam could be adjusted by the E-144 data acquisition system using a stepping motor which controlled the length of an optical trombone located in the laser's path prior to the interaction point.

3.2 Beam diagnostics and other data acquired by the "Beam computer"

One of the Intel i486-based, industry-standard-architecture-bus ("IBM-PC-compatible") front-end computers in the E-144 data acquisition system was able to control a CAMAC crate via a National Instruments peripheral card, and was able to communicate over a serial interface with the SLC control system main computer. This E-144 front-end computer was called the "Beam computer," which is somewhat of an incomplete label since Cherenkov detectors used to measure products of the laser-electron interaction were among the data acquired by from the CAMAC crate. All of the data acquired from modules in the CAMAC crate are described below, after which the serial connection protocol is described and the data acquired and devices controlled by that mechanism are listed.

3.2.1 Timing system and trigger word

As described in the Overview[Chapter 1], the pattern of events acquired was irregular: three events were acquired in every two seconds, separated by intervals of 1600 ms, 200 ms, 200 ms, synchronized with the accelerator, with the laser firing on the event preceding the long interval. Once every 6300 ms, the electron beam would be stopped well upstream of the FFTB; thus, in every 126 seconds, three events would be acquired with no electron beam present: one with the laser firing and two with no laser. This allowed signal, electron-beam backgrounds, laser-only pedestals, and baseline pedestals to be monitored continuously throughout every run.

An SLC programmable delay unit (PDU) module was configured to generate triggers for the experiment synchronized with the accelerator. Most of the PDU's 16 output channels went to gates on detectors, and these were programmed with the peculiar 1.6 s, 0.2 s, 0.2 s pattern. A few went to diagnostic devices which only acquired data on laser shots, *i.e.*, the laser diagnostics and the CCD spectrometer, and were therefore programmed to only fire on the beam pulses immediately preceding the 1.6 s interval. One was specially programmed to fire only on the "pedestal" pulses, when the beam was stopped far upstream. These triggers relied on proper operation of the local SLC microcomputer and the global pattern and timing systems.

In order to provide security against under- or over-triggering of the laser due to malfunction or operator error, the laser was not fired from a trigger programmed as above, but was instead

triggered by a signal derived PDU channel configured as a "base rate" trigger. Such triggers are uninterruptable by normal control system functions and, once initialized, operate independently of the host microcomputer, requiring only the master timing signal to operate properly, and no global pattern information, but unfortunately, the PDU hardware limits the minimum rate of base rate triggers to 10 Hz. The base rate trigger was divided down to a 0.5 Hz using a divide-by-10 and a divide-by-2 counter. The "Beam computer" was able to reset the divide-by-10 counter by strobing an output line on its parallel printer port; by this means, the 0.5 Hz laser trigger could be synchronized with the triggers intended for laser data collection, *i.e.*, the triggers preceding the 1.6 s interval.

In order to tag each event according to laser-on, laser-off, electron-beam-on, electron-beam-off, etc., and to confirm proper synchronization of the laser with the data acquisition triggers, a latching trigger register module was installed in the CAMAC crate, and the status of three triggers was acquired on every event. The three triggers were the 0.5-Hz trigger for the data acquisition system (bit 0), the upstream pedestal trigger (bit 1), and a pulse returned from the laser room indicating actual firing of the laser (bit 2). The module was gated by the 1.5-Hz event trigger. Thus, the most common trigger code was 0, indicating electron beam on and laser off; the second most common code was 5, indicating both electrons and laser were present; 2 and 7 indicated laser-off and laser-on pedestal pulses, respectively; and any other code indicated gross missynchronization of the laser.

3.2.2 Toroid charge monitor

No beam charge measuring device was installed in the FFTB beamline specifically for the purpose of monitoring the beam charge for this (or any) experiment. Each beam position monitor (BPM) provides an "intensity" signal (TMIT) to the SLC control system along with the beam position information, but BPMs are not primarily intended to measure absolute beam charge. The only other devices in the FFTB beamline capable of measuring beam charge were the Beam Containment System (BCS) toroid charge monitors designated I6 and I7. Since the BCS is a critical safety system, it seems reasonable to assume it would provide generally reliable charge measurements. The BCS group was kind enough to provide E-144 with processed toroid charge monitor signals from spare fan-out ports located in the beam containment electronics racks in the Main Control Center (MCC). This was the primary way charge was monitored during the experiment.

There are also toroids in the linac and North Ring-To-Linac transport lines, which are not BCS toroids but are read out by electronics designed especially for the SLC. These toroids are believed to be generally trustworthy, in some sense. Although a pedigree for their calibration can generally not be obtained, they are used to estimate the luminosity for SLC, which is very often within 10% of the luminosity actually observed by the SLD detector.[5] Upstream SLC toroids of this type were used to provide the absolute calibration scale of I6 and I7 prior to August 1996.

The BCS toroid charge monitor system is shown in Fig. 3.6. Toroid I6 is located in the Beam Switchyard (BSY), immediately upstream of the FFTB; toroid I7 is located near the end of the FFTB, about 1 meter below and 2 meters downstream of the CCD's, or 2 meters upstream of the dump. All BCS signals originate from or return to the Main Control Center (MCC). Each toroid has a one-turn calibration winding, connected on one end to the toroid housing and ground, and on the other to the center conductor of a standard coax connector. One or more calibration pulses are sent 360 times a second to each toroid's calibration winding to test system integrity,

[5] They have also historically been the source of the beam charge numbers quoted by operators when safety officers use beam-induced signals to set trip points on BCS devices.

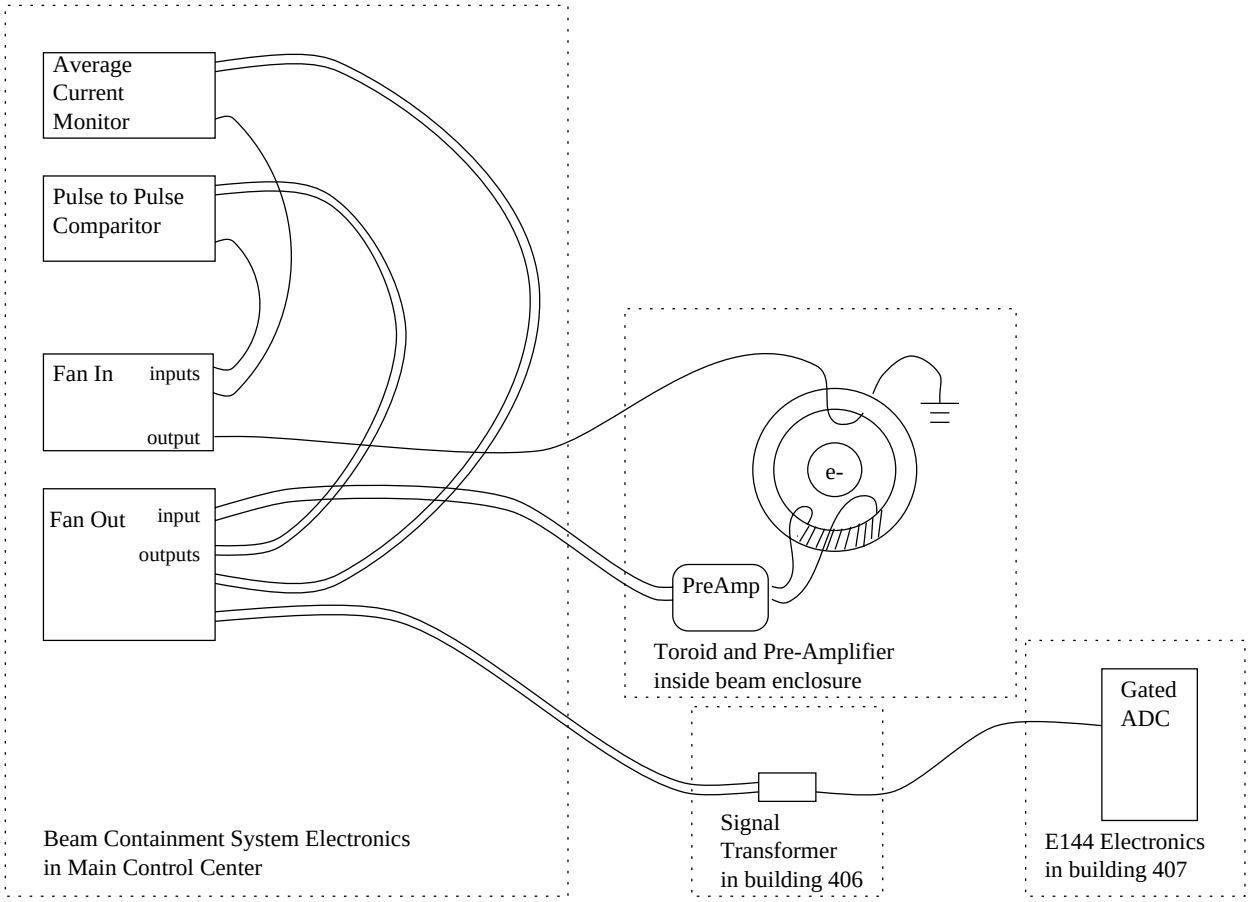

Figure 3.6: Diagram of the Beam Containment Toroid Charge Monitor System

Table 3.1: Toroid calibration data.

Date	Toro.	k	γ
March 1995	I6 (only)	12.28×10^6	0.900
December 1995	both	$\simeq 11 \times 10^6$	$\simeq 1.0$
August 1996	I7	32.50×10^6	0.894
	I6	71.09×10^6	0.894

but these calibration pulses are removed a couple of milliseconds from beam time and cause no interference. Additional signals can be injected into the calibration circuit through a front-panel connector on each fan-in module in MCC. The primary windings on the toroid are not connected to ground. The signal from each toroid is connected via twin-axial cable to a nearby preamplifier. The preamplified signal returns to MCC, again by balanced feed of a twin-axial cable, where it is processed and amplified and split by one or more fan-out modules. The output of each fan-out module also is to twin-axial cable. The BCS group identified spare fan-out ports for the I6 and I7 toroid signals and connected the balanced outputs to twisted pairs in a standard wiring trunk such as might be used for a telephone system. This trunk is generally intended for carrying DC signals from BCS interlocks to and from building 406, adjacent to the E144 counting house in building 407, but the signal quality as received at building 406 was surprisingly good. In building 406, the signal was processed by an inverting transformer to help remove common-mode noise, and then taken via standard RG-58 coax to building 407, where it was attenuated and fed into a standard LeCroy 2249A gated analog-to-digital converter (GADC) module in the CAMAC crate controlled by the "Beam computer." A single 2.5 μs gate was used to capture the entire pulse from both signals.

The system was calibrated in two ways. Prior to August 1996, the best calibration came from a two-step process in which non-linearities in the electronics were measured using signals injected into the fan-ins in MCC, an appropriate correction was applied to the toroid signals recorded by the E144 data acquisition system, and the average of the corrected signals recorded by E144 was compared to the average of the charges recorded by upstream SLC toroids.[43] In August 1996, at the start of the seven-month downtime that immediately followed the final E-144 run, the I7 calibration cable was temporarily disconnected from the I7 toroid, and a 68-nsec long pulse containing 1-nC of charge was directly injected into the calibration winding. Knowing that the calibration winding is a single turn and that the primary winding has 10 turns,[41, 44] the absolute calibration was then easily determined for I7, and the calibration for I6 could be determined from data recorded when beam losses between I6 and I7 were known to be low.

The calibration curves fit the form

$$Q = k(C - C_0)^\gamma, \tag{3.5}$$

where Q is the bunch population in particles/pulse, C is the number of raw ADC counts, C_0 is the reading of the toroid with no beam present, and k and γ are constants.

3.2.3 Beam parameters feedback data

A standard SLC beam steering feedback[45] was implemented to monitor and correct the position, angle, and energy of the beam. The feedback states were calculated to reflect these properties of

the beam at IP1. To measure the beam states, the feedback used bedrock-mounted beam position monitors (BPMs) downstream of the last quadrupole: one immediately upstream of the first soft bend, one immediately downstream of the last soft bend, and two downstream of the hard bends. These BPM locations resulted in a very simple measurement-to-state matrix, with no dependence on any quadrupole settings, so the beam trajectory states were extremely reliable. Dipole steering magnets were controlled by the feedback to maintain a standard reference trajectory. The energy of the beam was regulated by a separate feedback by adjusting klystrons in the last sectors of the linac, but the energy of the beam was independently monitored by the E-144 beam parameter feedback.

In addition to regulating and monitoring the electron beam, the feedback copied the computed state values to registers in a digital-to-analog converter (DAC) in an SLC CAMAC crate within 8 msec after every beam pulse. These analog signals were read out by the Beam computer LeCroy 2249A gated analog-to-digital converter (GADC) in the E-144 CAMAC crate on every pulse; the average levels of these signals were also monitored by a smart analog monitor (SAM) module.

Additionally, the beam phase cavity immediately downstream of IP1 was monitored by this feedback, yielding information on timing of the beam relative to the radio-frequency reference used to synchronize the laser, and the position of the beam inside a specially-built BPM inside the IP1 box itself was also read out by the feedback. These signals were not used to control any device, but average states for each were computed and stored in the SLC history buffers as well as being available on-line and through the E-144-to-SLC interface.

Thus, miscellaneous information was gathered, the beam trajectory was stabilized, and most importantly, beam trajectory information was acquired into the E-144 data stream on every pulse.

3.2.4 Miscellaneous counters read by the Beam computer

A number of particle detectors were read out by the "Beam" computer. Most important of these were the Cherenkov monitors[6]: CCM1, CCM2, EC37, EC31, N2MO, and N3MO. The Cherenkov monitors played a crucial role in measuring the total Compton photon rate and estimating the laser focal parameters, as described elsewhere[Chapter 1, Chapter 4, Chapter 5]. They were read out using GADCs in the CAMAC crate controlled by the "Beam" computer.

Other detectors read out in similar fashion include miscellaneous and experimental single-wafer silicon pads (intended to measure such things as the total beam charge, the $n = 1$ and $n = 2$ rate, and backgrounds in various places, but mostly seeing only low-energy backgrounds everywhere) and the four scintillators around the beampipe downstream of the CCD spectrometer.

3.2.5 Beam polarization readout

The SLC injector, which also supplies beam for FFTB and E-144, produces a polarized electron beam, flipping between left and right polarization randomly on every pulse. The state of the polarization is communicated on every pulse via a special carrier on the SLC cable network, and can be read out using a special CAMAC module. One early project of E-144 was to measure the SLC beam polarization [46], and there was significant interest in the possible effects of beam polarization on physics processes in E-144 [3]. For this reason, a PMON module was installed in the E-144 CAMAC crate and the Beam computer was programmed to read its state continuously and to write the polarization state to the E-144 data stream for beam pulses acquired as E-144

[6] These were generically referred to as "Clive's Cherenkov monitors," after their creator, R. Clive Field.

Table 3.2: DC signals acquired by the Beam computer.

	Channel	Description	tolerances	
			absolute [ADC Volts]	relative
0	IP1_ROT_POT_1	FFTB mover cam potentiometer	.005	.001
1	IP1_ROT_POT_2	FFTB mover cam potentiometer	.005	.001
2	IP1_ROT_POT_3	FFTB mover cam potentiometer	.005	.001
3	CCM1_HV	Cherenkov mon. high voltage	.001	.005
4	CCM2_HV	Cherenkov mon. high voltage	.001	.005
5	THEO_PHASE	Phase detector signal	.001	.001
6	ECAL_POS_POT	ECAL mover potentiometer	.001	.005
7	BEAM_ENERGY	Beam feedback state	20	20
8	BEAM_YPOS	Beam feedback state	20	20
9	BEAM_YANG	Beam feedback state	20	20
10	BEAM_XPOS	Beam feedback state	20	20
11	BEAM_XANG	Beam feedback state	20	20
12	BEAM_PHASE	Beam feedback state	20	20
13	GC0D_HV	Cherenk. mon. high voltage	.001	.005
14	EC31_HV	Cherenk. mon. high voltage	.001	.005
15	EC37_HV	Cherenk. mon. high voltage	.001	.005

events.[7] The data acquired included the state of the polarization, information on the status of the polarized source system, a check sum, and a 32-bit beam pulse identification code.

3.2.6 Miscellaneous DC signals

An SLC "smart analog monitor" (SAM) module[47] was installed in the E-144 CAMAC crate. This ingenious module combines a precision voltage reference, a 12-bit ADC, an internal 4-bit DAC, a 32-channel input multiplexer, and a programmable gain instrumentation amplifier with an on-board microcontroller to create a device capable of monitoring DC voltage and 60-Hz ripple on 32 channels, with 14 bits of floating-point precision over a range as large as ± 10.24 V with 0.6 mV resolution, or as small as ± 10 mV with 0.6 μV resolution.[8]

The E-144 SAM was used to read out a number of important DC signals, as listed in Table 3.2. A SAM data record was written to the E-144 data stream at the start of each run, and updated only when a channel varied by more than a specified tolerance from its last reported value.

Perhaps the most important voltages monitored by this device were the high voltages on the Cherenkov monitor photomultiplier tubes, since voltage has a dramatic effect on a tube's gain. Another channel monitored the output of a heterodyne mixer comparing the phase of the laser oscillator with the linac reference frequency[25], and still other channels provided redundant means of monitoring the ECAL and IP1 positions and the average feedback states.

[7] The programming in support of the PMON module was done by Christian Bula.

[8] I think SAMs are just great, in case you couldn't tell.

3.2.7 Serial port link to SLC control system

In general, many things known to the SLC control system are also of interest to the experiments to which SLC delivers beam, and many things that SLC controls over which experiments would like to have direct control. The feedback states are examples of data desired by E-144, and since the data was desired on a pulse-by-pulse basis, the DAC-to-GADC connection described above was implemented. However, it would be inconvenient and uneconomical to implement such connections for every piece of SLC data desired, and entirely unnecessary if such data varies slowly. Furthermore, due to the architecture of the SLC control system, it would be extremely awkward to implement experimenter control over SLC devices through such a means. A direct connection between the E-144 control system and the SLC control system master computer (MCC VAX) was required.

Such a connection was implemented by writing a program to be run on the MCC VAX that accepted specified messages from a terminal (TTY) port and responded either by providing information as requested or by controlling a device as directed. This program was called `E144DAQ`. The Beam computer's serial port was connected to a TTY port on the MCC VAX. Whenever the Beam computer front-end program was initialized or information was requested from the terminal, it would check to see if the E144DAQ program was running by sending a standard query to the serial port: if the `E144DAQ` program was not running on MCC VAX, a terminal emulator was started and the E-144 operator was prompted to log into MCC (if necessary) and start `E144DAQ`, after which the Beam computer would resume normal operation.

To guard against unintended actions, only specific functionality was implemented in `E144DAQ`, as opposed to creating a general-purpose command-line interface to the control system. As an additional guard against garbled communications, a strict protocol was implemented for communication between the Beam computer and the `E144DAQ` program: every message from the Beam computer to the MCC VAX consisted of a single line of form

`E2S` *n keyword data*

where the 3-character string `E2S` identified the line as a message from E-144 to the SLC control system, *n* was a 1-digit checksum, *keyword* was replaced by a 4-character command code, and *data* was additional data specific to the command. Responses from the `E144DAQ` program to the Beam computer consisted of lines of the form

`S2E` *n keyword data*

where the 3-character string `S2E` identified the line as a message from the SLC control system to E-144, *n* was again a 1-digit checksum, *keyword* was a 4-character response code, and *data* was additional data specific to the response. In practice, the terminal connection was error-free: since the system was implemented in February 1995, not a single garbled message was detected on either end.

The `E144DAQ` slept while awaiting input from the terminal, and therefore did not load the MCC VAX significantly. Since `E144DAQ` "spoke only when spoken to," the Beam computer sent an update request once every minute or so during data acquisition; however, the Beam computer accepted update responses asynchronously, so slow control system response did not normally affect the speed of the E-144 data acquisition. A record of this data was sent to the E-144 data stream only when a parameter had changed significantly. Although no autonomous responses were generated by `E144DAQ` *per se*, broadcast and process-specific messages from the SLC control system were written to the terminal and therefore appeared in the Beam computer's serial port buffer: such messages

Table 3.3: Messages exchanged between Beam computer (E-144) and `E144DAQ` program on MCC VAX (SLC control system). Items in parenthesis are replaced by the numeric or text data as described. *See Table 3.4 for definition of data indices used in `UPDT` response.

Command (`E2S`)	Response (`S2E`)
`CONN`	`CACK` 12345
`STRT` (run no.) (run type)	(updates using `UPDT` code, terminated by `SROK` code) `SROK` (run no.)
`ENDR`	(no response)
`UPDT`	`UPDT` (data index*) (data) (repeated as needed)
`BOX`x (setting in μm) [where x is either `X` or `Y`	`BXAK` (μm) `BXDN` (final x, μm) (final y, μm) (final roll, μm) (status)
`ECAL` (setting in mm)	`ECAK` (mm) `ECDN` (mm) (status)

Table 3.4: Definition of data indices used by `UPDT` response.

index	data returned
1	(IP1 x position, μm) (IP1 y position, μm) (IP1 roll, μr) (FFTB mover control process status, 4 characters) (MCC VAX system time, seconds since 00:00 Jan 1, 1970)
2	(CCD magnet current, Amperes) (ECAL position, mm) (FFTB magnet control job status, 4 characters)
3	(LTR spin rotator strength, kG-m) (RTL spin rotator strength, kG-m) (linac spin rotator solenoid strength, kG-m)
4	(feedback "hardware" status, hexadecimal) (feedback "software" status, hexadecimal)

were simply printed verbatim to the Beam computer's console, giving the E-144 operator additional information about the functioning of the SLC control system.

The commands and responses implemented are summarized in Table 3.3. They included commands to indicate start and end of run, information which was sent to the SLC log file and recorded in an SLC history buffer; commands to query the status of many things as noted in Table 3.4; and commands to change the positions of the IP1 box and ECAL. These last commands were used to implement focal position scans and ECAL position scans.

In all respects, the E-144 to SLC interface via the terminal connection worked fantastically well, providing a low-cost solution to the need to negotiate complex transactions that could be completed at a leisurely pace, and significantly expanding the range of capabilities of the experiment.

Chapter 4

Cross-calibration of E-144 Cherenkov monitors with the electron calorimeter

Synopsis

Previously-obtained calibration constants for the N2MO and N3MO detectors led to inconsistencies between our model predictions and the reported number of electrons observed by the detectors in the data from August 1996, the first run for which the detectors were available. This disagreement led us to distrust the calibration constants as given by Field[31], and to rescale those constants by large factors (about eight) when estimating the laser intensity with "constrained fits" of the Cherenkov monitor data[6]. Here, a cross-calibration of N2MO and N3MO with the electron calorimeter is presented. The cross-calibration gives results within a factor of three of the original, unscaled calibration. It is further shown that no large scale factors are required with the new calibration.

4.1 Introduction

Previously-obtained calibration constants led to inconsistencies between our model predictions and the reported number of electrons observed by the Cherenkov detectors N2MO, N3MO, EC31, and EC37 in the data from August 1996, the first run for which the detectors N2MO and N3MO were available. This disagreement led us to rescale the calibration constants by large factors: roughly a factor of 8 for both N2MO and N3MO, and a factor of 1.6 for "linear" electron counters.[6]

Our justification for this rescaling was that the ECAL data from March 1995, after "reconstruction" and normalization by the total number of backscattered photons, showed substantial agreement with the model, to within experimental uncertainties. The experimental uncertainties in the March data were roughly $\pm 15\%$ in η and $\pm 10\%$ in the reconstructed ECAL signal, leading to an overall $\pm 30\%$ uncertainty in the normalized $n = 2$ data and $\pm 60\%$ (i.e., a factor of 2) in the normalized $n = 3$ data. The rescaling required for N2MO and N3MO to obtain consistency in the August 1996 data was substantially greater than the experimental uncertainty in the March 1995 confirmation of the nonlinear Compton scattering theory, and so it was concluded that the cause of the discrepancy must lie in the calibration of the detectors. However, the large rescaling factor was not understood, and all of us found that disturbing.

It should be noted that the March 1995 results were obtained with experimental parameters somewhat different than in August 1996. All data in August 1996 was taken with the green, linearly-polarized laser, at η ranging from 0.2 to 0.4 (according to the Cherenkov monitor fits). N2MO intercepted electrons from 12.3 to 13.1 GeV/c. Our best data confirming the correctness of the accepted theory in modeling the flux of electrons scattered off the green laser into this momentum range comes from approximately 5000 events taken with green, *circularly*-polarized laser at η from 0.02 to 0.2. (Due to manpower limitations, the data taken with linearly-polarized green laser in March 1995 has not been as carefully studied as the circular.) We analyze N2MO based on theory not tested by experiment. This is not necessarily an improper thing to do, since we are still testing the overall internal consistency of the theory, but it is a point to be borne in mind.

It should also be remembered that the agreement between simulation and experiment is not particularly good in the March 1995 data, as discussed in a note presented at our September 9, 1996 meeting[48]. A χ^2/(degree of freedom) of 141/44 was calculated for the green-circular data which we still intend to present in the Phys. Rev. D paper.

Finally, an arbitrary shift of the Cherenkov monitor calibration rescaling factors can result in different predictions for η with equally good internal consistency of the constrained fits. For example, rescaling the "linear", N2MO, and N3MO counters by 1.6, 4.0, and 2.0, respectively, instead of 1.6, 8.0, and 8.0, would also give internally consistent readings but increase the intensity estimate by a factor of two.

All of these considerations and uncertainties motivated a second-look at the Cherenkov monitor calibration, this time using actual data taken in August 1996 to cross-calibrate N2MO and N3MO with ECAL, whose calibration is not in question.

4.2 Cross-calibration of N2MO and N3MO with ECAL

Below, the measurements of the acceptance functions of N2MO and N3MO are reviewed; a cross-calibration of N2MO, N3MO, and ECAL using a bremsstrahlung spectrum from wire scans and PCAL "wire data" runs is presented; a direct cross-calibration of N3MO with ECAL using green laser data is presented; and the effect of different assumed acceptance functions is analyzed.

4.2.1 The acceptance of N2MO and N3MO

The primary purpose of the Cherenkov monitors is to tell us the scattering rate of electrons into a relatively narrow range of momentum near some central momentum, *i.e.*, dN_{e^-}/dp near 12.7 GeV/c for N2MO and near 8.9 GeV/c for N3MO. For this reason, the monitors were designed with narrow acceptances and placed in the flattest portions of the electron spectrum.

The "as-built" acceptances of the monitors were measured using test-beam momentum scan data. Figures 4.1 and 4.2 show the data from these runs. [Runs 15017 and 15026 for N2MO and 15009 for N3MO.] The horizontal axis is test beam momentum as communicated to the E-144 DAQ through the "SAM" (Smart Analog Monitor) channel, and the vertical axis is ADC counts per incident electron as measured by the ECAL. The data is fitted to the sum of two square apertures "smeared" by the test-beam momentum spread and vertical size. The fitted acceptance with test-beam size set to zero is also shown.

The "standard" E-144 analysis used a single square aperture with different edges. Figures 4.3 and 4.4 show the best fit to the data obtained by varying normalization and smearing parameter

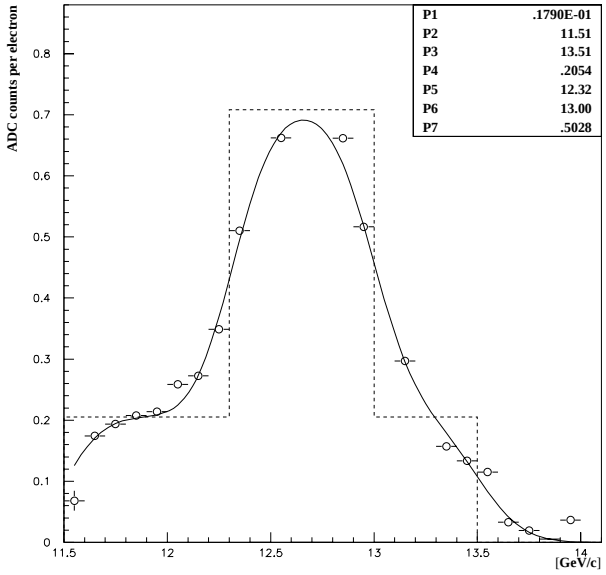

Figure 4.1: Acceptance of N2MO. Data points come from test beam momentum scans. Dashed line shows the two-box acceptance function, solid line the fit assuming a Gaussian shape to the test beam. The parameters of the fit are identified in Table 4.1.

while maintaining the edges at the "standard" acceptances. These parameters are specified in the `n2n3_par.lut` file used for the last round of n-tuple generation.

The parameters of the fit are identified in Table 4.1. The results of the two-box fit are given in Table 4.2, and the "standard" acceptance parameters are given in Table 4.3.

As noted previously, the precision of the calibration in terms of number of electrons in some unremarkable acceptance is less relevant than the accuracy of the estimate of dN_{e^-}/dp provided. The dN_{e^-}/dp result should be more-or-less unchanged by re-analysis of the acceptance. Likewise, such re-analysis should produce an equal scaling of the constant for converting ADC counts to observed electron count and of the range over which the simulated spectrum is integrated to calculate the simulated electron count. Therefore, in the following subsections, calibration constants will be given in units of electrons per GeV/c per ADC count, rather than total electrons per ADC count. Once a specific acceptance is chosen, the former type of calibration constant can easily be converted to the latter. The effect of different acceptance functions on the analysis will be discussed in a later subsection.

4.2.2 The indirect cross-calibration of N2MO and N3MO with ECAL using bremsstrahlung

In the midst of taking laser-collision data, time was taken for acquisition of "wire" data for calibration of the PCAL energy-vs-position map. This data also provides a known spectrum for calibrating the N2MO and N3MO detectors, with only one degree of freedom: the total flux of electrons in the data sample. This single unknown is well measured by the ECAL.

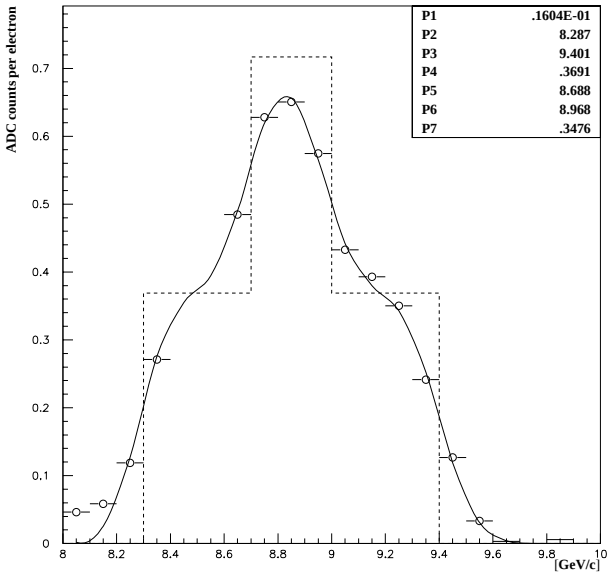

Figure 4.2: Acceptance of N3MO. Data points come from test beam momentum scans. Dashed line shows the two-box acceptance function, solid line the fit assuming a Gaussian shape to the test beam. The parameters of the fit are identified in Table 4.1.

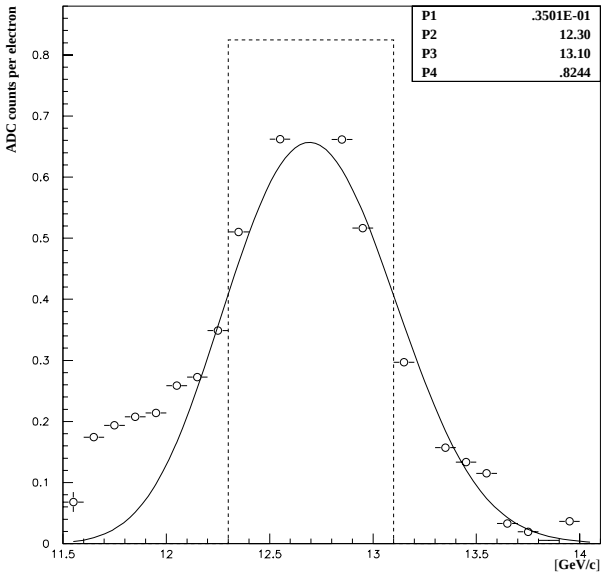

Figure 4.3: Same data as in Fig. 4.1, with the square acceptance used for the most recent round of n-tuple generation. The normalization for this figure was chosen by fitting for minimum chi-squared.

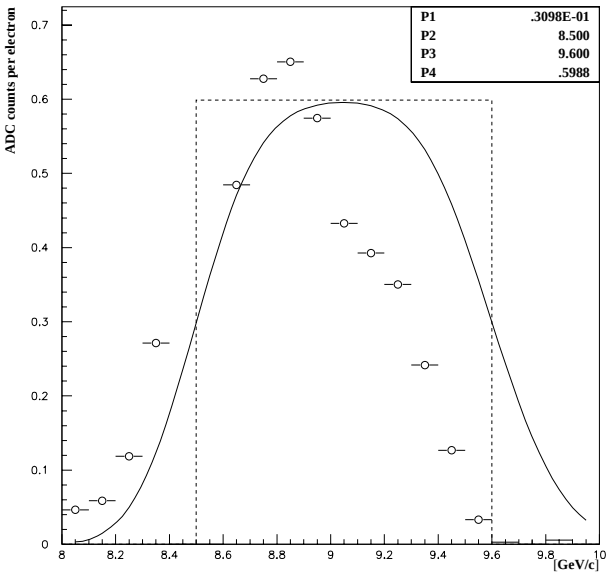

Figure 4.4: Same data as in Fig. 4.2, with the square acceptance used for the most recent round of n-tuple generation. The normalization was chosen by fitting for minimum chi-squared, with normalization chosen by fitting with upper edge at 9.3 GeV/c. (See footnotes to Table 4.3.)

Table 4.1: Description of the parameters of the fits in Figs. 4.1 through 4.4. Parameters 5-7 do not appear in the single-box fits.

parameter number	parameter description	units
1	fractional momentum spread (rms/mean)	-
2	low edge, box 1	GeV/c
3	high edge, box 1	GeV/c
4	constant, box 1	ADC counts per electron
5	low edge, box 2	GeV/c
6	high edge, box 2	GeV/c
7	constant, box 2	ADC counts per electron

Table 4.2: Values of the fitted parameters for the two-box fits to the test-beam acceptance scan data.

detector	low edge 1 [GeV/c]	high edge 1 [GeV/c]	const. 1[a] [ADC/e-]	low edge 2 [GeV/c]	high edge 2 [GeV/c]	const. 2 [ADC/e-]
N2MO	11.5	13.5	0.21	12.3	13.0	0.50
N3MO	8.30	9.40	0.37	8.70	9.00	0.35

[a] The constants are from the fits to the test beam data; they are not the calibration constant for the August 1996 data as such, although similar data was combined with other knowledge to obtain the original calibration constants[31].

Table 4.3: Values of the standard parameters for the single box acceptance of N2MO and N3MO.

detector	low edge [GeV/c]	high edge [GeV/c]	mean momentum [GeV/c]	width [GeV/c]	2-box mean mom.[a] and eff. width[b] [GeV/c]	[GeV/c]
N2MO	12.3	13.1	12.7	0.8	12.6	1.1
N3MO	8.50	9.60[c]	9.05[d]	1.10[e]	8.85	0.70

[a] The "mean momentum" is the mean value of the momentum p weighted by the acceptance function.
[b] The "effective width" is the total area under the acceptance function divided by the peak value of the (unsmeared) acceptance function.
[c]N.B. Some early versions of the `n2n3_par.lut` file have 9.3 GeV/c for the high edge of N3MO's acceptance, which would give a better match to the data and also give a width closer to the two-box "effective width." The parameter that was used for the last n-tuple generation is 9.6 GeV/c as quoted. The effect of different acceptances is discussed in a later subsection.
[d]Central momentum is 8.80 GeV/c if high edge is 9.3 GeV/c.
[e]Width is 0.8 GeV/c if high edge is 9.3 GeV/c.

The runs used in this analysis are wire scans 15140 and 15142, taken August 3, 1996, and PCAL calibration data runs 15530 through 15537, taken August 8, 1996. The ECAL position in the latter set of runs was -120 mm, so the top of the top row was at -245 mm in E144's "PCS" coordinate system. According to the spectrometer map, the top row intercepted electrons from 7.62 to 8.49 GeV/c, so it did not quite overlap the core of the N3MO acceptance. Because no ECAL row was directly behind the counters we wish to cross-calibrate, we have no direct measurement of the number of electrons passing through the counters in these runs. However, the average flux can be inferred and compared to the average ADC reading. In wire scans 15140 and 15142, the ECAL was behind N3MO but still below N2MO.

In this analysis, the top four rows of the ECAL are used to measure the total rate. Beam charge data from the Beam Containment System toroids was not available for these runs, so the fifth row of ECAL is used as a somewhat-independent quality cut to eliminate beam pulses with low rate. The total energy deposited in each row is divided by the mean energy of individual electrons accepted by that row and the width of the momentum acceptance for the row to obtain an estimate of the differential scattering rate at each momentum. This data is plotted in Figs. 4.5 and 4.7 (for the August 8 and August 3 data, respectively) and fitted to a bremsstrahlung spectrum. The fitted bremsstrahlung spectrum is extrapolated to the central momenta of the two counters, N2MO and N3MO. From this data, the average number of electrons per GeV/c is known for each detector for this data set.

Using the same events from the same data set, the pedestal-subtracted ADC readings from the two counters are plotted in Fig. 4.6 and 4.8. (Pedestals were measured using events for which the ECAL read little or no deposited energy.) From this data, the average ADC reading of each detector is known for the corresponding average number of electrons per GeV/c measured with the ECAL data. Dividing one by the other gives the calibration in terms of electrons per GeV/c per ADC count. This data is summarized in Table 4.4.

There is a factor of two discrepancy between the two calibrations. This cannot be due to the position of the calorimeter, as the effect of ECAL position on N3MO operation, if any, is very small, as shown in Fig. 4.9, and the vacuum chamber window extends below 6.0 GeV/c. The difference

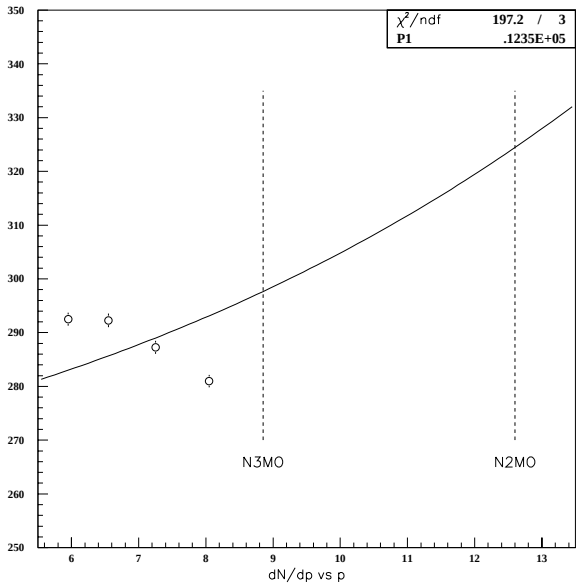

Figure 4.5: Bremsstrahlung spectrum as measured by the ECAL, from PCAL wire data taken August 8, 1996. Each data point represents one row of ECAL: the horizontal axis is the row central momentum, and the vertical axis gives average number of electrons per GeV/c in the row's momentum acceptance for this run. (The data is not normalized by beam charge or foil thickness, but can be compared to the "raw" data in Fig. 4.6 in order to obtain the calibration constant.) The error bars on the points give the root-mean-square of the data divided by the square root of the number of events. The solid curve is a bremsstrahlung spectrum with best-fit normalization. Central momenta of N2MO and N3MO are indicated with dashed lines. The deviation of the data points from the fit exceeds the statistical spread of the data. These residuals are all less than 6% of the signal and may be due to systematic effects in ECAL reconstruction or calibration.

Table 4.4: Results of the indirect calibration of N2MO and N3MO with ECAL using bremsstrahlung. The quoted errors represent statistical errors only.

detector	inferred flux [e-/(GeV/c)]	average reading [ADC counts]	calibration constant [e-/(GeV/c)/count]
N2MO, Aug. 3	204 ± 4	$12.4 \pm .7$	16.5 ± 1.0
N3MO, Aug. 3	187 ± 4	257 ± 12	$0.73 \pm .04$
N2MO, Aug. 8	325 ± 1	$11.8 \pm .1$	27.5 ± 0.3
N3MO, Aug. 8	298 ± 1	219 ± 2	$1.36 \pm .01$

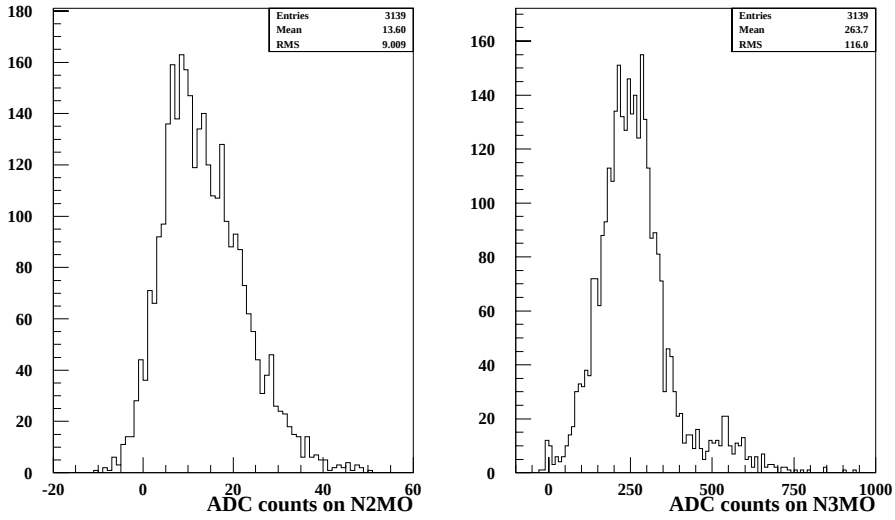

Figure 4.6: Frequency distribution of ADC readings, after pedestal correction, for runs 15530-15537 (August 8, 1996 data). N2MO is on the left, N3MO on the right.

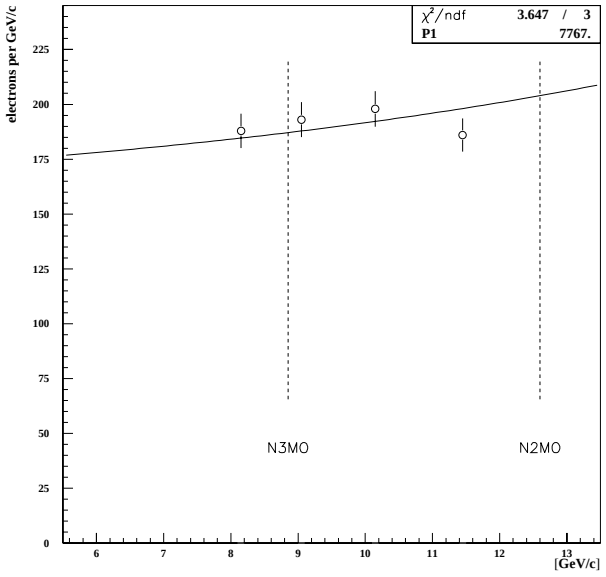

Figure 4.7: Bremsstrahlung spectrum as measured by the ECAL for wire scans 15140 and 15142 (August 3, 1996).

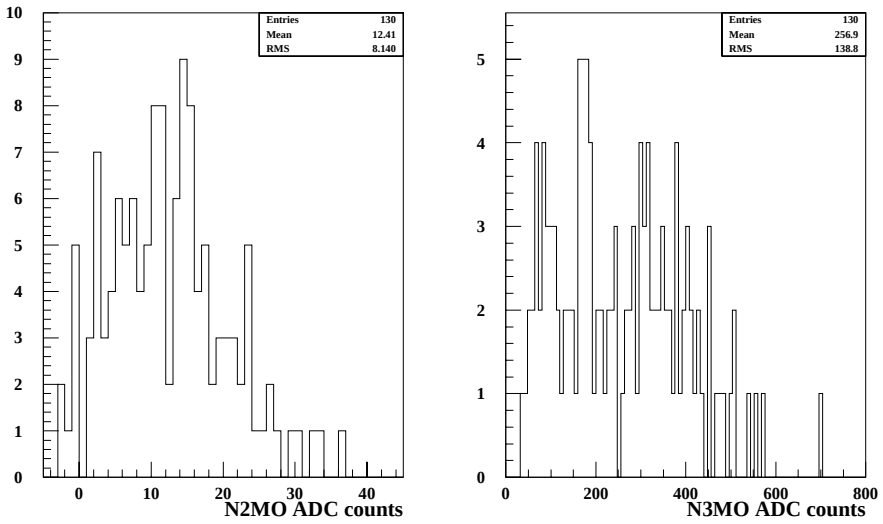

Figure 4.8: Frequency distribution of ADC readings for wire scans 15140 and 15142.

in calibration constants could represent a real change in N3MO calibration between August 3 and August 8. It could also be an effect of the unusual steering in place on August 8, although the horizontal "bump" of 1 mm should have been negligible for these detectors.

It will be noted that N2MO reads very low in these conditions, a result of it being configured for its normal mode of operation where it sees an order of magnitude more flux than N3MO. Extrapolating from low ADC counts should not be too much of a problem, however, since N2MO and N3MO were read out with a LeCroy 2249 ADC, which is much more linear at its low end than the Phillips QDC used for the $n = 1$ rate monitors. The residuals on the ECAL fit suggest systematic uncertainties of at least $\pm 6\%$ may be present.

As a further aid to visualizing the cross-calibration data, two-dimensional density plots of N2MO and N3MO readings versus the top row of ECAL are shown in Fig. 4.10, with lines drawn representing the major axis expected from the calibration constants given above. The scatter in the data arises from the statistics of the electron to photo-multiplier-tube photo-electron conversion and the fact that the top row intercepts different electrons than the detectors, but it can be clearly seen that the cross-calibration is correct.

The calibration constants found above will be compared with the "original" calibration in a later section.

4.2.3 The direct cross-calibration of N3MO with ECAL using nonlinear Compton data

A direct cross-calibration can be done with any run for which part of the ECAL was directly behind the monitor of interest. Unfortunately, there is no data for which the ECAL was directly behind N2MO. (The ECAL mover would have to have been set to above -60 mm, and the highest position

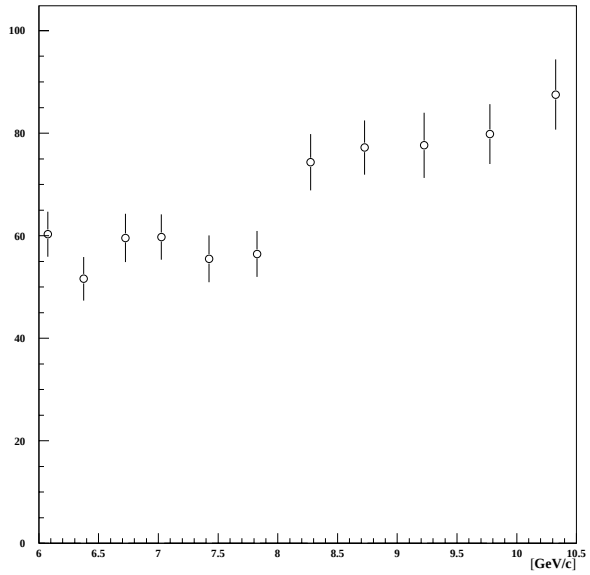

Figure 4.9: N3MO ADC normalized by linear detector, plotted versus central momentum of the top ECAL row, for ECAL scan 15296. A slight (less than 10%) shift in average reading may occur as the top of the ECAL passes behind N3MO, but this could also be a coincidental change in collision conditions.

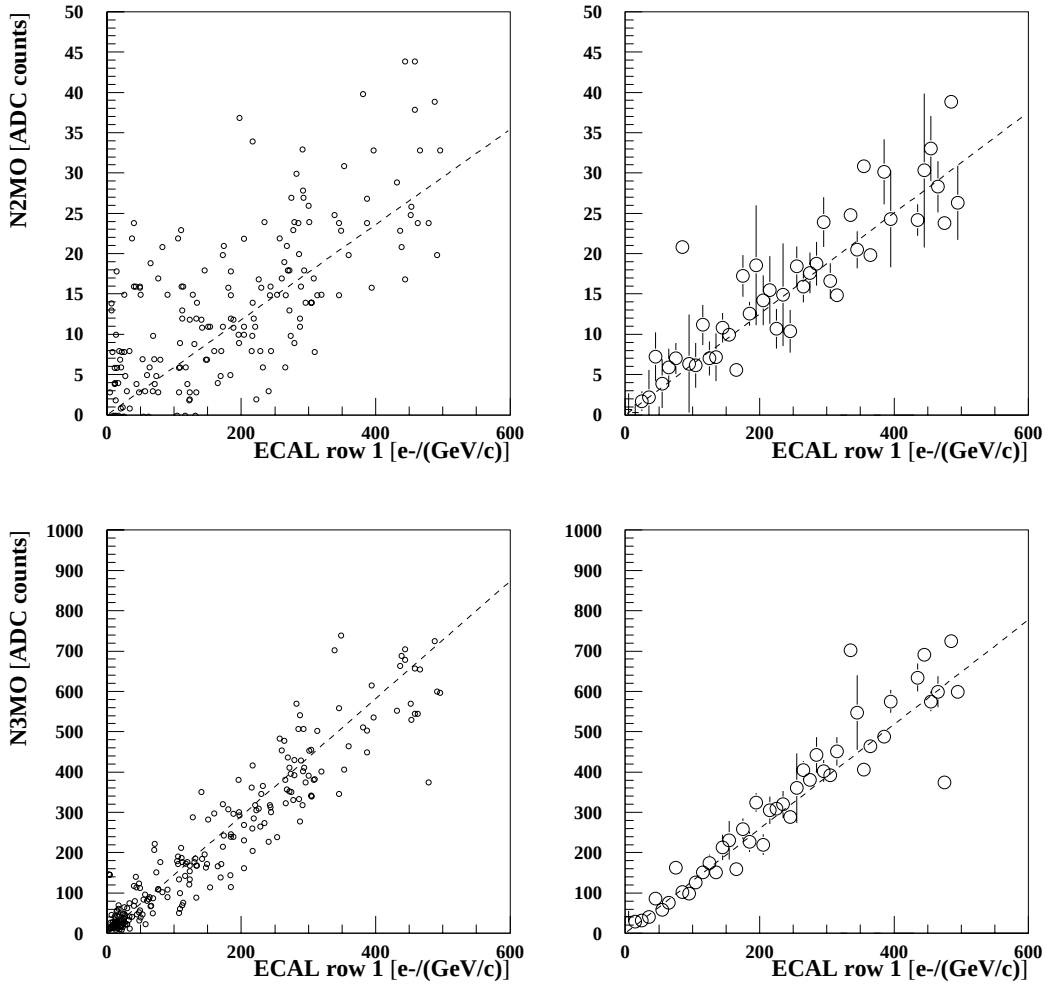

Figure 4.10: Two-dimensional scatter plots (left) and profile histograms (right) of ADC readings versus energy deposited in the top row of ECAL. The lines are calculated based on the calibration constants in Table 4.4, the acceptance of the top row, and the shape of the bremsstrahlung spectrum. Data is from the wire scans 15140 and 15142.

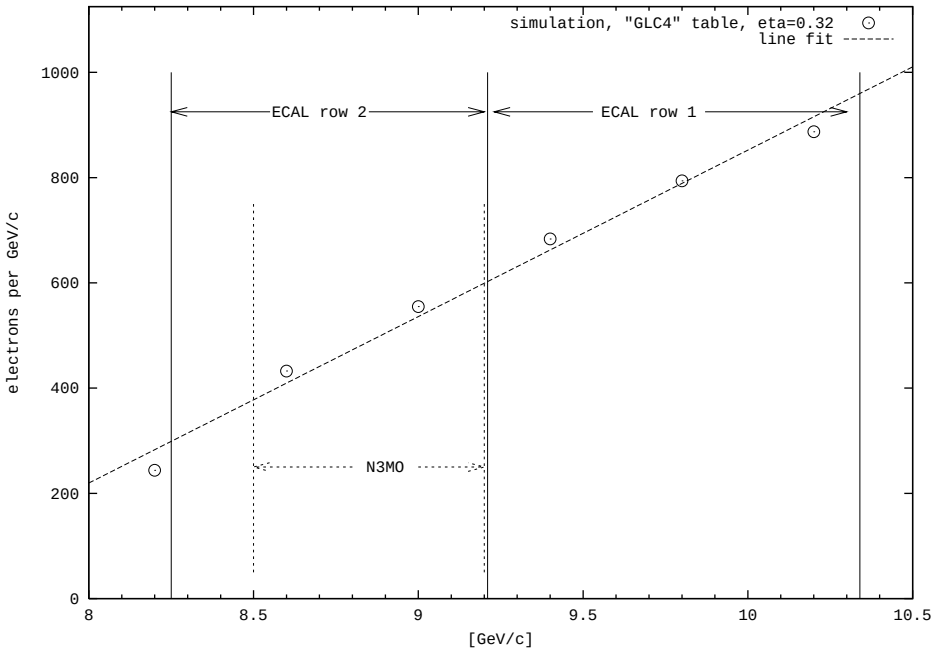

Figure 4.11: Simulated spectrum for the linear-green laser with edges indicated for N3MO and ECAL at -92 mm (top of top row at -217 mm in "PCS"). Parameters were $\eta = 0.3162$, $A_{\rm eff} = 29.9\ \mu^2$, $\tau = 1.5$ psec (FWHM), $E_{e-} = 46.6$ GeV, $\sigma_{e-,x} = 25\ \mu$, $\sigma_{e-,y} = 40\ \mu$, $\tau_{e-} = 7$ psec (FWHM).

recorded was -70 mm.) However, there is plenty of data to provide a direct N3MO calibration.

The analysis below uses data from ECAL scan 15296, which is an actual, positron-producing run (arguably our first and best), with N2MO and N3MO in their final data-taking configurations. This data was taken on August 5, 1996. Events were required to have at least 10^6 forward photons as detected by EC31/EC37. The spectrum varies nearly linearly with momentum in the range of interest, both in theory (Fig. 4.11) and in observation (Fig. 4.12). The number of electrons per GeV/c averaged over the acceptance of the counter is therefore the same as the number of electrons per GeV/c at the center of the counter's acceptance, which can be determined by interpolating between the two nearest calorimeter rows.

Figure 4.13 shows the N3MO ADC plotted versus the ECAL estimate of $dN_{e-,\texttt{N3MO}}/dp$, for events when at least one row of ECAL was above and one below the N3MO central momentum of 8.85 GeV/c. The inverse of the slope is $0.65 \pm .01$ electrons per GeV/c per ADC count, where the error comes from the fit covariance matrix. Significant scatter of the data is evident. Figure 4.14 shows the electron yield in the ECAL in the $n = 3$ region normalized by the N3MO ADC. The intercept of the fitted line with 8.85 GeV/c is at $0.62 \pm .01$ GeV/c per ADC count, where again the error quoted is from the canonical fit error estimate. Folding in likely ECAL reconstruction systematic errors, the cross-calibration using laser-scattered electrons with ECAL directly behind N3MO gives a calibration constant of 0.63 ± 0.01 stat. $\pm$ .04 syst., which is consistent with the August 3 cross-calibration.

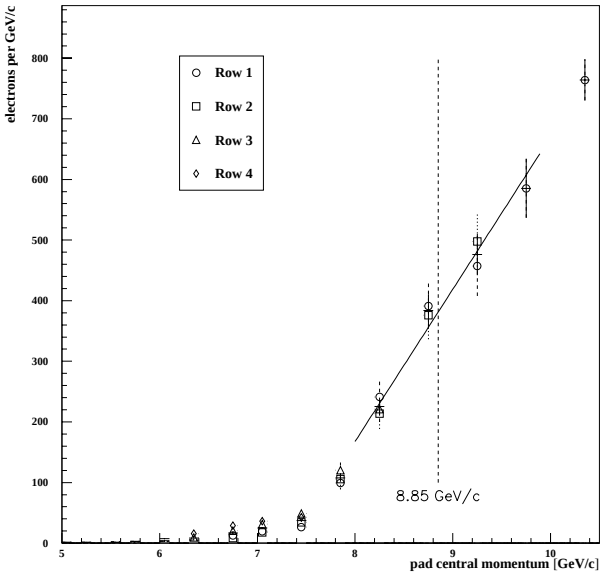

Figure 4.12: Observed spectrum from run 15296. All reconstructed rows give the same answer in the region of interest.

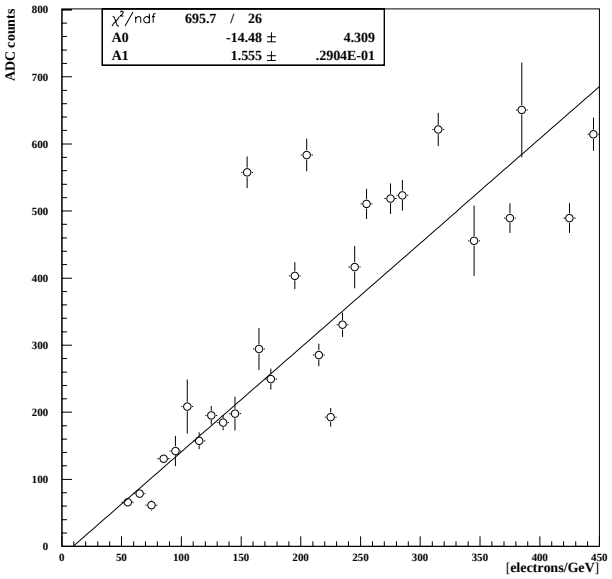

Figure 4.13: N3MO ADC averaged and plotted in bins of ECAL-based estimate of number electrons per GeV/c at 8.85 GeV/c.

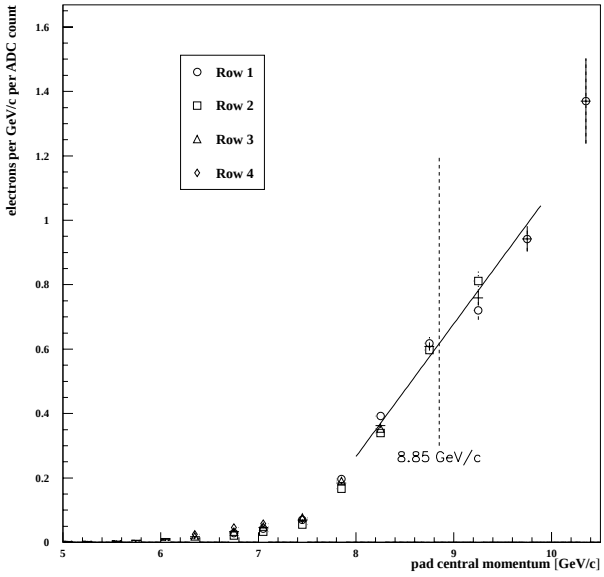

Figure 4.14: Observed spectrum from run 15296 normalized by N3MO ADC. The intercept of the fitted line with the N3MO central momentum gives the calibration constant.

4.2.4 Effect of different assumed acceptances of N2MO and N3MO

The calibration constants calculated above convert from ADC counts to electron yield per unit momentum at the central momentum of the detector's acceptance. Any systematic errors in this conversion due to the finite width of the momentum acceptance must come from second-order and higher derivatives of the spectrum w.r.t. momentum, and since both the bremsstrahlung and nonlinear Compton spectra are very linear with momentum in the ranges of interest, this error is small. Figures 4.15 and 4.16 show the effects of varying momentum acceptances on N2MO and N3MO, based on simulated spectra.

Systematic calibration errors could also come from misidentification of the detector central momentum. Figures 4.17 and 4.18 show the fractional error introduced by central momentum misidentification, obtained by shifting the acceptance functions described in Table 4.2 and applying them to simulated spectra. Again, the possible systematic errors are small.

4.2.5 Calibration of EC31/EC37 with bremsstrahlung

Calibration runs 15460 - 15463, taken August 7, 1996, provide good data for calibrating the linear monitors. This data was taken with an electron beam of 7.1×10^9 e-/pulse incident on the 30-μm-thick aluminum foil at IP1. N2MO readings for this run are consistent with these conditions, assuming the 0.7 e-/(GeV/c)/(ADC count) calibration is correct. (N3MO apparently had its filter removed for this run.) The EC31/10 ADC averaged 3317 counts for this data, and the EC37 ADC averaged 1508 counts. The electron flux from the 46.6-GeV electrons incident on the foil can be calculated as 1.64×10^5 e-/(GeV/c) and 2.92×10^5 e-/(GeV/c) at 31 GeV/c and 37 GeV/c, respectively. The Compton spectrum for the green laser gives values of $dN/dp/N_\gamma$ of 0.0276 e-

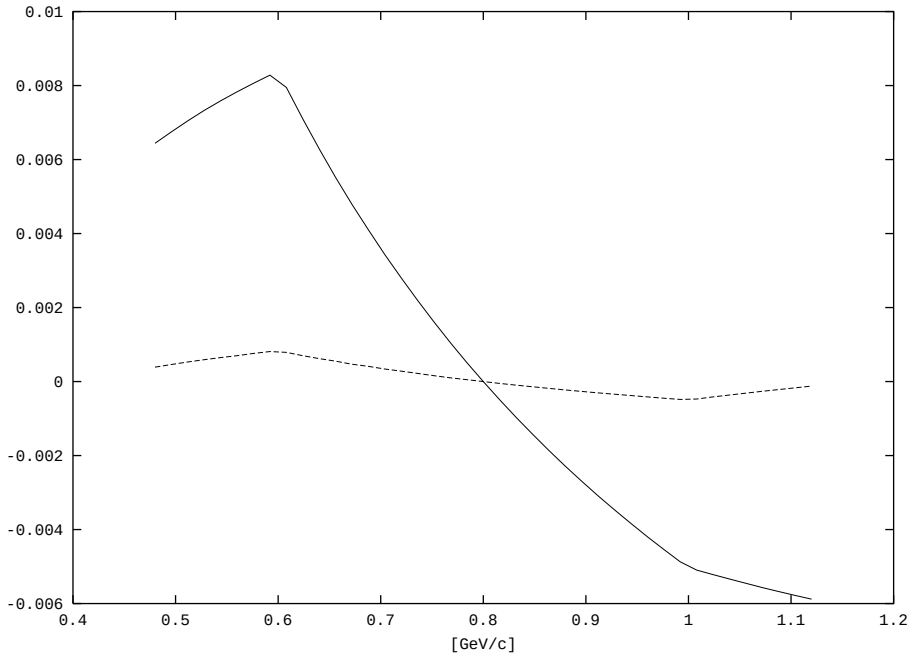

Figure 4.15: Simulated effect of different acceptance widths on the accuracy of the calibration of N2MO. The horizontal axis is the (effective) width of the acceptance function, and the vertical axis is the fractional deviation of the estimated electron yield from its actual value at the central momentum of the detector, as calculated by the `sum_bin` function from the MSDOS-based E-144 analysis software. The solid line shows the simulation for a nonlinear Compton spectrum and a square aperture of the given width; the dashed line shows the effect with square aperture and a bremsstrahlung spectrum.

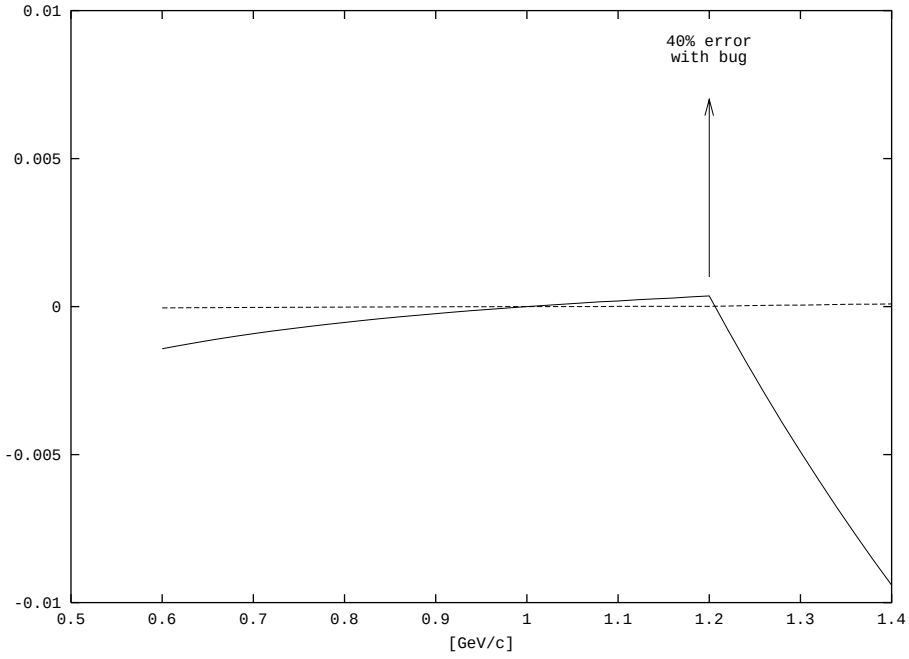

Figure 4.16: Simulated effect of different acceptance widths on the accuracy of the calibration of N3MO. (See caption for Fig. 4.15.) A bug in some early versions of the `sum_bin` function caused an extra bin to be added when the upper end of the acceptance fell exactly on a bin edge.

/photon/(GeV/c) at 31 GeV/c and 0.0320 e-/photon/(GeV/c) at 37 GeV/c. The calibration constants are therefore 179 photons/(ADC count) for EC31 (*i.e.,* 1790 photons per count of the EC31/10 ADC) and 6045 photons/(ADC count) for EC37. Systematic errors may be of order $\pm 10\%$ due to beam charge toroid calibration, accuracy of foil thickness information, and/or N2MO cross-calibration accuracy.

4.2.6 Calibration constants used in the last round of n-tuple generation

In the most recent round of n-tuple generation, the N2MO and N3MO ADC readings were converted to total electron counts by multiplying by constants 49 e-/count and 1.8 e-/count, respectively. The corresponding simulation results for these counters were found by integrating the simulated spectrum over the ranges indicated in Table 4.3, which have widths 0.8 GeV/c and 1.10 GeV/c, respectively. Expressed in the terms of the analysis in this paper, the last round of n-tuple analysis compared the simulated yield of electrons per GeV/c averaged over the given momentum range versus the ADC readings times the constants 61 e-/(GeV/c)/count (for N2MO) and 1.6 e-/(GeV/c)/count (for N3MO). These constants are different from those found here by cross-calibration with ECAL, but not by a factor of eight.

The calibration constants for EC31 and EC37 were equivalent to 289 photons/(ADC count) and 8874 photons/(ADC count), respectively.

The "old" calibration constants and those found in the subsections above are summarized in Table 4.5, and further discussed in the next section.

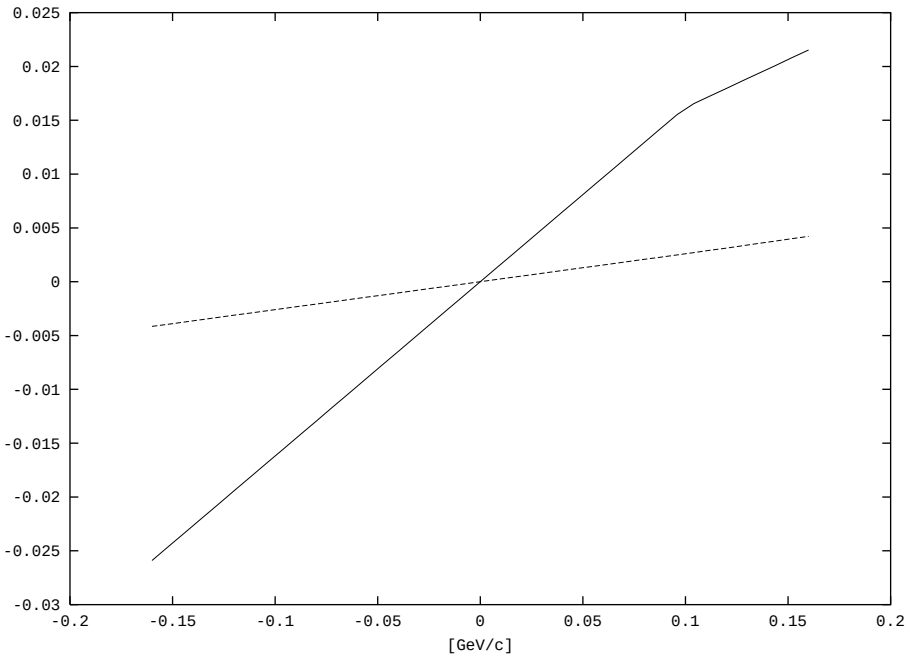

Figure 4.17: Simulated effect of different central momenta on the accuracy of the calibration of N2MO. Horizontal axis is the shift in GeV/c relative to the nominal central momentum of 12.6 GeV/c. Vertical axis is fractional deviation of the detector yield estimate from its actual value at the nominal central momentum. Solid line is nonlinear Compton data, dashed line is bremsstrahlung.

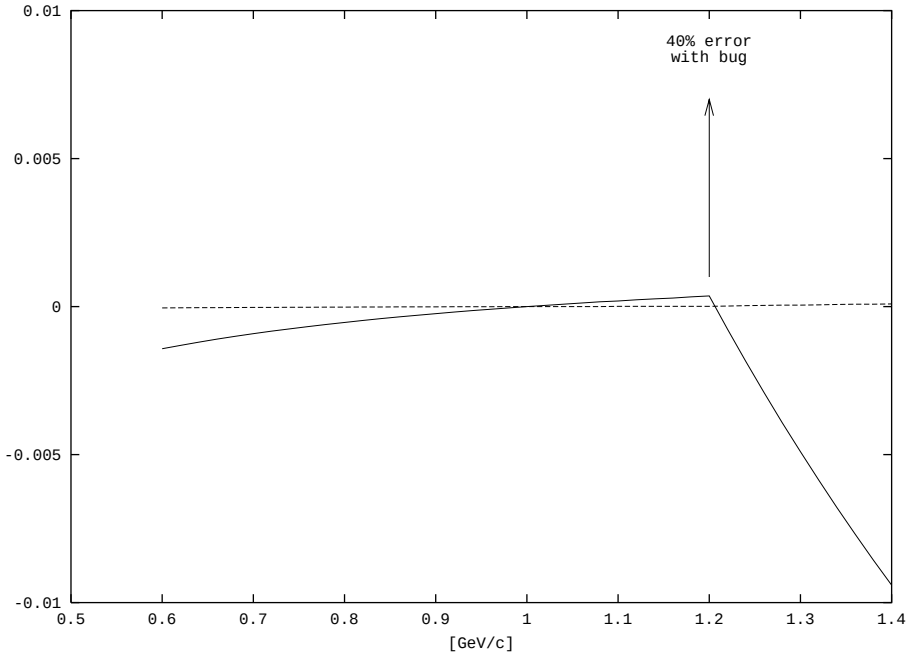

Figure 4.18: Simulated effect of different central momenta on the accuracy of the calibration of N3MO. (See caption for Fig. 4.17.) A bug in some early versions of the `sum_bin` function caused an extra bin to be added when the upper end of the acceptance fell exactly on a bin edge.

4.3 Conclusions

The results of the cross-calibration described above are presented in Table 4.5 along with the equivalent constants from the original calibration. Table 4.6 lists the scaling factors used in the original Cherenkov-monitor-based constrained fits [6] and the equivalent scaling accomplished by using the new calibration constants. Although the cross-calibration does not give the same constants as the original calibration, the disagreement is much smaller than that required to explain the factor of eight rescaling done for the constrained fits. According to the cross-calibration, the N2MO and N3MO detectors readings using the old calibration should be scaled by no more than 3.6 and 2.5, respectively.

Furthermore, no adjustment of these newly-obtained calibration constants is required in order to allow good fits of Cherenkov monitors with the simulation. Using the calibration constants of 17 e-/(GeV/c)/count for N2MO and 0.65 e-/(GeV/c)/count for N3MO results in agreement as good as that obtained by large scaling of the original calibration, as detailed in Table 4.6. Use of these new calibration constants corresponds to scaling the original calibration of N2MO by 3.6 and scaling N3MO by 2.5. The average fitted value of η is $\sqrt{1.8}$ times higher when using the new "adjusted" calibration, a result of the effective scaling with respect to the original differing by a factor of approximately $(1.8)^n$ in each detector. It is interesting that both previous and "new" average η values are less than that calculated from the "standard" laser area, pulse length, and measured energy.

To demonstrate the internal consistency of the new calibration and associated η fits, in Fig. 4.19 the reconstructed signal (Horton-Smith method) in the top row of ECAL is plotted versus the

Table 4.5: Comparison of original calibration with cross-calibration results. Quoted errors include best-guess systematic errors.

	calibration constant [e-/(GeV/c)/count]				
detector	original	bremss., Aug. 3	Compton, Aug. 5	bremss., Aug. 8	adjusted for fit
N2MO	61	17 ± 2	—	28 ± 2	15
N3MO	1.6	0.7 ± 0.1	$0.64 \pm .04$	1.4 ± 0.1	0.71

	calibration constant [photons/count]				
detector	original	—	—	bremss., Aug. 7	adjusted for fit
EC31	289			179 ± 20	185
EC37	8874			6045 ± 600	5900

Table 4.6: Comparison of effect of the previously-used scaling with the effect of using the new "adjusted" calibration constants in Table 4.5, based on analysis of runs 15296 and 15323. Note that the average value of η calculated from "standard" laser area, pulse length, and measured energy is 0.37.

	scaling of readouts w.r.t. original				
	NGAMMA	N2MO	N3MO	fit $\langle\chi^2\rangle$	$\langle\eta_{\rm fit}\rangle$
previous scaling	1.6	8	8	0.9	0.23
new calibration	1.5	4.0	2.3	0.9	0.31
ratio	1.1	2.0	3.5	—	$\sqrt{1.8}$
$0.6 \times (1.8)^n =$	1.1	1.9	3.5		

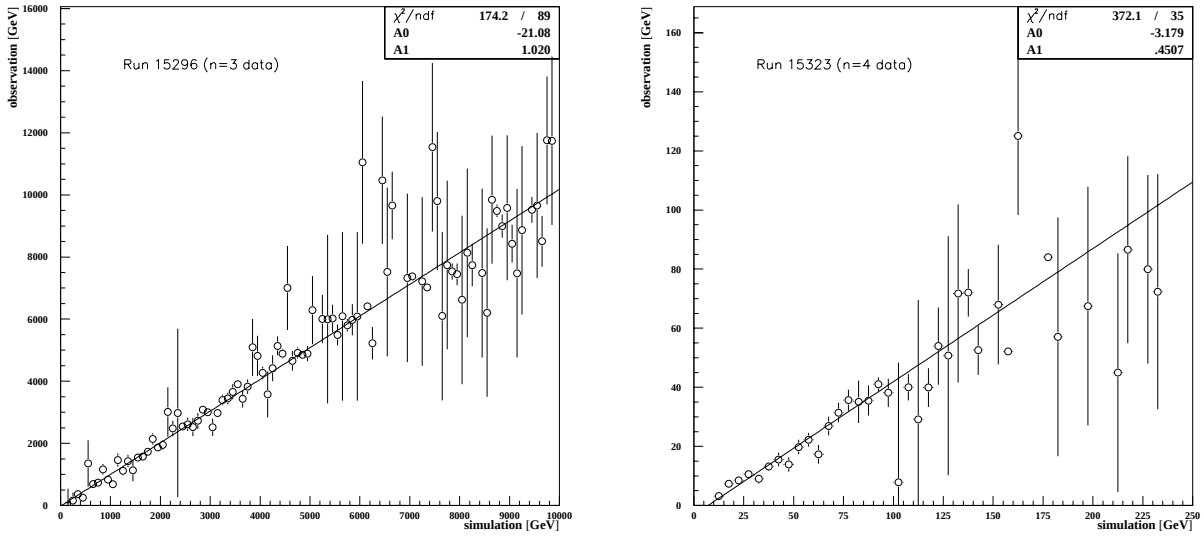

Figure 4.19: Comparison of experiment with simulation calculated for Cherenkov-monitor-derived η, with new calibration, for two sample runs: 15296 on left, and 15323 on right.

simulation prediction for runs 15296 and 15323, using all data with only a single cut: $0 < \chi^2_{\eta\text{-fit}} < 10$. Agreement should be almost guaranteed for run 15296, where ECAL is in the $n = 3$ position: the η and overlap are chosen to reproduce the observed signal in N3MO, which has been cross-calibrated with ECAL. In actuality, the observed signal is the same as predicted, demonstrating internal consistency.

The top half of the top row of the ECAL is above the $n = 4$ edge for run 15323. In this case, the observed signal is about half of the simulation. This would not have been considered a "disagreement" in the past, since our intensity estimates were uncertain by $\pm 30\%$, giving a factor of 2.5 uncertainty in normalized $n = 4$ rate and even greater uncertainty in absolute rate. The new calibration reduces systematic uncertainties to the point where this deviation from theory deserves a closer look.

Curiously, the laser intensities from constrained fits with the previous scaled calibration were also about a factor of two lower than the new estimates, which would precisely match the relative rates of the ECAL observations in the two orders, though not the absolute rates.

Shmakov [49] has previously compared the observed ECAL data to spectra simulated for laser intensities obtained from the Cherenkov-monitor η fits using the previous scaling. He has found the observation and simulation to be in good agreement, even for cases when the ECAL is directly behind N3MO. Naively, one might not expect this given the new calibration data. If the estimated electron yield from the unscaled N3MO is only a factor of 2.3 too high, and an η is chosen that gives a simulated N3MO reading lower than the actual N3MO reading by a factor of 8, one would expect the simulated ECAL reading of rows in the $n = 3$ region to be lower than observation by about factor of 3.5. This would follow directly from the cross-calibration of N3MO with ECAL using actual $n = 3$ non-linear Compton data. However, since Shmakov compares ECAL data normalized by *unscaled* number of forward photons (using the old calibration) to simulated ECAL data normalized by the simulated number of photons, the expected disagreement between simulation and theory is only $3.5/1.5 \simeq 2$. This is still larger than seen in his analysis. It is just possible that there may have been a compensating factor of two error in the calculation of the simulated N2MO and N3MO signals, or there may be some other subtlety of the analysis that is not obvious upon naive consideration. As mentioned above, any effect bringing the $n = 3$ yield into agreement would also result in excellent agreement for the $n = 4$ agreement due to the lower η estimates in the "scaled" analysis.

In conclusion, the good news is that the cross-calibration apparently eliminates the need for *ad hoc* scaling factors and therefore reduces systematic uncertainties, so that previously-uninteresting inconsistencies between data and theory now become potentially interesting. Additionally, the increased confidence these cross-calibration results provide for the Cherenkov monitors and associated η information is extremely welcome in light of the CCD data, which shows an unexpected relationship between the $\eta_{\text{CCD-fit}}$ and $\eta_{\text{Chernkov-fit}}$. These issues will be addressed in a later paper.

Chapter 5

Estimation of laser focal properties using Cherenkov monitor data

Synopsis

A method is described for finding E-144 interaction geometry parameters consistent with event-by-event readings of the Cherenkov monitors. Information from the laser pulse energy monitor may optionally be incorporated. This method should be essentially equivalent to that developed by Koffas and Bula [6], but without any iterative adjustment of scaling factors. The algorithm for implementing the method is fully described, and results of its application to real data are presented and discussed.

5.1 Introduction

The E-144 experiment is unusual in that the physics result depends strongly on the geometry of the interaction. From past observation, it is known that the laser focal area and pulse length can vary widely event-to-event. In our most recent and most successful data-taking run, event-by-event measurements of the laser spot-size and pulse length were not available. The laser pulse energy and electron pulse charge were measured event-by-event, while the electron transverse sizes were measured periodically and the electron bunch length was inferred from the ring-to-linac compressor settings and data obtained by Holtzapple[21]. The laser is believed to have had a focal area of roughly $30 \times 3^{\pm 1}$ μm^2 and a pulse length "typically" about 1.5 psec but possibly as low as 1 psec on "good" pulses or much larger on "bad" pulses.

Since it was not possible to measure all parameters of the interaction geometry directly on an event-by-event basis, we test the accepted theory by finding a geometry consistent with a few selected detectors assuming the standard theory, and then look at the remaining detectors to see if they also are consistent with the accepted theory. The detectors we choose are the $n = 1$, $n = 2$, and $n = 3$ fixed-momentum Cherenkov detectors. In addition, the laser pulse energy measurement from the joule pyrometer may be regarded as a sort of "$n = 0$" measurement; it is useful to regard the laser energy data as an observation to be fit rather than a fixed input parameter because the measurement uncertainty in the energy measurement is comparable to that in the Cherenkov monitors. Having a limited number of measurements, only a limit number of unknowns can be found. The unknowns of choice are timing offset ("overlap", Δt), focal intensity (or η^2), and focal area (A). These three parameters are sufficient to fit the four detector readouts with $\langle \chi^2 \rangle < 1$.

The final section gives further details of the method of finding the three unknowns $(\Delta t, \eta^2, A)$ that best fit the four measurements ($n = 1, 2, 3$ and laser energy). The resulting theory predictions are guaranteed to match the observations in these detectors, assuming the χ^2 is low. The question of the agreement of other detectors with the theory will be addressed elsewhere.

5.2 Review of simulation lookup tables

The simulation data in the standard data summary n-tuples is obtained using "lookup tables" computed by Bula's NUMINT program, described fully in his documentation[26]. What is most important to know here is that the predicted rate for any given final state is approximated as a product of functions of the different interaction parameters:

$$N_i(\eta, A, \tau, \sigma_x, \sigma_y, \tau_e, \Delta t) \cong N_i|_\eta(\eta) R_i|_A(A) R_i|_\tau(\tau) R_i|_{\sigma_x}(\sigma_x) R_i|_{\sigma_y}(\sigma_y) R_i|_{\tau_e}(\tau_e) R_i|_{\Delta t}(\Delta t), \qquad (5.1)$$

where

$$\begin{aligned} N_i|_\eta(\eta) &\equiv N_i(\eta, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0) \\ R_i|_A(A) &\equiv N_i(\eta_0, A, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0)/N_i(\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0) \\ R_i|_\tau(\tau) &\equiv N_i(\eta_0, A_0, \tau, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0)/N_i(\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0) \\ &\vdots \\ R_i|_{\Delta t}(\Delta t) &\equiv N_i(\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, \Delta t)/N_i(\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}, 0), \end{aligned} \qquad (5.2)$$

and $\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}$ are "standard" values of the interaction geometry parameters. By taking the logarithm of Eq. 5.1, one sees that the approximation is equivalent to a first-order Taylor expansion of N_i for small deviations of the parameters from their "standard" values. Clearly, the approximation is accurate for large deviations of the parameters if the function really does separate in variables as assumed.

In fact, this approximation is reasonably accurate in accounting for variations in electron beam sizes, overlap, and laser pulse length at fixed intensity and area, since these parameters affect the rates of all processes approximately linearly.[1] At fixed area, variations in rate with intensity are also well approximated since the rate for any given final state tends to be dominated by the process or processes with the lowest "n". The functional dependence of rate on area is perhaps the most complicated due to the changing dominance of multiple scattering versus single-scattering as area changes, and so any failure in the separation of variables in Eq. 5.1 is most likely to appear with area.

In order to minimize such errors, the "standard" parameters $\eta_0, A_0, \tau_0, \sigma_{x0}, \sigma_{y_0}, \tau_{e0}$ are chosen to be well-centered in the distribution of parameters actually required for simulation of the observed events. Eleven lookup table sets were produced at one time or another, each with their own standard parameters. Different table sets are given designations such as "GLC4" or "ICA1"; the origin of these codes is described in Bula's note[26]. The standard parameters for several lookup table sets are summarized in Table 5.1. An idea of the scale of errors introduced by approximation 5.1 can be obtained by looking up spectra in one lookup table set at another set's standard parameters, and comparing the estimate from one with the exact simulation results recorded in the other. Figure 5.1 shows the electron spectrum at the "GLC4" standard parameters as estimated from a different lookup table set.

[1] With fixed energy, laser pulse length has a nonlinear effect due to its effect on intensity, but at fixed intensity the effect is approximately linear in pulse length for all processes.

Table 5.1: Standard parameters for several lookup table sets.

table set	laser pol.	wavelength [μm]	η_0	A_0 [μm^2]	τ_0 [psec]	$e-$ energy [GeV]	σ_{x0} [μm]	σ_{y_0} [μm]	τ_{e0} [psec]
GCA1	circular	0.527	0.0976	35	2	46.6	60	60	6.8
GLA1	linear	0.527	0.0976	35	2	46.6	60	60	6.8
ICA1	circular	1.054	0.171	65	2	46.6	60	60	6.8
ILA1	linear	1.054	0.171	65	2	46.6	60	60	6.8
GLC4	linear	0.527	0.326	29.9	1.5	46.6	25	40	7.0
GLD1	linear	0.527	0.309	29.9	1.5	49.1	34	34	6.7

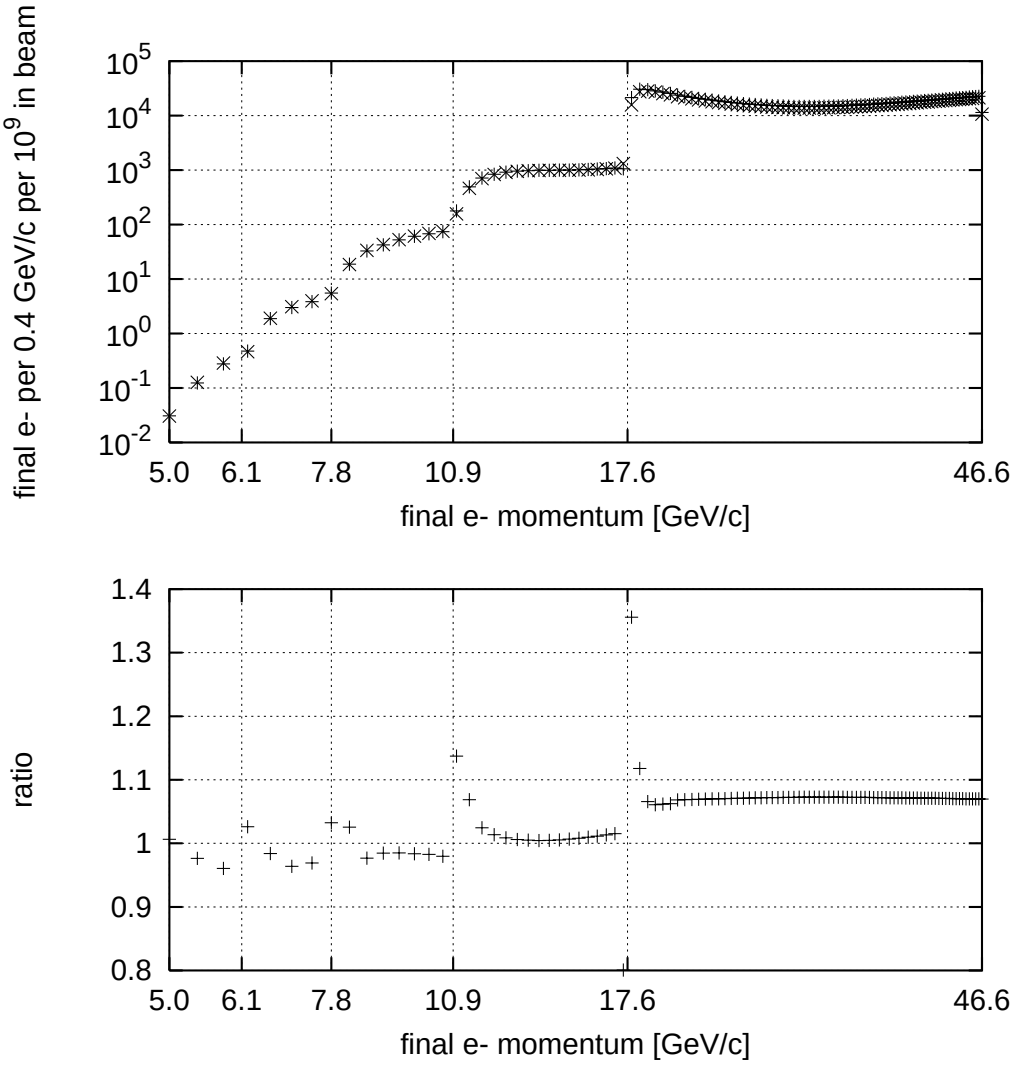

Figure 5.1: The top graph shows the electron spectrum at the "GLC4" standard parameters as estimated from "GLA1" (+) and "GLC4" (×) lookup table sets plotted with. The lower graph shows the ratio of the "GLA1" table extrapolation to the unextrapolated simulation results stored in the "GLC4" table.

Table 5.2: Naive expectation for power law coefficients of rates of various processes as a function of interaction geometry parameters. The quantity m is the number of scatterings in a successive-scattering process, and n is the sum of the number of incident photons absorbed at each scattering in the sequence.

	$n=1$	$n=2$		$n=3$			$n \lesssim 3$	$n \gtrsim 3$
	$m=1$	$m=1$	$m=2$	$m=1$	$m=2$	$m=3$		
η^2	1	2	2	3	3	3	n	n
A	1	1	1.5	1	1.5	2	$(m+1)/2$	$(m+2)/2$
τ	1	1	1	1	1	1	1	1

5.3 Power-law dependence of total scattering rates on interaction geometry

The dependence of scattering rates on intensity is the easiest to understand. If n is the sum of the number of incident photons absorbed, then for low or moderate intensities and fixed focal sizes the rate is proportional to intensity to the n-th power, or η^{2n}, to leading-order in η^2. The dependence of rates on geometry for fixed intensity is somewhat less intuitive.

Figures 5.2-5.3 show cut-away views of the geometry of the interacting beams for optimal temporal and spatial overlap, at the time of the intersection of the centroids of the two pulses. It can be seen that the electron sizes are all larger than the corresponding laser sizes. Because the laser pulse is shorter than the electron pulse and most scattering occurs far from any edge of the electron beam, all scattering rates are nearly proportional to the pulse length at fixed intensity.

The dependence on area is more complicated, and is best considered one dimension at a time. The vertical extent of the laser focus encountered by the electron is proportional to $\sqrt{A}$. The depth of the laser field encountered is proportional to $\sqrt{A}$. The horizontal extent of the contributing part of the focus is the most hard to intuit: for low-n scattering, say $n = 1$ or 2, the decrease in intensity due to moving off-axis horizontally is partially compensated for by the increase in the depth of field, so that the horizontal extent of the contributing part of the focus is determined mostly by the laser angular divergence for low n, and is therefore roughly independent of the focal area. However, for high n, the decrease in intensity with horizontal position is the primary factor limiting the horizontal extent of the contributing portion of the focus, so that the crossing-angle-foreshortened Rayleigh range of the laser determines the focal horizontal extent, which introduces another factor of $\sqrt{A}$ for large n. Thus, at fixed intensity and pulse length, the rate of scattering by m successive scatters with n photons absorbed is roughly proportional to $(\sqrt{A})^{m+1}$ for $n \lesssim 3$ and $(\sqrt{A})^{m+2}$ for $n \gtrsim 3$. These rough expectations are summarized in Table 5.2.

The above discussion is an attempt to provide some sort of intuitive basis for the variation of rate with focal parameters. Quantitative theoretical predictions are obtained from the NUMINT program. The lookup tables written by NUMINT contain information on the variation of rate as a function of each parameter. This information can be fitted to a power law. Figure 5.4 shows there results of such an analysis of the "GLC4" lookup table set, plotted versus momentum for each bin of the electron and photon spectrum. Table 5.3 gives the power law coefficients for several particularly interesting portions of the spectra: the linear portions of the electron and photon spectra, the $n = 2$ portion of the photon spectrum, and the portions of the electron spectrum corresponding to the $n = 2$ and $n = 3$ monitors. Note that our electron detectors can not tell how many times an electron was scattered in the laser focus, and neither is this information recorded in

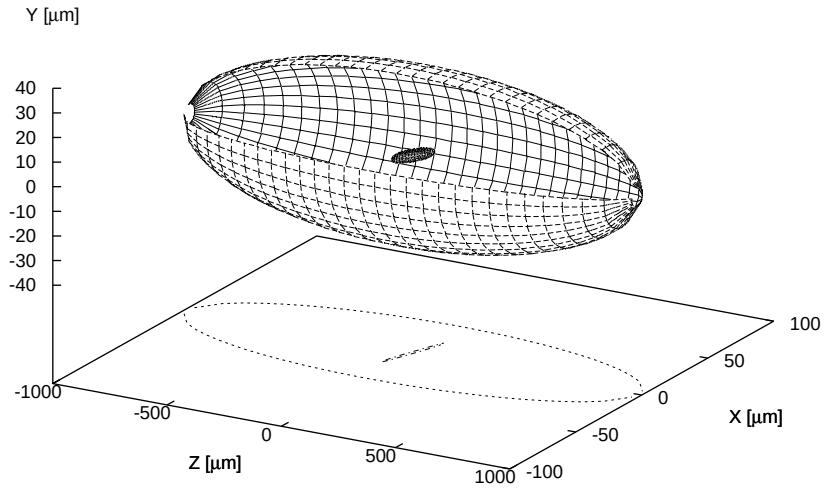

Figure 5.2: Three-dimensional, cut-away view of the geometry of the interacting beams showing the half-intensity surfaces of the electron beam and the focussed laser beam.

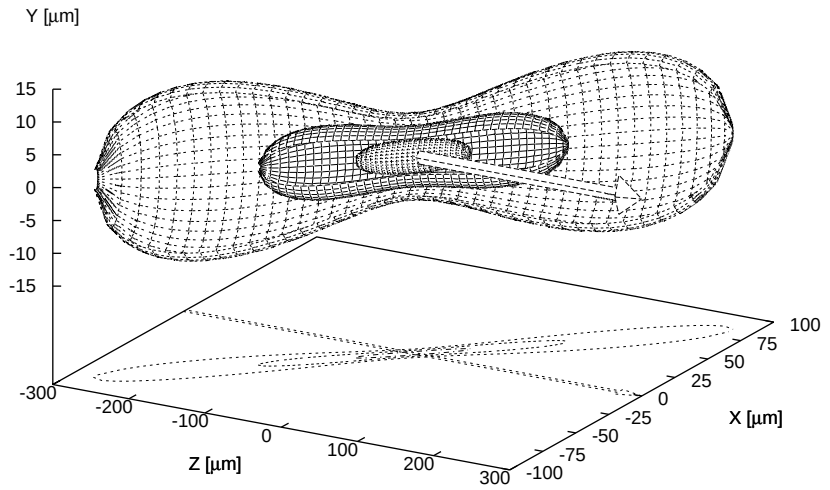

Figure 5.3: Enlarged cut-away view of the focussed laser, showing surfaces at 1%, 10%, and 50% of maximum intensity.

Table 5.3: Fitted power law coefficients from GLC4 lookup tables for various interesting regions of the electron and photon spectra. The intensity (η^2) is held constant as area (A) or pulse length (τ) are varied.

	electrons			photons		positrons
	19-46.6 GeV/c	12.2-13.2 GeV/c	8.5-9.5 GeV/c	0-28 GeV/c	30-34 GeV/c	(all)
η^2	1.0	2.0	2.9	1.0	2.0	5.1
A	1.0	1.5	1.7	1.0	1.3	2.0
τ	1.1	1.1	1.0	1.1	1.0	1.1

Bula's lookup tables, but the behavior of rate with focal area at fixed intensity gives an indication of what order m is dominant.

With the information above, one can calculate the change in the detectors as a parameter is varied. The deviation of the logarithm of the detector reading will be linearly related by the power law coefficient to the change the logarithm of the parameter. For example, if the area changes from A_1 to A_2, then the rates in the EC31, N2MO, and N3MO detectors will change by factors of $(A_2/A_1)^{1.0}$, $(A_2/A_1)^{1.5}$, and $(A_2/A_1)^{1.7}$, respectively.

5.4 Details of the algorithm

The algorithm consists of a golden section search for the minimum χ^2 fit to the observed monitor readings as η is varied. At each η, the overlap and focal area are chosen as follows:

1 Initialize area to default.

2 Initialize pulse length to default.

3 Look up rates for given $\eta, A, \tau, \sigma_x, \sigma_y, \tau_e$ and measured total number of $n = 1$ scatters. Note that Bula's standard lookup routines automatically find the Δt that yields the observed $n = 1$ rate, if possible.

4 If computed $\Delta t = 0$ and computed $n = 1$ rate is greater than observed, multiply τ by factor of $[(\text{observed } n = 1)/(\text{computed } n = 1)]^{0.95}$ and repeat step 3.

5 Compute logarithms of observed rates divided by calculated rates for each monitor, i.e., deviations of logarithms of observed rates from logarithms of calculated rates.

6 Calculate χ^2.

7 Perform linear fit of deviations to area power-law coefficients. If area correction is large, implement correction and reiterate from step 2. The focal area parameter is required to stay within the range of the lookup tables. Note if χ^2 increases after correction, the lowest χ^2 and corresponding area, offset, and pulse length are restored and iterations ceased. Iterations also cease after the number of loops exceeds `cfit_max_it`, which defaults to 2.

8 Return best χ^2 for this value of η to the golden section search procedure, which continues the golden section search until the minimum is found.

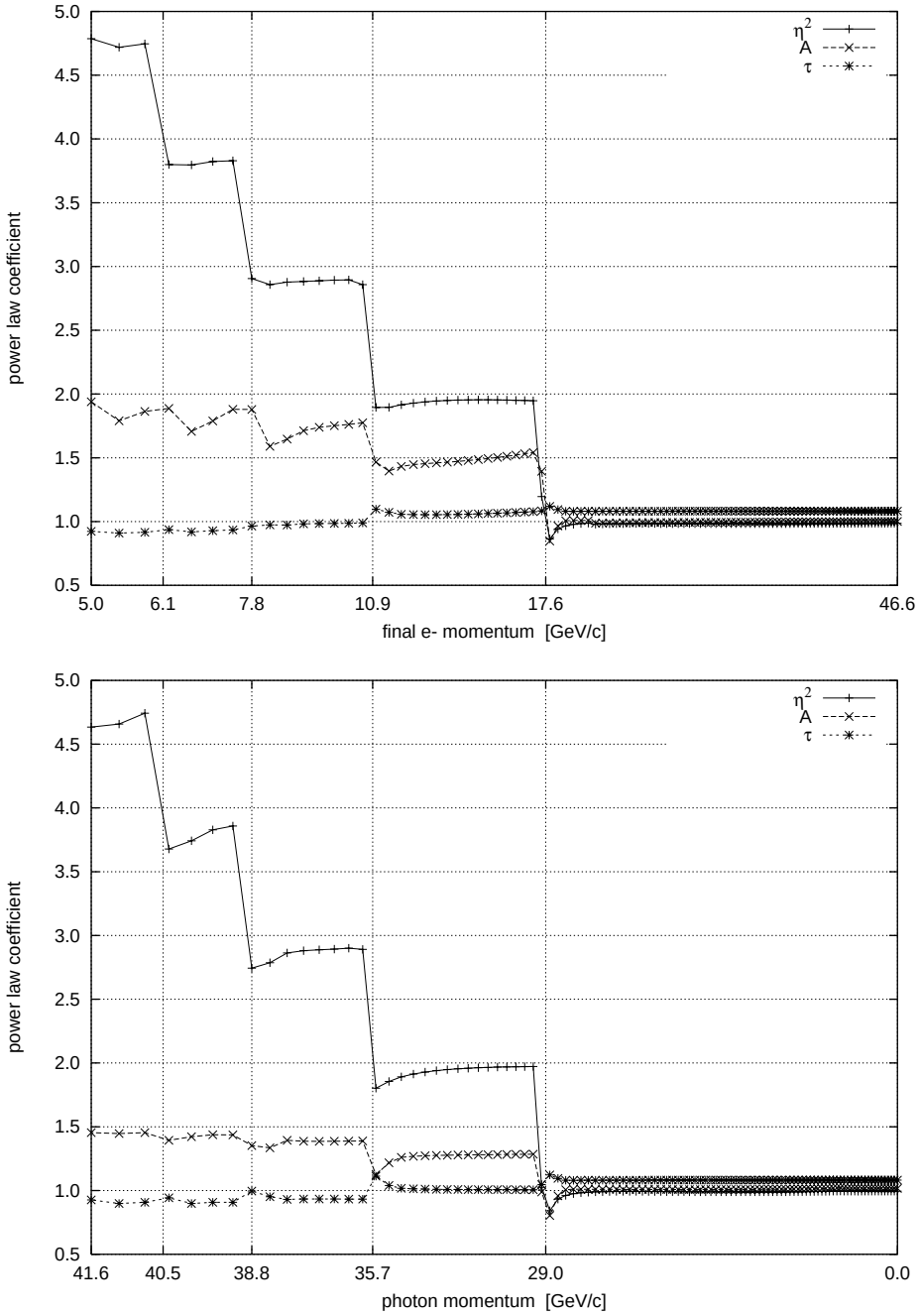

Figure 5.4: Fitted power law coefficients for the electron spectrum (top) and the photon spectrum (bottom).

The χ^2 is calculated as follows:

$$\chi^2 = \left(\frac{\delta \log E}{\sigma_{\delta \log E}}\right)^2 + \left(\frac{\delta \log n_1}{\sigma_{\delta \log n_1}}\right)^2 + \left(\frac{\delta \log n_2}{\sigma_{\delta \log n_2}}\right)^2 + \left(\frac{\delta \log n_3}{\sigma_{\delta \log n_3}}\right)^2, \tag{5.3}$$

where the symbols E, n_1, n_2, n_3 denote the pulse energy, the "EC31" or "EC37" estimate of total Compton scattering rate, the "N2MO" measure of $n = 2$ electron flux, and the "N3MO" measure of $n = 3$ electron flux, respectively; $\delta \log X \equiv \log(X_{(\text{measured})}/X_{(\text{calculated})})$. Units of E are mJ, and n_1 is measured in number of photons. Units of n_2 and n_3 are electrons per GeV/c, where the measurement is converted from raw ADC counts according to the cross-calibration-derived conversion constants [Chapter 4], and the simulation result is obtained by summing over a 1-GeV/c-wide flat acceptance centered at the central detector momentum.

The scaling factors for the deviations are

$$\sigma_{\delta \log E} = 0.33 \tag{5.4}$$

$$\sigma_{\delta \log n_1} = \sqrt{9000/n_1} \tag{5.5}$$

$$\sigma_{\delta \log n_2} = 2.26/\sqrt{(\texttt{N2MO ADC counts})} \tag{5.6}$$

$$\sigma_{\delta \log n_3} = 2.78/\sqrt{(\texttt{N3MO ADC counts})}. \tag{5.7}$$

The standard deviation $\sigma_{\delta \log n_3}$ is that calculated by Field [31] for the expected fluctuation of the detector signals due to the statistics of the Cherenkov photon collection and photoelectron conversion; the value of $\sigma_{\delta \log n_2}$ is Field's estimate multiplied by a factor of two in accordance with the practice of Bula and Koffas [6], motivated by a desire to balance the fit residuals of the two detectors. The value of $\sigma_{\delta \log n_1}$ is only a crude guess which reflects the expectation that the $n = 1$ measurement should be more precise; the choice if $\sigma_{\delta \log n_1}$ is largely immaterial in any case since the algorithm tries to match observed and calculated $n = 1$ signal exactly. The claimed precision and accuracy of the laser pulse energy monitor is $\pm 10\%$; however, a value of $\sigma_{\delta \log E}$ avoid corresponding to 33% accuracy was chosen to avoid unnaturally tight correlations of fitted pulse energy with other focal parameters and to give the most weight to the Cherenkov monitor data.

As will be discussed further below, the data can be almost as well fit by one fewer parameter: for example, by setting `cfit_max_it` to 0 and $\sigma_{\delta \log E}$ to 10000, the fitting algorithm above finds good fits using a fixed area while allowing arbitrary pulse energy. Likewise, one could take $\sigma_{\delta \log E} \to 0$ and allow the algorithm to fit area while forcing the pulse energy to be exactly as read by the energy monitor. This implies a certain degree of degeneracy between the focal parameters. This degeneracy is of concern only if a shift along the degenerate direction in parameter space can change the theoretical calculation for some other aspect of the experiment even as it leaves the calculated Cherenkov monitors largely unchanged. Whether this is or is not the case is addressed below.

The partial degeneracy among focal parameters and the arbitrariness in the choice of $\sigma_{\delta \log E}$ and $\sigma_{\delta \log n_1}$ might raise questions about the interpretation of χ^2 as defined in Eq. 5.3. This concern is allayed by the tight association between overlap with the $n = 1$ rate and fitted energy with the observed energy, leaving just $\delta \log n_2$ and $\delta \log n_3$ as the remaining contributors the χ^2 and η as the remaining fit parameter. Both $\sigma_{\delta \log n_2}$ and $\sigma_{\delta \log n_3}$ have a legitimate basis, and so χ^2 as defined above should be considered as being properly scaled and having one remaining degree of freedom after the fit.

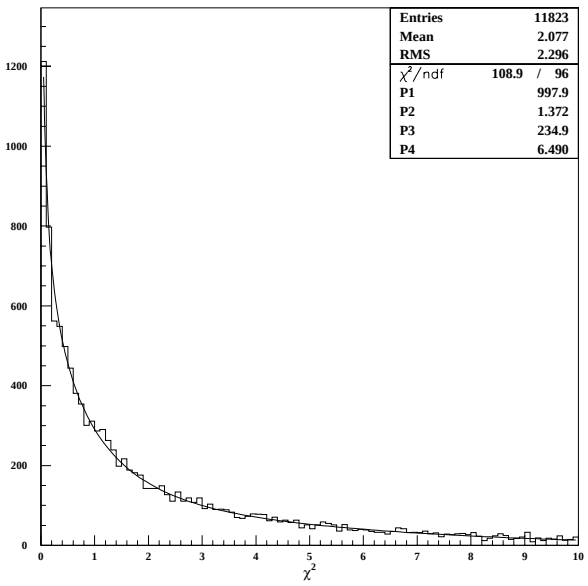

Figure 5.5: Fit χ^2 distribution for August 1996 data runs selected as described in text. The distribution appears to be the sum of two χ^2 distributions: one with $\langle\chi^2\rangle \simeq 1$ and one with $\langle\chi^2\rangle \simeq 6$.

5.5 Performance of the algorithm on real data

Figures 5.5 through 5.7 show the fit χ^2, fitted η, laser area, pulse energy, and $n = 1$ overlap ratio, and the fit residuals for low χ^2. All results in these figures are from a set of runs from August 1996 selected according availability of CCD data with either the "thick" aluminum or "thin" nitrocellulose converter foil in the photon beam line. (The CCD data is irrelevant to this analysis, but the Cherenkov monitor analysis is intended to be used to interpret the CCD data, thus motivating this choice of runs.)

Figures 5.8 through 5.10 show the same information for the same data set with area set to a fixed value and a large $\sigma_{\delta \log E}$ to essentially eliminate the laser pulse energy measurement information from the fit. The χ^2 distribution is not appreciably different. This shows that multiple solutions for the parameters $\eta, A, \Delta t$ can be found with statistically meaningless differences in the Cherenkov monitor residuals.

To assess the impact of this degeneracy among parameters, Fig. 5.11 shows the distribution of simulated rates for positrons, "$n = 4$"-region electrons, and "$n = 2$"-region photons for the standard and fixed-area fits. The two distributions are obviously quite similar and are in fact statistically indistinguishable. Evidently, multiple parameter sets can successfully model a given set of observed Cherenkov monitor signals, but these parameter sets also yield very similar calculated signals for the remaining detectors.

Note that if we trust the laser energy monitor absolutely, then there is no question of any non-trivial degeneracy in the focal parameters. The degeneracy between the "beam overlap factor" and laser pulse length is trivial, since the simulation shows linear dependence on these parameters

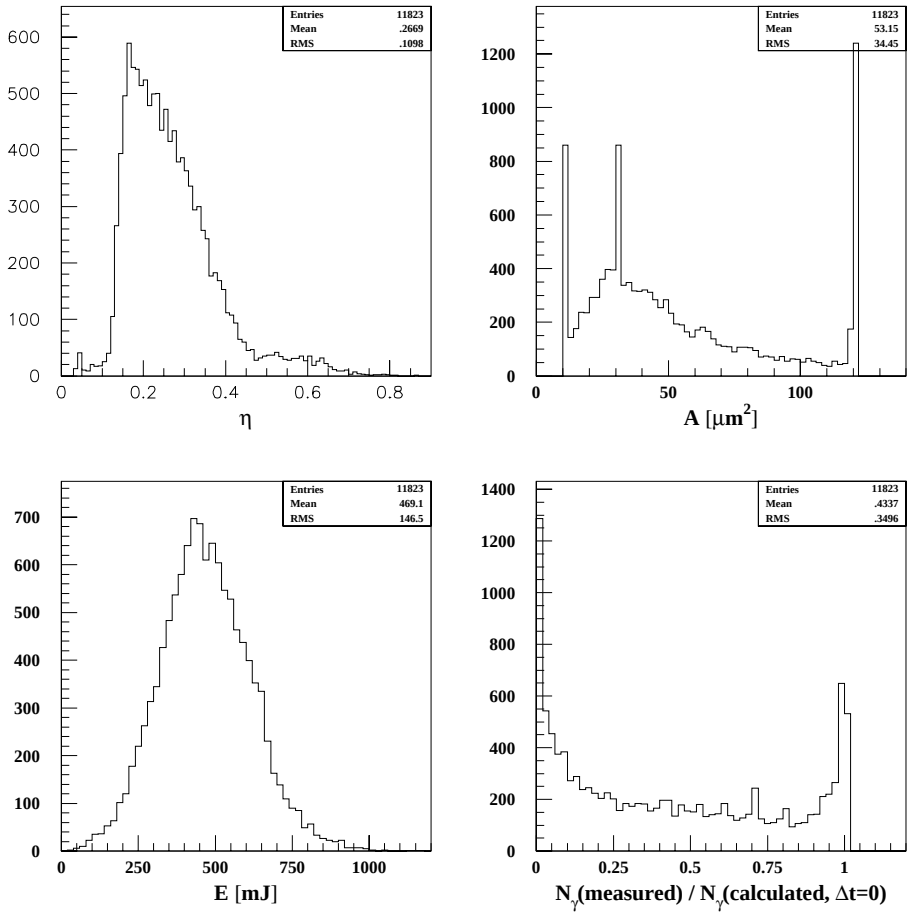

Figure 5.6: Distribution of fitted η, laser area, pulse energy, and $n = 1$ overlap ratio for August 1996 data runs selected as described in text. The spikes at the ends of the allowed range in the area plot are from events where a smaller- or larger-than-allowed area would further improve the χ^2; the spike at the default value of 30 μm^2 contains events in which no adjustment of area improved the fit significantly.

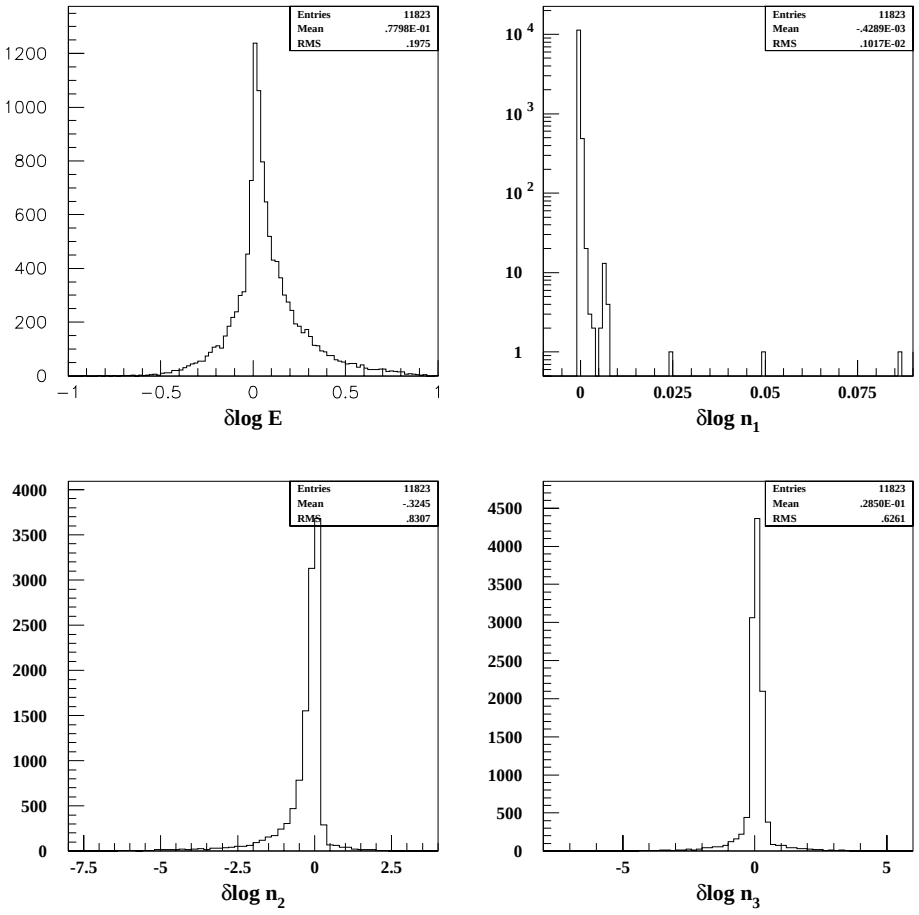

Figure 5.7: Distribution of fit residuals $\delta \log E, \delta \log n_1 \delta \log n_2, \delta \log n_3$ for August 1996 data runs selected as described in text.

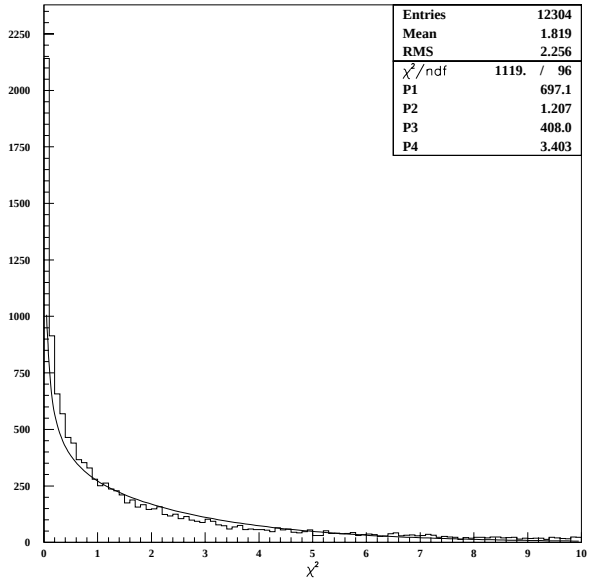

Figure 5.8: Fit χ^2 distribution for the same data set with fit forced to use fixed area and $\delta \log E$ eliminated from χ^2.

for all detectors at all orders: a single effective parameter comprises both overlap and pulse length. With pulse energy measured by a trusted monitor, the only other free parameter is the focal area. The fit of focal parameters to the observed signals at three different orders n then becomes a simple line fit of two parameters to three unknowns, which has no degeneracy.

5.6 Comparison with previous focal parameter fits

Previous focal parameter fits were done not only with different (arbitrarily scaled) calibration constants for the Cherenkov monitors, they were in fact done by different workers (Christian Bula and Thomas Koffas were these brave pioneers) using different software. Comparing the two estimates therefore helps check the internal and mutual consistency of the algorithms, as well as showing the effect of the rescaling.

In Fig. 5.12, the logarithm of the ratio of values of η from the previous and new simulations has been plotted as a function of η from the new simulation, using data from the thick foil CCD runs. Where the returned value of η is very small or very large, the routines behave differently: this is generally because such extreme values are actually from poor-statistics, low-overlap events. Only events with η between 0.15 and 0.40 were used in the final analysis: in this range, both algorithms returned compatible answers, with the "old" (scaled) η values consistently factor of $e^{-.2934}$ below the "new" (calibrated) η values.

The Koffas-Bula fits used in the positron analysis maintained the fixed relationship $A = (20\ \mu\mathrm{m}^2/\mathrm{ps})\,\tau$. For purposes of simplifying the comparison, the algorithm described here was also contrained to vary A and τ together according to this relationship. The "new" values of

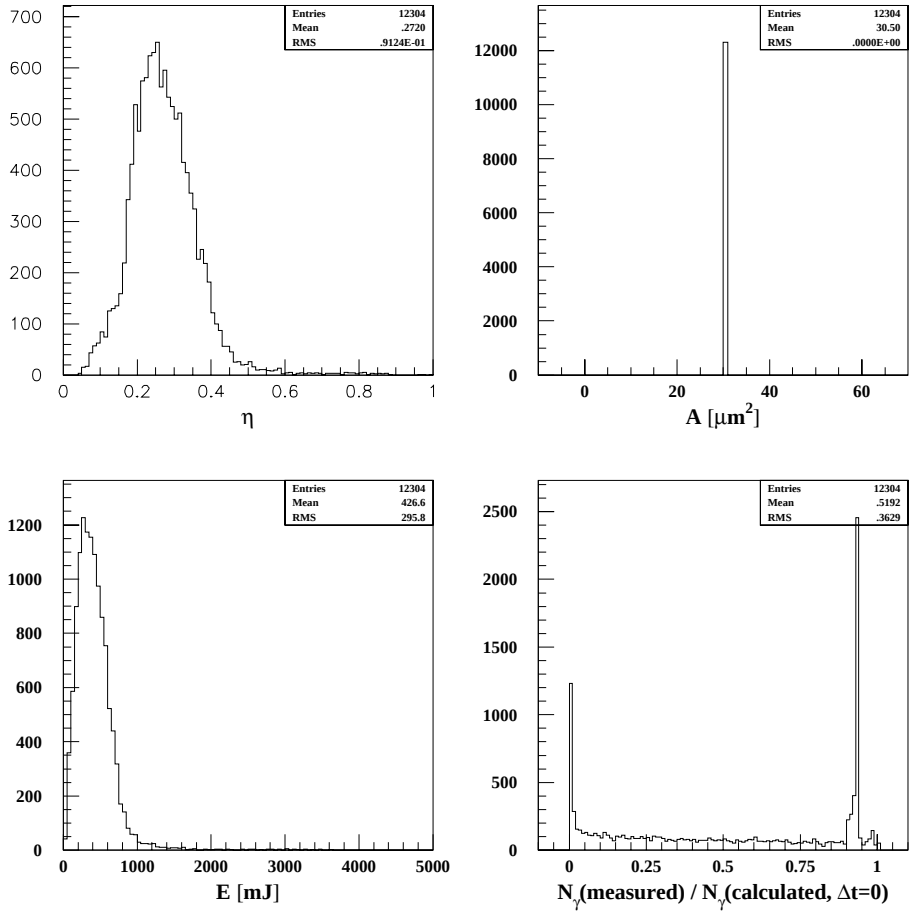

Figure 5.9: Distribution of fitted η, laser area, pulse energy, and $n = 1$ overlap ratio for the same data set, with fit forced to use fixed area and $\delta \log E$ eliminated from χ^2.

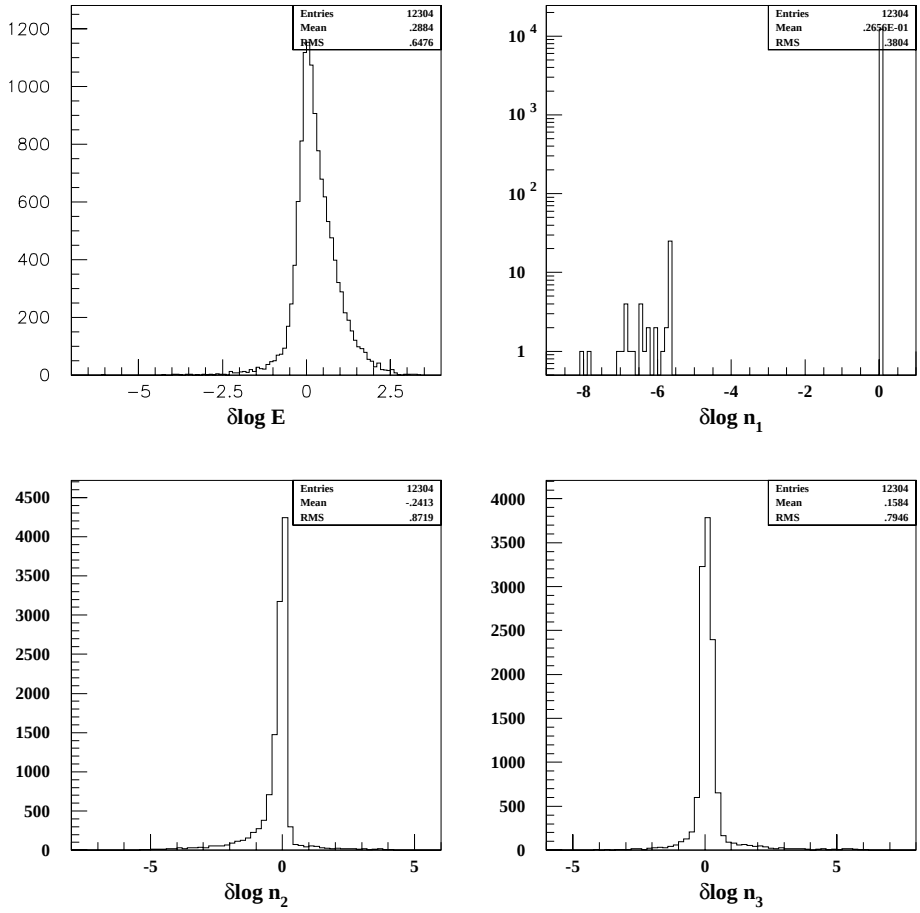

Figure 5.10: Distribution of fit residuals $\delta \log E, \delta \log n_1 \delta \log n_2, \delta \log n_3$ for the same data set, with fit forced to use fixed area and $\delta \log E$ eliminated from χ^2.

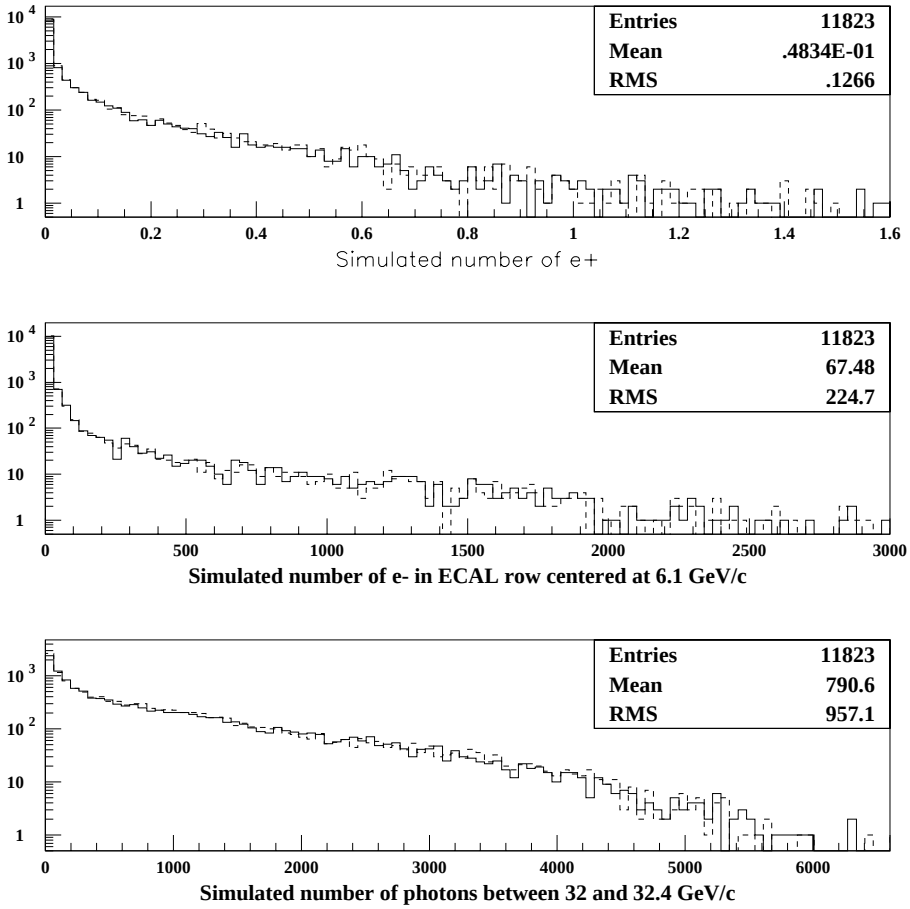

Figure 5.11: Comparison of simulated rates for positrons (top), $n = 4$-region electrons (middle), and $n = 2$-region photons (bottom) calculated using fit results from standard Cherenkov monitor fit (solid line) and fixed-area fit (dashed line), using Cherenkov monitor data from the data set as described in the text.

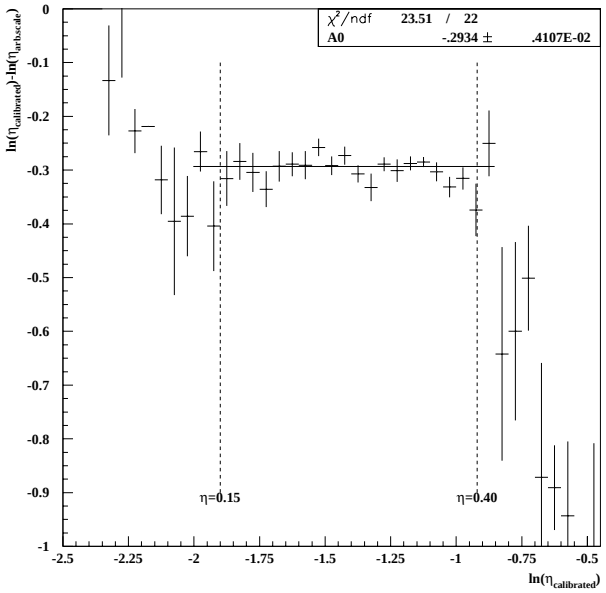

Figure 5.12: Comparison of η obtained by original Koffas-Bula algorithm with old scaling constants and algorithm presented here with new calibration.

A, τ, and η thus obtained are also used for the CCD track spectrum analysis, again to simplify understanding of the impact of the new analysis on the old results.

5.7 Reduction of degeneracy by definition of a new parameter

A simple "linear model" expresses the scattering rates as a product of various quantities raised to fixed exponents. (Taking the logarithm makes the linearity manifest.) The detailed simulation fits this approximation quite well over a wide range. For example, near the "standard parameters" for the August 1996 run, we have

$$\begin{aligned}
N_1 &= k_1\,\eta^{2.0}A^{1.0}\tau v &&= k_1\,E^{1.0}v \\
N_{2e} &= k_{2e}\eta^{3.9}A^{1.5}\tau v &&= k_{2e}E^{2.0}A^{-0.5}\tau^{-1.0}v \\
N_{3e} &= k_{3e}\eta^{5.8}A^{1.7}\tau v &&= k_{3e}E^{2.9}A^{-1.2}\tau^{-1.9}v \\
N_{4e} &= k_{4e}\eta^{7.6}A^{1.7}\tau v &&= k_{3e}E^{3.8}A^{-2.1}\tau^{-2.8}v \\
N_{2\gamma} &= k_{2\gamma}\eta^{4.0}A^{1.3}\tau v &&= k_{2\gamma}E^{2.0}A^{-0.7}\tau^{-1.0}v \\
N_{e+} &= k_{e+}\eta^{10.2}A^{2}\tau v &&= k_{e+}E^{5.1}A^{-3.1}\tau^{-4.1}v,
\end{aligned}$$

where η, A, τ have their usual meanings, E is the laser pulse energy, and $v \cong \exp(-(\Delta t/\sigma_t)^2/2)$ is the overlap factor.

The pulse energy E is measured, the constants $k_{1,2e,3e,\ldots}$ are given by the theory, and the overlap factor v estimation is dominated by the linear rate measurement, leaving the following two quantities as the non-linear luminosity estimators:

$$L_{2e/1} \equiv \frac{N_{2e}/k_{2e}}{N_1/k_1} = EA^{-0.5}\tau^{-1} \tag{5.8}$$

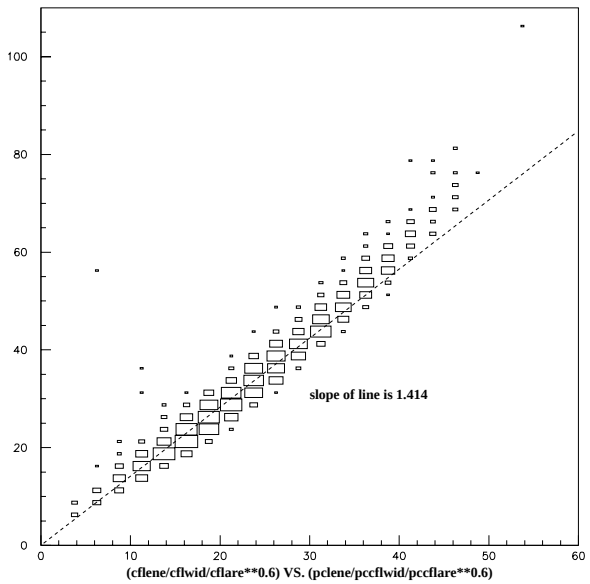

Figure 5.13: Comparison of "original" and "new" values of $y \equiv \frac{E}{\tau A^{0.6}}$.

$$L_{3e/1} \equiv \frac{N_{3e}/k_{2e}}{N_1/k_1} \quad = \quad E^{1.9} A^{-1.2} \tau^{-1.9}. \tag{5.9}$$

Noting how similar $L_{3e/1}$ is to the square of $L_{2e/1}$, let us define

$$y \equiv \frac{E}{\tau A^{0.6}}. \tag{5.10}$$

In terms of the parameter y, the luminosity measurements are

$$L_{2e/1} = y^{1.0} A^{+0.1} \tag{5.11}$$

$$L_{3e/1} = y^{1.9} A^{-0.1}. \tag{5.12}$$

For a given y, an order of magnitude change in the area A will only change the measured rates by about 25%. It is thus clear that our nonlinear luminosity monitors do not provide separate measures of A, τ, or η, but only two measurements of the single parameter y.

Expressed in terms of y, the net effect of the recalibration of the Cherenkov monitors is not as large as it first appeared. (The true magnitude of the changes was obscured due to comparing η, A, and τ separately, instead of comparing the single parameter y.) Figure 5.13 shows a scatter plot of the "original" and "new" values of y. The "new" values are higher than the original by about 40%, *i.e.*, a factor of $\sqrt{2}$.

The other normalized luminosities are given in terms of y as

$$L_{2\gamma/1} = y^{1.0} A^{-0.1} \tag{5.13}$$

$$L_{4e/1} = y^{2.8} A^{-0.4} \tag{5.14}$$

$$L_{e+/1} = y^{4.1} A^{-0.6}. \tag{5.15}$$

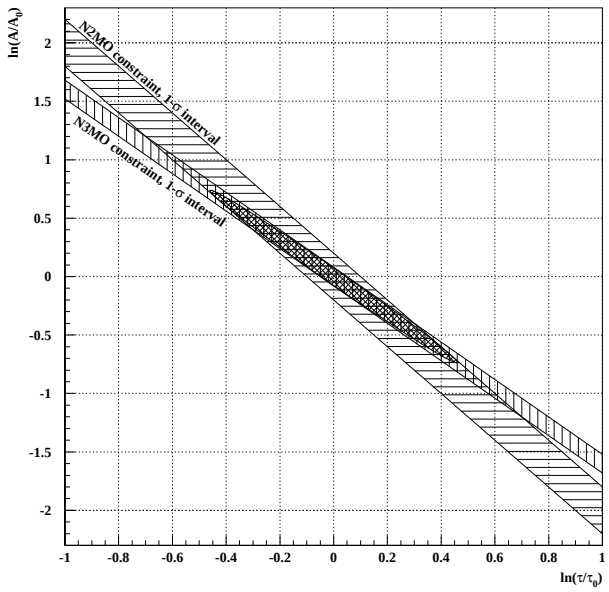

Figure 5.14: One standard deviation intervals from the N2MO and N3MO data plotted in the parameter space of focal area (A) and pulse length (τ), for a typical event in which both N2MO and N3MO have estimated standard deviations of $\pm 10\%$. The heavily shaded ellipse is the one-sigma error ellipse of the joint distribution, transformed into A and τ parameters according to Eqs. 5.8 and 5.9.

Thus, whether E, τ, or A is used as adjusted as a free parameter (or all of them) will not appreciably affect the prediction for $L_{2\gamma/1}$, the photon $n = 2$ to $n = 1$ ratio. This means that a discrepancy between prediction and observation in this detector cannot easily be resolved by changing assumptions regarding focal parameters. On the other hand, the choice of how to adjust the focal parameters has a profound affect on the agreement of the $n = 4$ electron rate and the positron rate with the prediction. This freedom somewhat reduces the extent to which we can claim to have demonstrated the correctness of the theory.

5.8 Graphical depiction of the problem

The fundamental difficulty in reconstructing focal area and pulse length information from N2MO and N3MO is illustrated in Fig. 5.14. In this case, both N2MO and N3MO have an estimated fractional standard deviation of 10%. However, comparable precision is not achieved in either the reconstructed pulse length or focal area: the 1-σ error ellipse projected on the τ axis represents a $\pm 50\%$ uncertainty in the pulse length, and the projection on the A axis gives a $\pm 75\%$ uncertainty in the area.

As depicted in Fig. 5.15, this uncertainty translates into approximately a $\pm 45\%$ uncertainty in the positron yield, but only a $\pm 7\%$ uncertainty in the $n = 2$ photon yield.

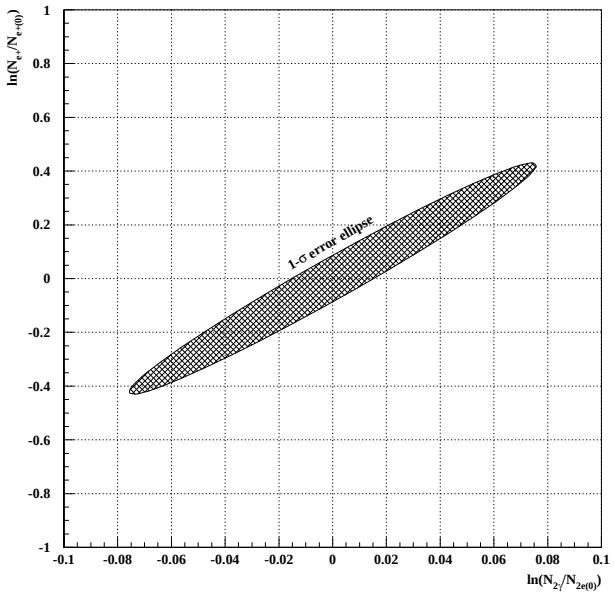

Figure 5.15: The one-sigma error ellipse from Fig. 5.14 transformed into the parameter space of positron and $n = 2$ photon yield according to Eqs. 5.11 through 5.15.

5.9 Conclusion

The Cherenkov monitors provide information that can be used to calculate rates in other detectors consistent with the accepted theory of intense field quantum electrodynamics. The method presented above is one way of arriving at such theoretical calculations: focal parameters are chosen which match theory to observation in the chosen detectors, and the rates in the remaining detectors are calculated using these focal parameters. Although the mapping of observed Cherenkov monitor signals to consistent focal parameters is one-to-many (when the laser energy monitor is ignored), the corresponding many-to-one mapping of focal parameters to other detectors of interest reduces or eliminates the theoretical uncertainty introduced by this method of inferred parameters. Specifically, measurement of the $n = 2$ and $n = 3$ electron yield to $\pm 10\%$ precision only constrains the area to $\pm 75\%$ and the pulse length to $\pm 50\%$; however, the $n = 2$ photon yield theoretically can be predicted with $\pm 7\%$ precision, and the positron yield with $\pm 45\%$ precision. Thus, the internal consistency of our calculations from quantum electrodynamics may still be tested against observation despite the lack of direct measurement of the properties of the laser focus.

Chapter 6

Features of the E-144 CCD Tracking Spectrometer and Monte-Carlo Simulation

Synopsis

The layout, hardware, software, and actual performance of the CCD tracking spectrometer system are presented. A Monte Carlo simulation of the system is also described.

6.1 Hardware and software

Photons produced at IP1 proceed down their own beamline through the converter foil and the CCD tracking spectrometer, as previously described [Chapter 1]. In this chapter, the CCD tracking spectrometer system will be examined in detail. Figure 6.1 shows the magnet and eight CCD packages inside their evacuated chambers. Figure 6.2 shows the precise CCD positions used for the August 1996 data; also shown are the trajectories of electrons and positrons of various momenta when the magnet is set to 247.5 MeV/c kick. Note that 29 GeV/c is the $n = 1$ kinematic limit, so the positrons and electrons intercepting the inner portions of the CCDs could only have been created from $n \geq 2$ photons, while signal in the outer portions of the CCDs were dominated by the $n = 1$ flux.

Each CCD plane consisted of a large area CCD image sensor and associated support electronics. The planes were mounted on remotely-controlled stages inside an evacuated chamber. Bulk cooling was provided by circulating chilled ethylene glycol in copper pipes brazed onto a copper backplane surrounding the CCD sensor; thermo-electric coolers further lowered the temperature of the sensors. Cooling the CCDs to below 0°C significantly reduced leakage current and associated noise. Thermocouples allowed the temperature of each plane to be remotely monitored. The external DC bias voltage on each plane could also be remotely controlled.

The eight CCDs were clocked from a single source: the read-out began shortly after each laser shot, and was completed within one second. The CCDs output a standard raster signal, with "blacker-than-black" frames around each scan line. These signals were digitized and processed by eight commercially-available frame-grabber boards with built-in digital signal processing (DSP) capability. The on-board DSP chip subtracted pedestal frames (acquired during dedicated pedestal

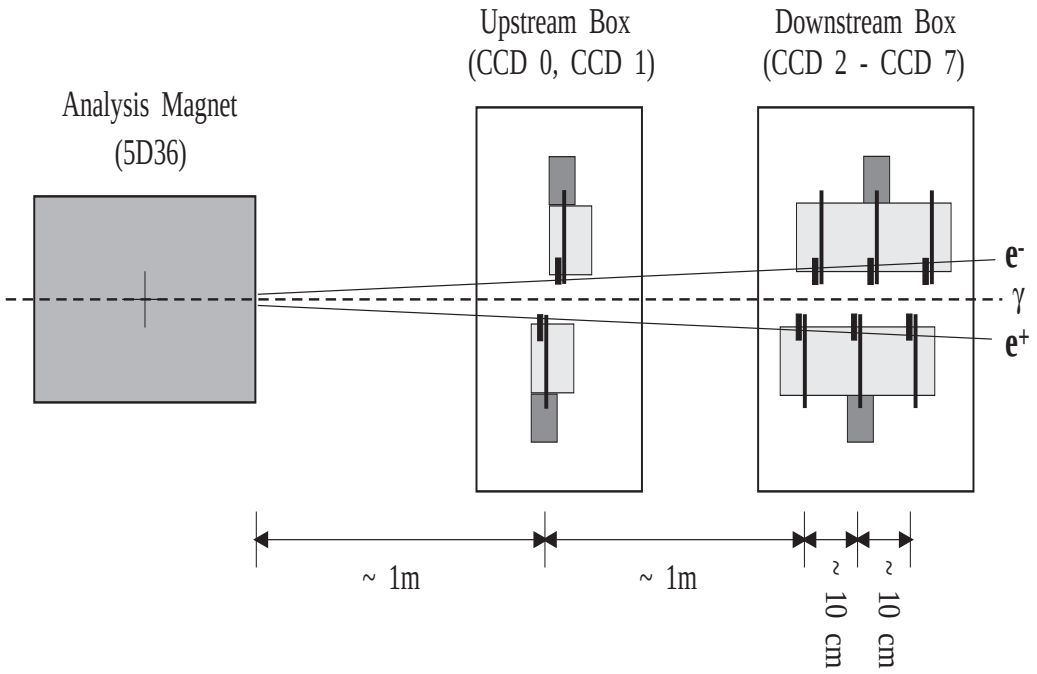

Figure 6.1: The CCD tracking spectrometer.

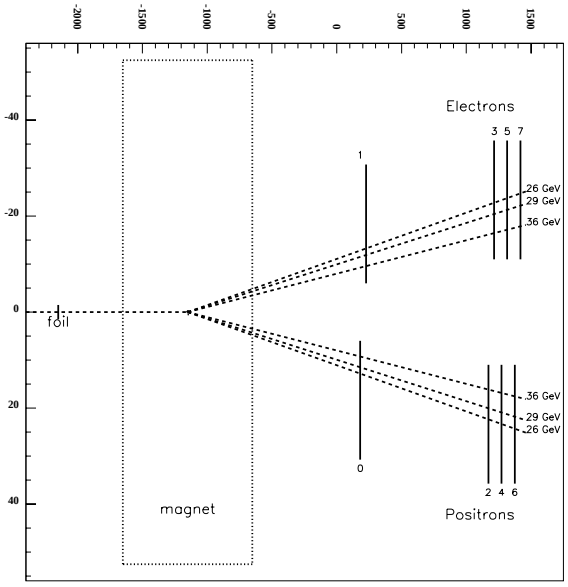

Figure 6.2: Positions of the active CCD areas and several particle trajectories for magnet set to 247.5 MeV/c kick. Dimensions are in millimeters.

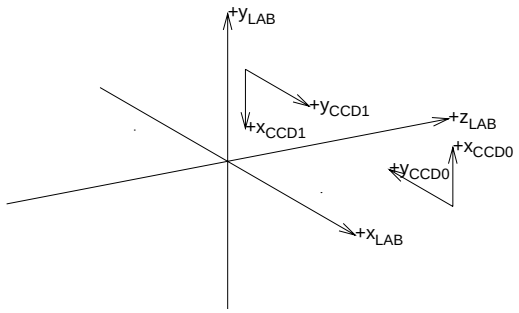

Figure 6.3: Coordinate systems.

runs), calculated line-by-line DC offset corrections, and converted signal level information into a stream of hit coordinates.

The eight frame-grabber/DSP boards were installed in a standard EISA-bus micro-computer which ran as an E-144 front-end data acquisition computer. The DSP boards wrote the stream of hit coordinates directly to the computer's memory. The CCD front-end computer wrote the complete event data to its hard disk, and sent summary information to the E-144 back-end computer. Because of the long read-out time for the CCD system, the CCD front-end computer was not required to return data to the back-end on every event, and was allowed to return its summary data one laser shot late. A common event number was used for the CCD hard disk file the data acquisition stream, and the data was recombined off-line.

6.1.1 Coordinate systems

The CCD tracking spectrometer system enjoys a rich diversity of applicable coordinate systems (Fig. 6.3).

Each CCD plane has its own internal coordinate system, aligned with its raster pattern. The CCDs are mounted with their light-sensitive surfaces facing downstream, away from the incident beam. (This does not reduce their ability to detect particle tracks.) They are rotated such that their major (slow) scan direction is towards the center of the beamline: thus, the CCDs on the positron arm are rotated approximately 180° as compared to the CCDs on the electron arm. The coordinate axes are labeled $x_{\mathrm{CCD}i}$ and $y_{\mathrm{CCD}i}$ in Fig. 6.3, where $x_{\mathrm{CCD}i}$ is the rapid scan direction (commonly the "horizontal" direction in a video system, but here vertical), and $y_{\mathrm{CCD}i}$ is the slow scan direction.

The global coordinate system for the spectrometer will be identified by the subscript LAB. The

$+\hat{z}_{\rm LAB}$ direction is parallel to the photon beamline (east), the $+\hat{y}_{\rm LAB}$ direction is up, and the $+\hat{x}_{\rm LAB}$ axis points towards the positron arm of the spectrometer. This is a left-handed coordinate system. These coordinates are essentially the same as the standard Stanford Linear Collider coordinate system except for a reversal of the sign of the x axis.

An alignment procedure described below determines the origins and slight rotations of the CCDi coordinate axes within the LAB system.

6.2 Set-up and data-taking procedures

Prior to each block of beam time, the alignment of the fixed collimators in the photon beamline was checked with a portable, low-power laser. Most other systems were also checked at this time; however, long-term operation of the thermo-electric coolers was deferred until after the vacuum system was sealed and pumped down. Once the cooling system was fully operational, a final check was performed to verify CCD noise levels were acceptable and that bias levels did not exceed the available range.

Immediately prior to data-taking, the external DC bias levels were adjusted such that the unexposed pixels read slightly above 0 ADC counts while the "blacker-than-black" frames of the CCD signals read near 255 ADC counts. The number of noise hits per plane per event was noted for later reference: typically it was 100 to 200 hits per event.

Proper steering of the high energy photon beam through the collimators was required in order to minimize synchrotron backgrounds and undesirable beam-clipping effects. To accomplish this, a foil was inserted at the interaction point, which created a steady, intense beam of high-energy photons, and the electron beam's trajectory was adjusted manually to maximize transmission of photons to the end of the line. Then, the signal on four scintillators around the beamline downstream of the CCDs was monitored while the beam trajectory was scanned systematically by computer: the best trajectory was one which minimized the scintillator signals while maintaining transmission. A computer-controlled feedback was activated to maintain the beam steering. Finally, the foil at the interaction point was removed, and the CCD stages were sent to their "home" position, directly in the photon beam line. These CCDs were able to directly image the synchrotron light from the electron beam. The setpoints of the beam trajectory feedback were then fine-tuned such that the collimators blocked the synchrotron light from the 0.5-mrad bending magnets on either side of the interaction region. The synchrotron radiation from the 0.06-mrad bends remained, with a slight overlap which indicated the beam position.

After beam steering, the positions of the CCDs and gamma-converters were found to within 0.3 mm using the synchrotron light. The aluminum foil on the gamma-converter fork blocked the synchrotron light when inserted in the beam, and so its edge could be located precisely. The other materials could then be found by moving the converter fork a known amount in relation to this reference. Similarly, once the beam position was known from the synchrotron light image with the CCDs at their "home" position, it was possible to know their position with respect to the beam after moving them by counting steps on the stage motors.

The total alignment process took approximately 16 hours of beam time. Once this process was complete, the CCDs were moved into position to intercept tracks in the desired momentum range and data-taking was begun.

During the course of data-taking, the number of hits in each CCD was monitored in real time. The electron beam was periodically deliberately suppressed, allowing the number of noise hits to be

checked and compared with the previously-recorded levels. This provided a quick way to determine if the bias levels needed adjustment: when the number of hits became too low or too high, bias levels were adjusted and pedestal frames reacquired. The electron beam trajectory, beam current, laser performance, collision efficiency, and photon-loss scintillators were also monitored in real time.

A final off-line alignment of the CCD positions in their data-taking positions was performed[1] using the track data. Starting from the approximate alignment described above, offsets were adjusted to maximize the correlation of hit positions between planes. After track reconstruction [Chapter 7], slight relative rotations and offsets in the horizontal and vertical directions were made to minimize any systematic deviation of hits from their tracks.

The absolute position and angle of an entire arm of the spectrometer is inherited from the more basic alignment. Indeed, after all analysis was complete, it became apparent that the rear CCD stages must have been further out and had a slight rotation of their upstream faces towards the beamline, resulting in slightly too-shallow tracks. This is apparent in the offsets of the edges in the raw hit distributions, for example (as discussed in connection with Figs. 6.10 and 6.11 below), as well as in the location of the reconstructed kinematic edge. One way of addressing this issue would be to adjust the alignment data and reprocessing all the data, but an equally-valid method is to correct the reconstructed track data by adding 0.36 mrad to the magnitude of all track x angles; this later approach was used for the present data.

6.3 Performance

In this section, the performance of the system as a set of "hit" detectors is detailed. The track reconstruction performance is detailed elsewhere [Chapter 7].

6.3.1 Stochastic noise hits

One aspect of system performance is the presence of random hits from stochastic noise sources such as leakage current. Figure 6.4 shows the charge distribution of hits seen during a "pedestal run" (number 15319) during which no beam was present. These beam-off background hits are almost exclusively due to noise.[2] In Fig. 6.4, data has been summed over all events and all planes: the multiple peaks at the high-charge endpoint are due to different saturation levels (after pedestal subtraction) in the different planes.

Figure 6.5 shows the two-dimensional distribution of background hits in CCD plane 7 summed over the entire run. Physical positions in the lab coordinate system and the corresponding track momentum (for magnet set to 247.5 GeV/c kick) are indicated. The distribution is uniform, except at the leading edge of each scan line (physically at the $+\hat{y}_{\rm lab}$ edge, due to the rotation of the CCDs), where a transient "ghost" of the frame signal biases the hit threshold significantly: this transient has decayed to insignificance well before the interesting area ± 3 mm around $y_{\rm lab} = 0$. In short, the stochastic noise may be regarded as essentially uniform in its position distribution.

[1] Alignment analysis by Eric Prebys.

[2] In the 163 events of this run, 35878 hits were found in plane 7, and 4 tracks were found in the spectrometer overall. This number of tracks far exceeds the number expected from random chance, and suggests a horizontal particle flux of 30 $\mathrm{m}^{-2}\mathrm{s}^{-1}$, approximately one-third of the (mostly vertical) cosmic ray flux. These horizontal particles are probably muons produced by the electron beam in the upstream beam stopper. (The ionization energy loss of a high-energy muon in a CCD plane is identical to that of an electron.) In any event, the few extra hits in the rear CCDs due to this background particle flux are insignificant compared to the stochastic noise hits.

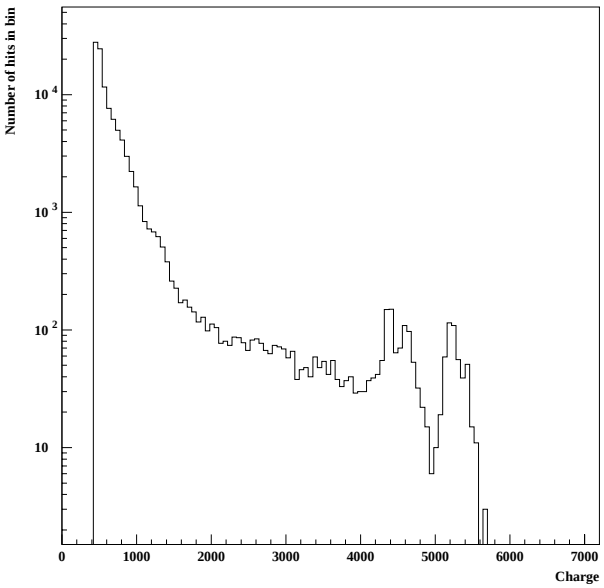

Figure 6.4: Distribution of charge of hits seen during a "pedestal run" 15319, during which no beam was present.

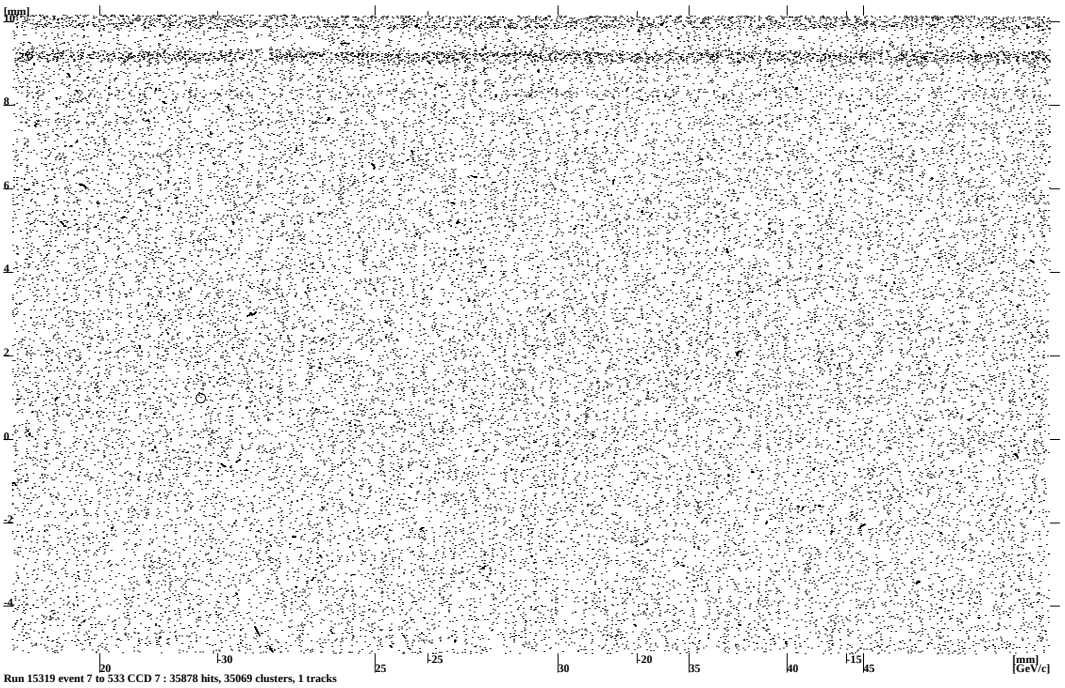

Figure 6.5: Two-dimensional distribution of hits seen in CCD plane 7 during pedestal run 15319, summed over the entire run. The circle indicates the intercept of the one track detected in this arm of the spectrometer.

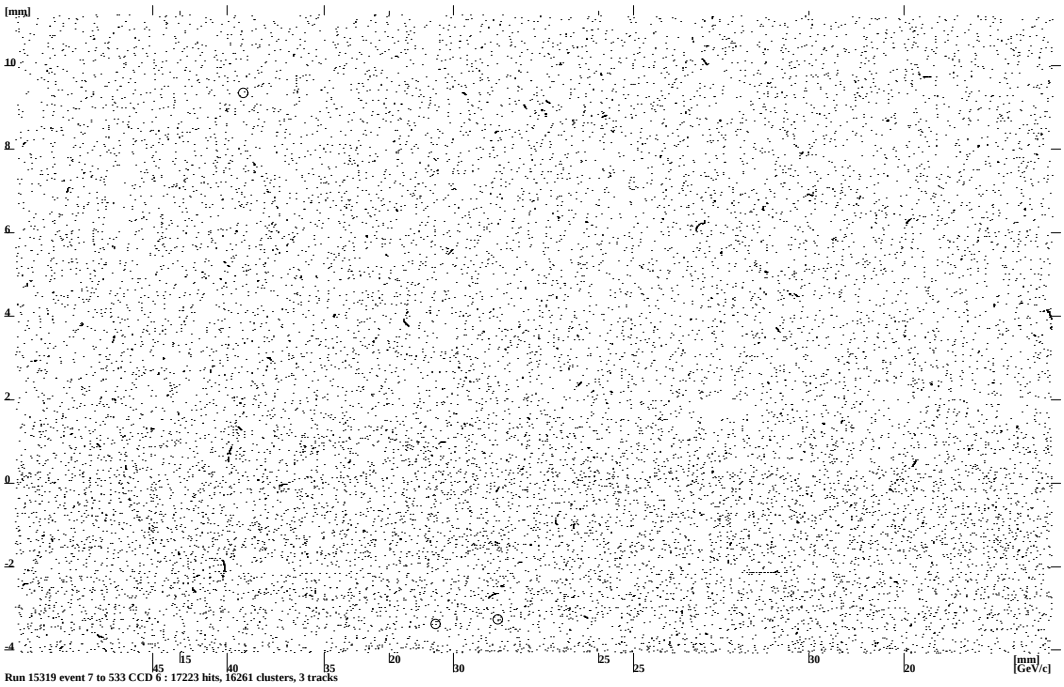

Figure 6.6: Two-dimensional distribution of hits seen in CCD plane 6 during pedestal run 15319, summed over the entire run. The three circles indicate the intercepts of the three tracks detected in this arm of the spectrometer.

6.3.2 Signal hits

Even more interesting than the properties of the stochastic noise hits are the properties of hits created by real particles passing through the CCDs. For laser-electron collision data, one way of distinguishing stochastic noise hits from "real" hits is by position: hits from particles originated by collisions at IP1 fall in a narrow band in each CCD, due to the small emittance and high energy of the electron beam. Another distinguishing feature is whether the hits can be combined to form tracks.

The narrow band referred to above may be seen by plotting the two-dimension distribution of hits separated according to the number of tracks on which they lie. Figure 6.7 shows the positions of raw hits and track intercepts in plane 7, from data taken with the thin foil converter and modest laser energy (run number 15228). Figure 6.8 shows the two-dimensional distribution of hits in a different form, where the data has been binned in two dimensions; each bin is 1 pixel tall and 22 pixels wide. At left, only hits used to form tracks are shown, and at right, hits which could not be matched are shown. The distribution of background hits remains uniform when the beam is present.

The hit density in the "signal" region of the CCDs can be extremely high, as shown in Fig. 6.9, which shows a *single* event in run 15296, a high laser energy, thick foil run. Figure 6.12 shows the distributions of matched and unmatched hits after summation over all events. The 22-pixel bin with the highest occupancy integrated 198 hits over the 1479 events in this run. The mean number of matched hits per event in this data is 184, while the peak number of matched hits per event exceeds 1800, so the peak hit density per pixel approaches 0.06 hits per pixel. The effect of high

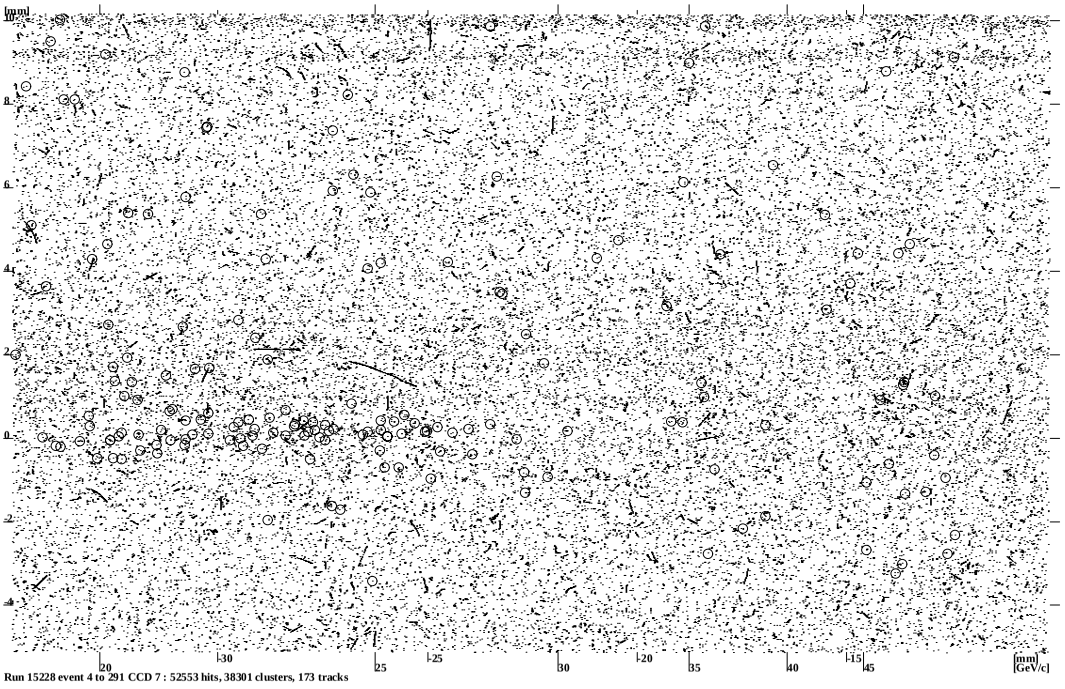

Figure 6.7: Two-dimensional distribution of hits and tracks seen in CCD plane 7 during run 15228, summed over the entire run.

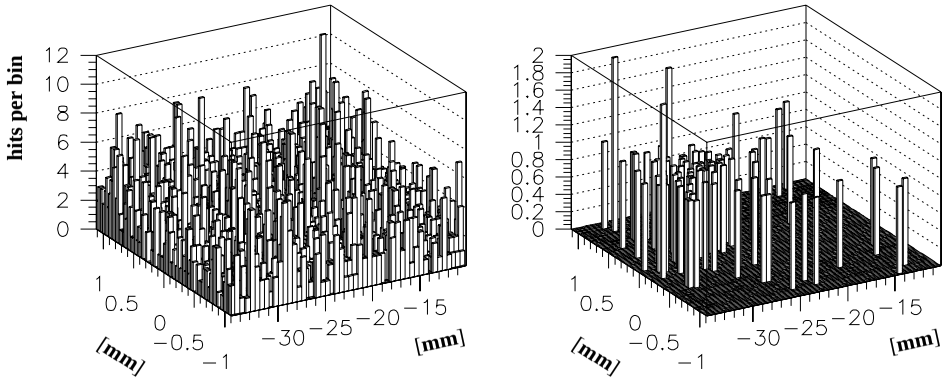

Figure 6.8: Lego-style histograms showing the two-dimensional distribution of hits for thin converter foil data (run 15228): at left, hits which could not be matched with tracks are used, and at right, only hits used to form tracks are included. Bin area is 22 pixels; 73 events are summed.

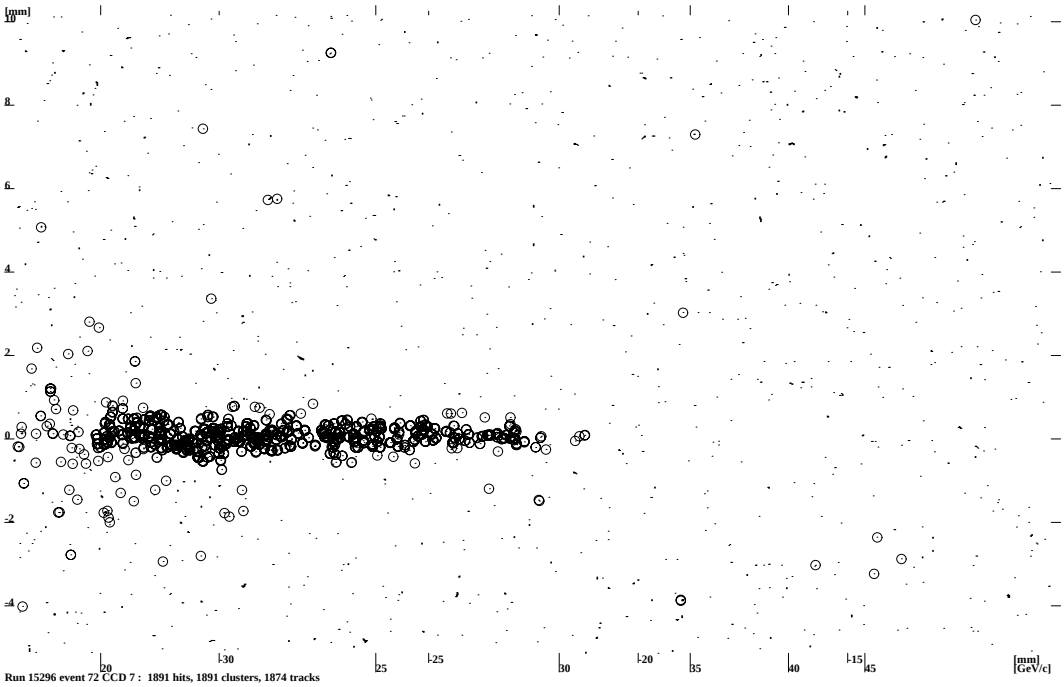

Figure 6.9: Hits in plane 7 for a single event in run 15296.

hit density is discussed further in [Chapter 7].

Fig. 6.10 shows the data plotted in Fig. 6.12 (and similar data for the other planes) projected onto the x axis, for each plane of the spectrometer, with the expected distribution shown with fitted horizontal offset, vertical scale factor, and noise hit level. Fig. 6.11 shows the hit distribution for data from run 15339, a similar data set in all respects except for the magnet setting: 219 MeV/c kick instead of 247.5 MeV/c. Note that in these figures, only hits lying within ± 49 pixels of $y_{\mathrm{LAB}} = 0$ have been used, thus including essentially all "signal" hits while reducing the number of uniform background hits by roughly a factor of 7. The estimated errors on the fitted horizontal shifts (denoted δx in the figures) are each ± 0.02 mm. The horizontal shifts are consistent with a systematic over-estimation of the separation of the two arms by 1.6 mm, with a compensating inward rotation of each arm.

The hit distributions show a sudden and unexpected cutoff on the outside. This is a general feature of the data, present in all CCDs. For a given magnet setting, the cutoff appears at the same momentum in each CCD in the back planes, at a different physical position in each plane, suggesting that the cutoff does not originate from a CCD-specific effect. The cutoff is present in the front planes as well, but the location of the cutoff cannot be determined accurately due to the heavy saturation of these planes. At different magnet settings, the cutoff is always at the same positions in each CCD, suggesting that the cutoff originates from some physical aperture upstream of the CCDs and downstream or inside of the magnet. According to the design, no such physical aperture is supposed to exist. This effect is not understood, but since the analysis only considers tracks with momenta over 26 GeV/c, it assumed that the phenomenon can be safely ignored.

Figure 6.13 shows charge distributions for unmatched hits, hits uniquely matched to a single track, and hits used in multiple tracks, for the thick-foil data in run 15296. Note that the charge of a hit is only weakly used in deciding whether a track is valid: in order to be accepted as a track,

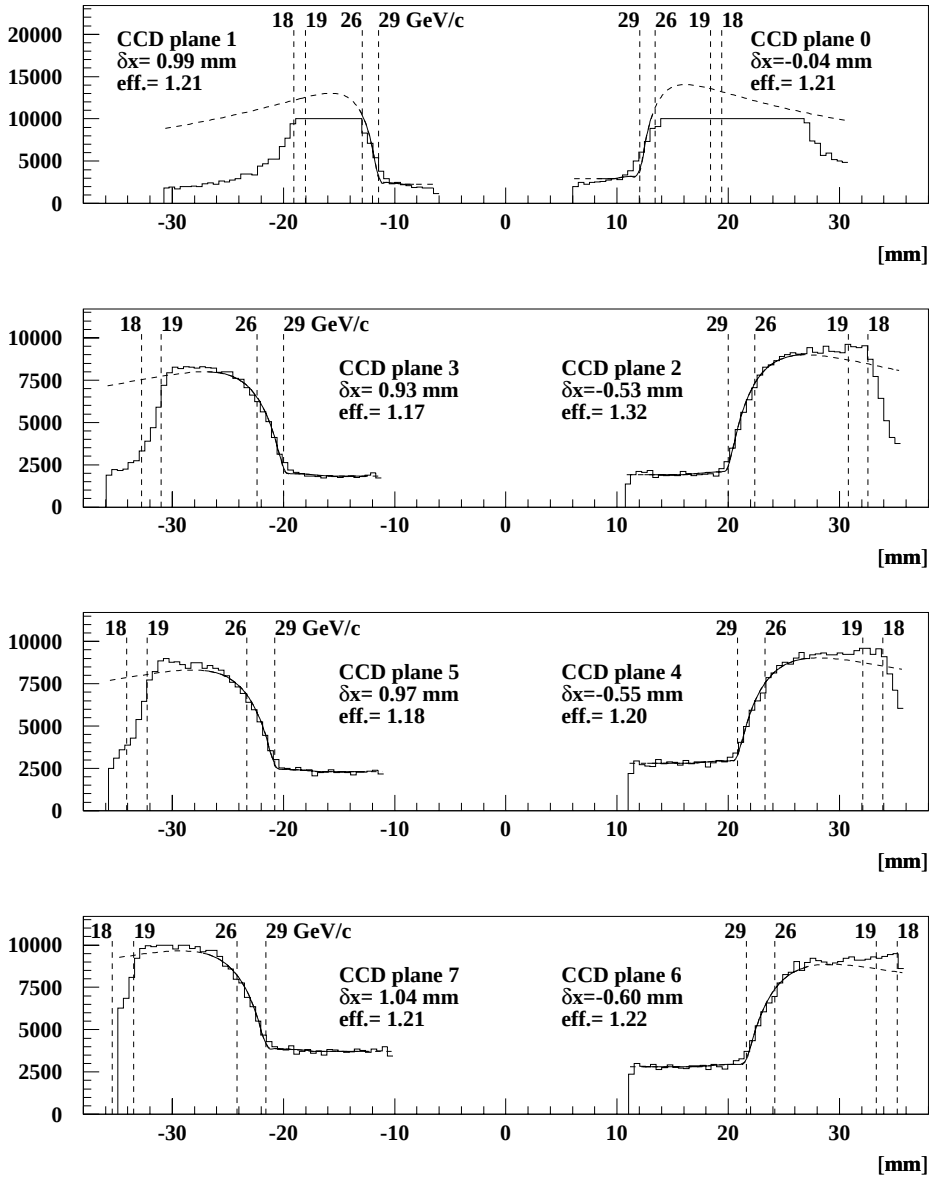

Figure 6.10: Hit position distribution in all planes in run 15296.

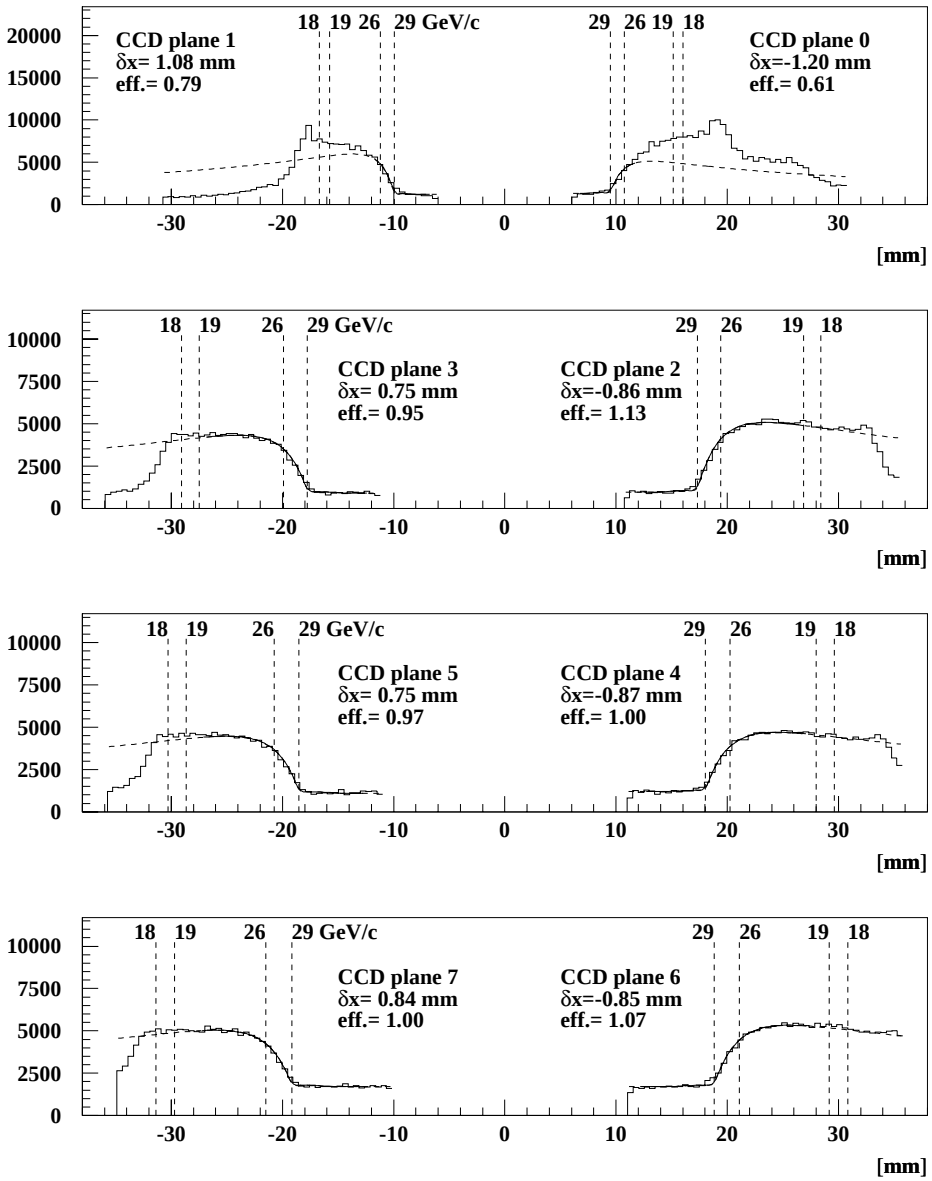

Figure 6.11: Hit position distribution in all planes in run 15339, a data set quite similar to run 15296 except that the spectrometer magnet 5D36 was set to 219 MeV/c kick instead of 247.5 MeV/c.

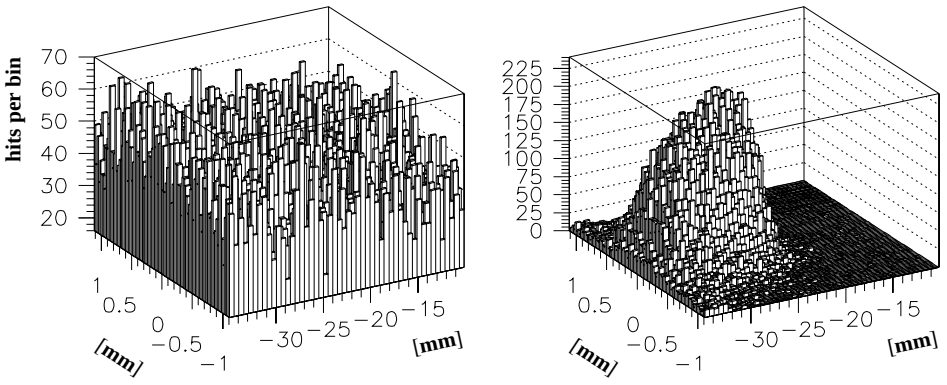

Figure 6.12: Lego-style histograms showing the two-dimensional distribution of hits for thick converter foil data (run 15296): at left, hits which could not be matched with tracks are included, and at right, only hits used to form tracks are used. Bin area is 22 pixels; 1479 events are summed.

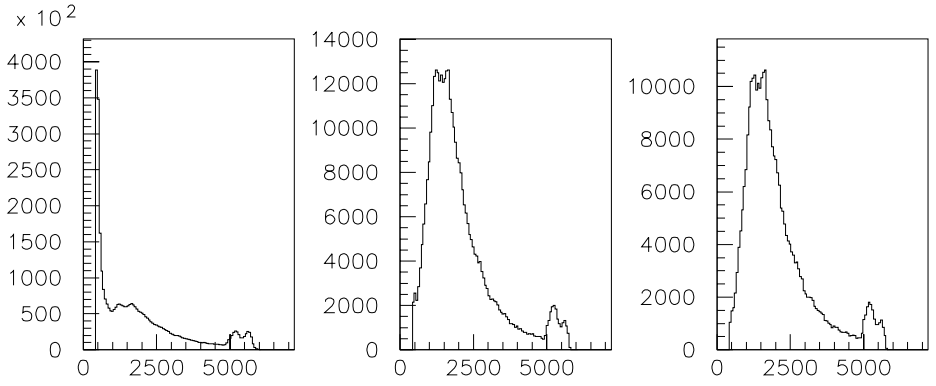

Figure 6.13: Charge distributions for unmatched hits (left), hits uniquely matched to a single track (middle), and hits used in multiple tracks (right), using the same data set as the previous figure.

only one of the hits used must have charge above 1000 counts. In principle, a track could have two out of three hits with charge below 1000 counts. In practice, it is seen that low-charge hits rarely if ever form a valid track.

6.3.3 Signal hit clusters

Another feature of the CCD system performance that could affect the track reconstruction is the creation of hit clusters. Occasionally, a particle will produce a hit in more than one pixel, either by striking right on the boundary between two pixels, or through showering, or through some other process. Additionally, occasionally a low-energy "knock-on" electron inside the CCD will have a direction of motion parallel to the plane of the CCD, resulting in a long, linear pattern of hits in a single CCD, as can be clearly seen in the raw hit figures above.

In analyzing low-rate data, it is often supposed that one should collect all hits within a certain distance of each other into a single logical point, and fit tracks to those points. This introduces dead space around each track, which is generally acceptable for low-rate data, but is unacceptable when processing the high-flux data of this experiment.

In order to properly include the multiple-hit cluster effect in the Monte-Carlo simulation discussed below, it is necessary to assess how often a single particle creates multiple hits in a CCD and to have some idea of the distribution of hits in the multiple-hit clusters. For a first estimate of the probability of multiple hits, it is sufficient to count the total number of hits near fitted tracks and the number of such hits which also have one or more immediately adjacent, lower charge[3] neighbor hit(s), and take the ratio: this ratio is found to be 0.3. Note that this agrees very well with the ratio of the number of raw hits seen to number expected, which is approximately 1.3, as plotted in Figs. 6.10 and 6.11.

To study the distribution of these neighboring hits, the "point spread function" of multiple-hit clusters used in fitted tracks is studied. The point spread function shows the distribution of hits around the track's hit: the bin at the (0,0) position just counts the total number of tracks, the bin at the (1,0) position counts the number of coincident hits offset by one unit along the $+\hat{x}$ direction, etc. Fig. 6.14 shows the point spread function summed over the six rear planes, using data from the representative high-flux run 15296. Only the pixels immediately adjacent to tracks show any excess over the uniform level expected from the overall number of hits.

6.3.4 Hit efficiency and flux-correlated background

The final aspect of CCD system performance to be considered is hit detection efficiency. Since the photon spectrum is dominated by the well-known Compton spectrum, and since pair production is also a well-known process, the number of particles intercepting a plane is easily calculated for a given total number of photons, converter foil thickness, and magnet setting. For the case where the magnet is set to 247.5 GeV/c transverse kick and the foil thickness is 5.6×10^{-4} radiation lengths, the calculation yields 4.0×10^{-5} tracks per photon for a low-momentum cut-off of 19 GeV/c, and 4.7×10^{-5} tracks per photon for a low-momentum cut-off of 18 GeV/c.

Fig. 6.15 shows the average number of hits seen in each CCD plane as a function of the total number of particles expected, using data from run 15296, where the total number of photons is estimated using the Cherenkov monitors EC31 and EC37. The data are fitted to a line in each case.

[3] Counting only the hits whose charge exceeds that of its neighbors avoids over-counting the number of multi-hit clusters.

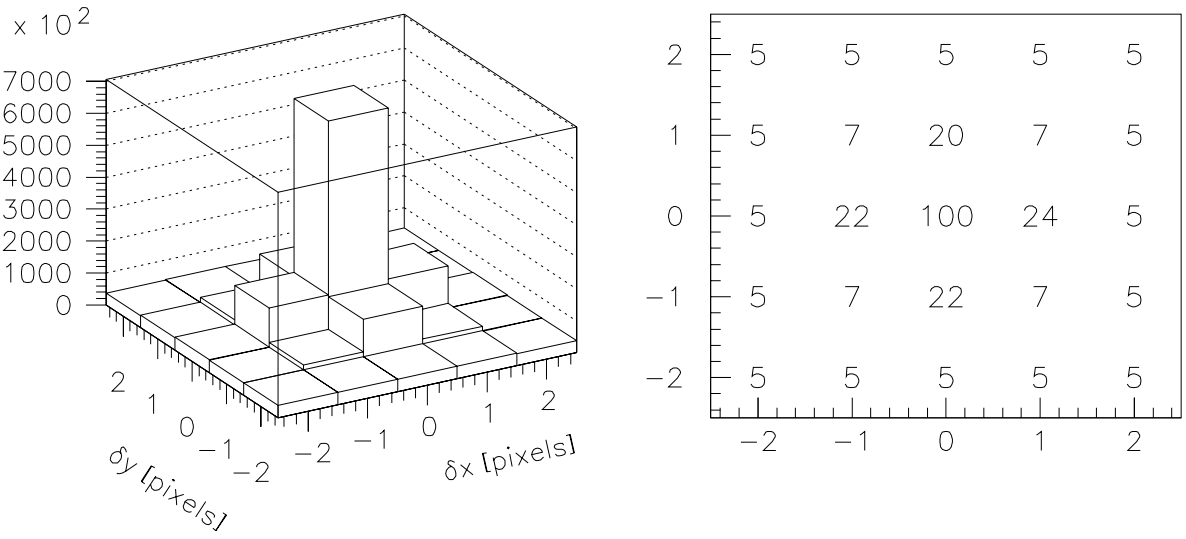

Figure 6.14: Point spread function summed over the six rear planes (see text). A Lego-style plot of raw hit counts is shown at left, and the values as a percentage of central peak are tabulated at right. Data is from run 15296.

The front planes' hit buffers saturated on the highest flux events during this run, but this is not a problem as only the back planes are used in the track reconstruction for high-flux data. The linear relationship between number of tracks expected and number of hits observed holds quite well, but the magnitude of the slope is quite interesting: in the back planes, approximately 3 hits in each plane are seen per particle expected, and 8 to 11 hits per particle expected are seen in the front plane. This is in contrast with the previous finding that the absolute scale of the hit distributions in the signal portion of the CCD, as plotted in Figs. 6.10 and 6.11, is consistent with the number of particles expected and a 30% probability of adjacent "neighbor hits" accompanying a given track. The explanation is that there is a uniform distribution of additional "noise" hits whose number is proportional to the number of "signal" hits. Supporting this explanation, Fig. 6.16 plots the number of hits seen in each CCD plane as a function of the number of hits seen in the "signal" region selected for Figs. 6.10 and 6.11, *i.e.*, a band 99 pixels tall centered around $y_{\mathrm{LAB}} = 0$. The total number of hits above the no-signal background is approximately twice the number of hits in the 99-pixel-tall signal band in the back planes.

In Fig. 6.17, the detection efficiency and background level are plotted for each run in the experiment, as judged by linear fits of the number of hits in each CCD to the number of photons produced at IP1. Clearly, there were excursions in the system efficiency, but the efficiency was generally good and not wildly unstable. One exceptionally unstable period was run 15280, during which the converter fork was moved starting approximately at event 800, and the electron beam rate was changed from 10 Hz to 30 Hz at event 1600. Data from this run is plotted versus event number in Fig. 6.18; the behavior of all planes is similar, but for clarity, only plane 7 is plotted. Increasing the beam rate only increases the level of the signal-uncorrelated background, while changes in the

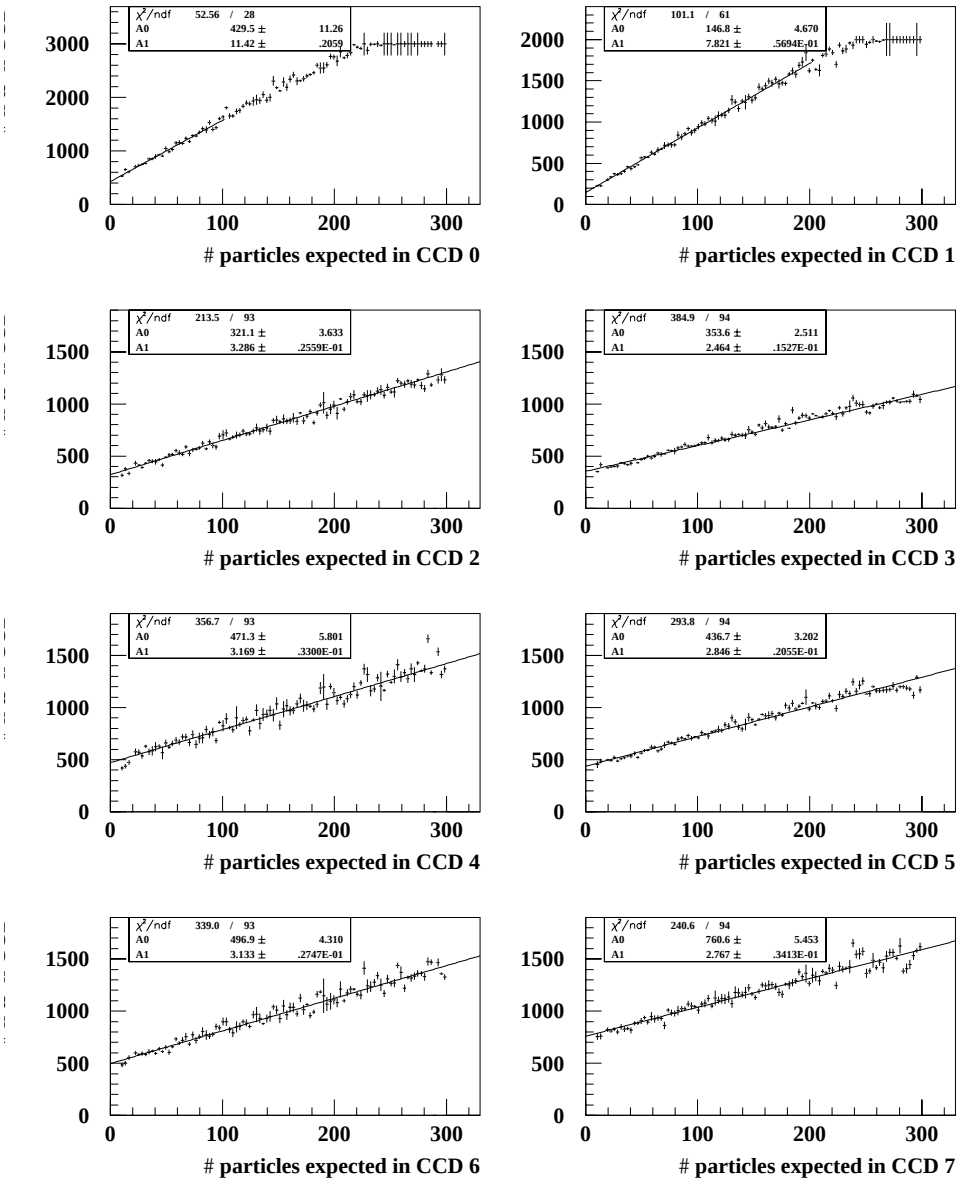

Figure 6.15: Average number of hits seen in each CCD plane as a function of the total number of particles expected, using data from run 15296.

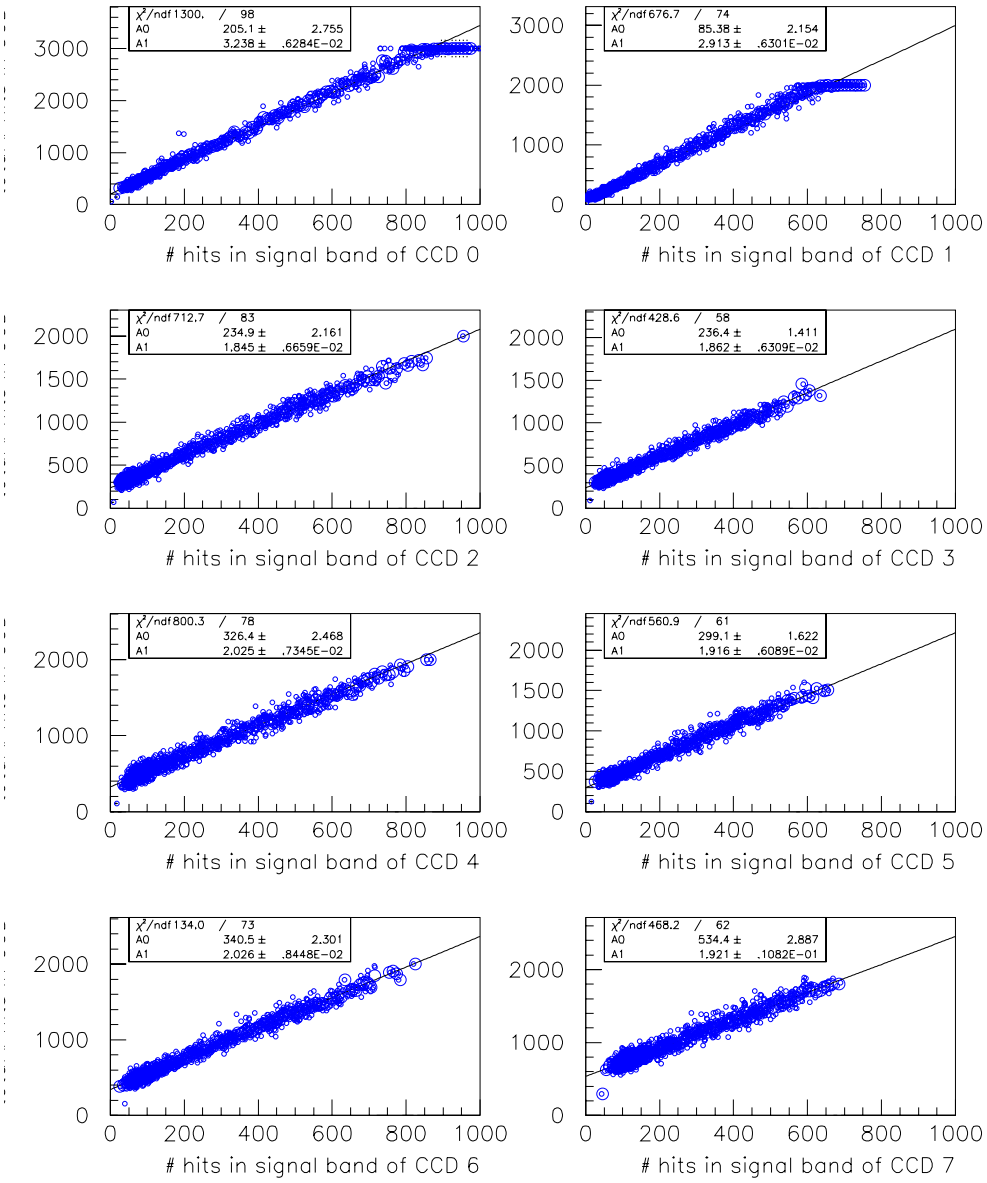

Figure 6.16: Number of hits seen in each CCD plane as a function of the number of hits seen in the region within ±50 pixels of the signal band, using data from run 15296.

converter fork thickness obviously affect the signal rate.

6.4 Monte Carlo simulation of the CCD tracking spectrometer system

The photon production, pair conversion, hit detection, and track reconstruction are simulated using straightforward Monte Carlo techniques. Multiple simulated events are generated for specified, fixed collision parameters. For each event, the following steps are performed:

1. The photon spectrum corresponding to the given interaction parameters is obtained from the look-up tables[26].

2. A list of photons to be converted is generated, distributed randomly according to the photon spectrum. The mean number of photons in the list is equal to the total number of photons in the spectrum times the thickness of the converter in conversion lengths.

3. The Compton scattering angle for each photon is calculated based on its momentum, with azimuthal angle chosen randomly.

4. An additional random contribution to each photon's horizontal and vertical angles is introduced according to the electron beam emittance.

5. The resulting position of each photon at the converter foil is calculated.

6. Each photon is split into an electron and a positron, with momentum distributed according to the Bethe-Heitler differential cross-section. The appropriate angular "kick" is introduced.

7. The positions of the charged particles at each CCD plane are calculated.

8. Track intercept positions are converted to CCD pixel coordinates, and charge depositions are calculated according to a semi-empirical model. Each CCD has a hit list in which the x-position/y-position/charge triplets are accumulated. In roughly 30% of the hits, a direction of motion in the plane is randomly chosen and neighbor hit is added one pixel offset in that direction; 30% of those neighbor hits themselves have neighbors offset along the same direction of motion, leading to streak-like features in a small fraction of events.

9. Background hits are added. The simulated background consists of 200 isolated hits with charges distributed according to the empirical stochastic noise hit charge distribution, plus an extra hit with the same properties as a "signal" hit for every track hit added in the previous step. (cf Fig. 6.16) These background hits are uniformly distributed in the CCD plane.

10. Finally, the hit lists are sorted, hits on the same pixel are combined, and the results written out in the same format as used by the data acquisition system. The resulting data file is later analyzed by the same reconstruction code that is used for the real data [Chapter 7].

The hits in a single simulated event are shown in Fig. 6.19. This data was simulated for $\eta_{\rm foc} = 0.30$, with all other focal parameters at the standard values specified in the GLC4 lookup table [Chapter 5]. Note that the vertical distribution of tracks is much tighter in the Monte Carlo data than in the real data (Fig. 6.9). The Monte Carlo includes all known contributions to vertical

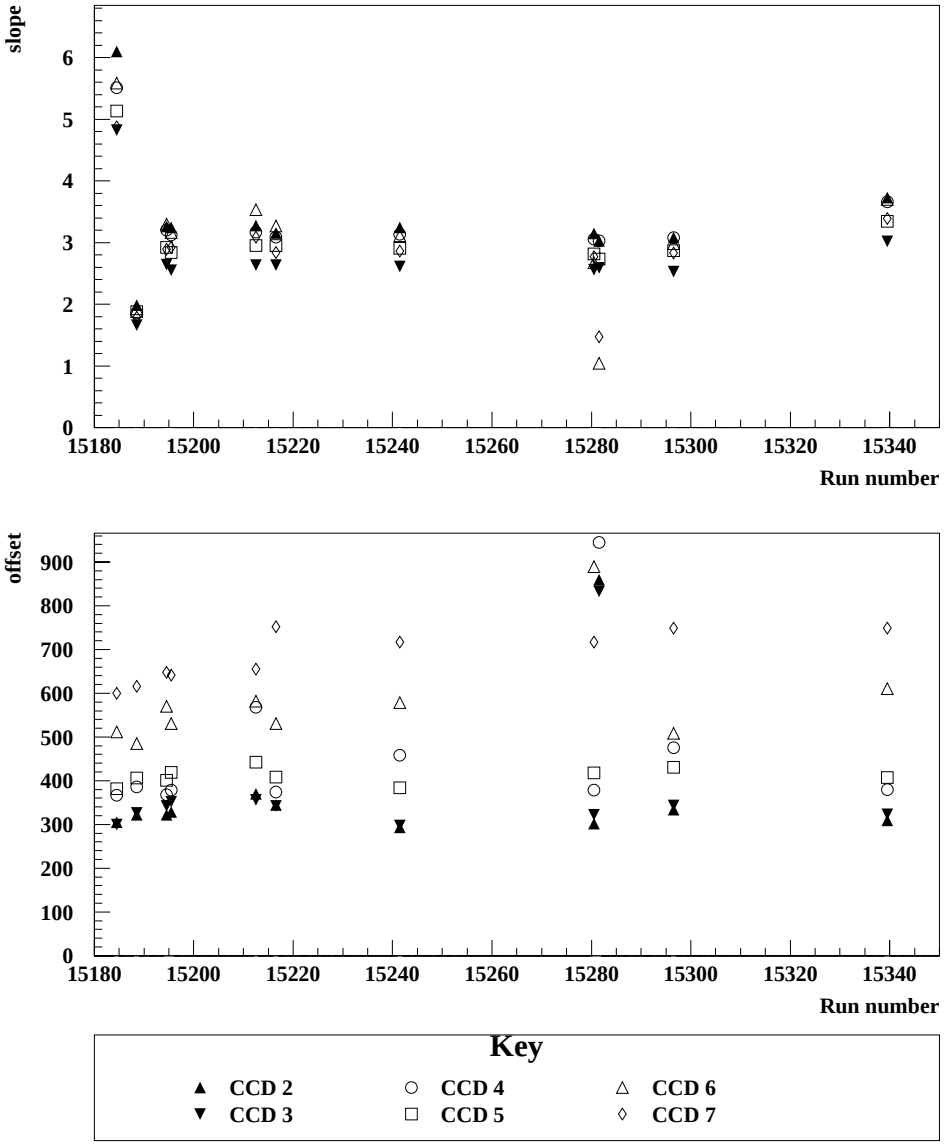

Figure 6.17: CCD performance versus time, as judged by linear fits of the number of hits in each CCD to the number of photons produced at IP1. Here, "time" has been expressed in terms of run number, and the slope of the linear fit has been normalized in the same manner as in Fig. 6.15.

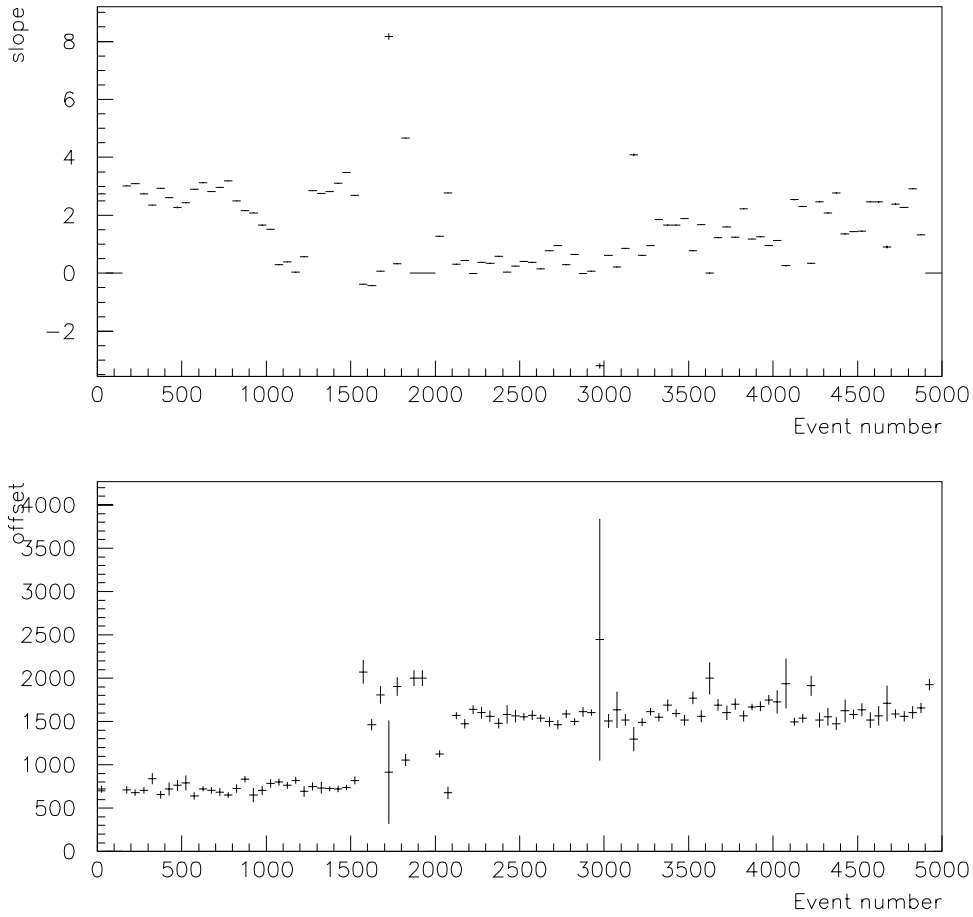

Figure 6.18: CCD performance versus time during run 15280. CCD plane 7 only. See text.

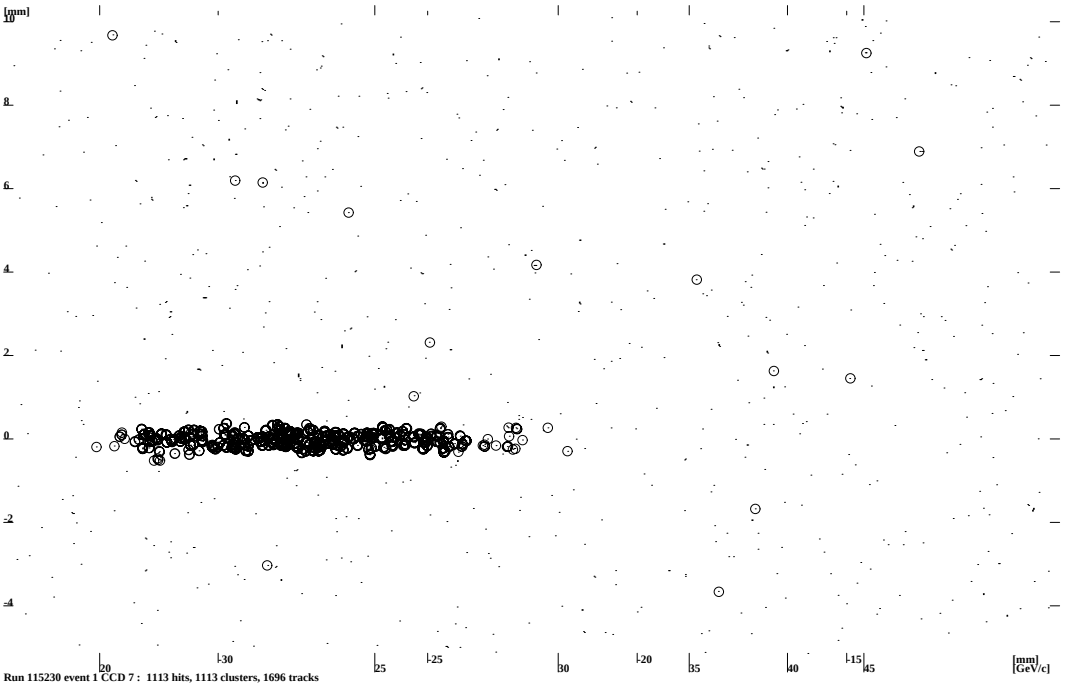

Figure 6.19: Hits in plane 7 for a single event in a simulated data set with $\eta_{\rm foc} = 0.30$.

dispersion of the tracks, and only one of these contributions relies on an experiment-specific measurement: the electron beam angular divergence at IP1, calculated from the upstream emittance and the dispersion at IP1 as measured by various standard beam optics techniques. It is likely the electron beam vertical angular divergence was larger than calculated. However, tighter clustering of the tracks only exaggerates the effect of high flux, and since a primary purpose of the CCD system simulation is to test the performance of the reconstruction algorithm at high track density, the vertical angular divergence was left as inferred from the beam diagnostics.

6.5 Conclusions

The CCD tracking spectrometer system performed largely in accordance with design expectations. The horizontal distribution of hits within the CCD planes is as expected within the range of interest; an unexpected low-momentum cut-off does not appear to be of concern. The vertical distribution of hits is wider than predicted, which can only improve the performance of the track reconstruction algorithm. A uniformly-distributed background correlated with the electron beam rate is seen, and an additional uniformly-distributed background correlated with the photon flux from IP1 is also seen. With the exception of the wider-than-expected vertical distribution, all of these effects, plus the observed charge distributions and cluster formation probabilities, are incorporated into the Monte Carlo simulation of the CCD system.

Chapter 7

CCD Track Reconstruction for E-144: Assuring Uniform Efficiency for High and Low Track Density

Synopsis

The problem of track reconstruction for E-144's pair spectrometer is unremarkable except in that the outer portion of each CCD plane has a much higher hit density than the inner portion, due to the large dynamic range between orders of the spectrum being studied. The physics result would be adversely affected if the efficiency of track reconstruction and fake rejection differed significantly between the two regions. Therefore, it is desirable that the track reconstruction efficiency should be well-matched in the two regions, or failing that, that the efficiencies should be well-known. Here, the reconstruction algorithm is discussed in detail, with particular attention paid to the issue of systematic variations in performance as a function of track density. Results of a Monte-Carlo study are presented and compared to conditions encountered in the August 1996 data. It is concluded that using raw hits with no "clustering" and allowing no more than one hit to be shared between two tracks provides the best performance at all encountered track densities.

7.1 Introduction

The bulk of the useful CCD pair spectrometer data comes from August 1996 positron search runs, in which 10^6 to 10^7 high energy photons were produced on each event. About 25,000 laser "shots" were sampled with the CCD systems active, using a $5.6 \times 10^{-4} X_0$ aluminum foil for the converter for roughly 2,300 of the shots, a thinner nitrocellulose foil for another 5,100 shots, and the remainder of the data devoted to using (or trying to locate) various much-thinner wires. So far, the most usable data has proved to be the aluminum foil data.

With 10^7 high energy photons incident on a $5.6 \times 10^{-4} X_0$ target, one naturally gets a large number of pairs produced: $(7/9) \times 10^7 \times 5.6 \times 10^{-4} \cong 4,400$ pairs per event, to be specific. Most of these are produced by photons in the linear ($n = 1$) portion of the Compton spectrum. This high number of tracks makes identification of a given track's oppositely charged counterpart practically impossible if either track falls in the densely populated portion of the spectrum below the $n = 1$ kinematic edge. Furthermore, all pairs have at least one component below the $n = 1$ kinematic edge, because the beam energy at 46.6 GeV is less than half the $n = 1$ kinematic edge at 29.0 GeV.

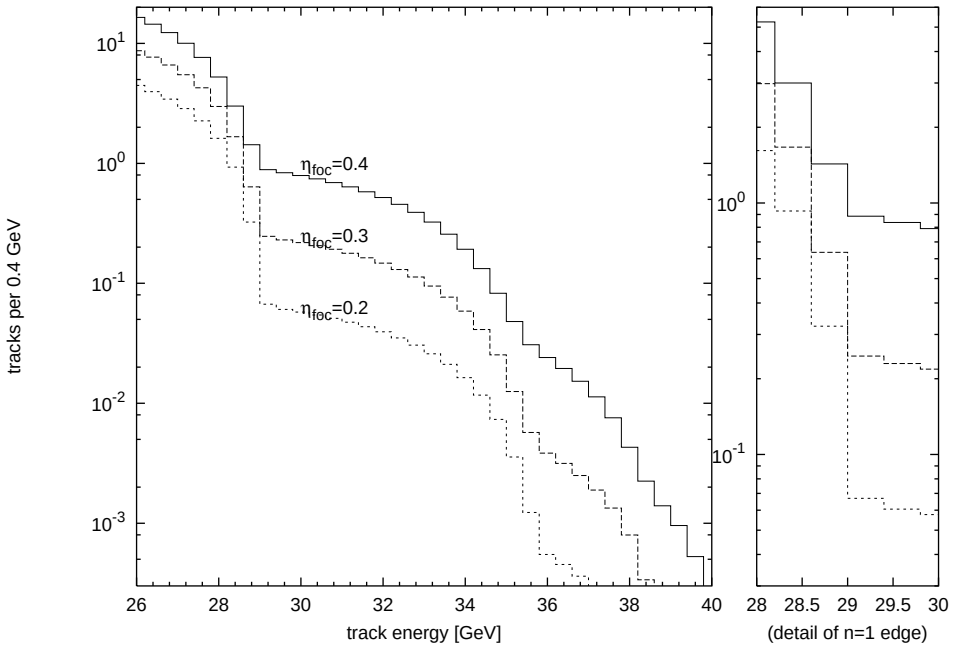

Figure 7.1: Simulated "single-arm" track spectrum for a single event using August 1996 ("GLC4") standard interaction parameters. A "close-up" view of the n=1 edge is provided at the right.

For this reason, we analyze the two arms of the spectrometer independently, comparing the track spectrum to a theoretical spectrum calculated by convolving the simulated photon spectrum with the Bethe-Heitler pair spectrum.

Figure 7.1 shows the calculated "single-arm" track spectrum for a single event using August 1996 standard interaction parameters, those used in the "GLC4" lookup table[26, Chapter 5]. At $\eta_{\rm foc} = 0.33$, only about 50 tracks per event have momenta over 26 GeV/c, and only 3% of those tracks have momenta exceeding 30 GeV/c. With only 50 tracks per event expected in the region of interest, one might wonder if there is really any significant challenge in the track reconstruction. In fact, there is, because the photons are near the endpoint of the Compton spectrum and therefore have small scattering angle, and the linear-collider-style "flat" beam has low vertical emittance, leading to a relatively high track density in the CCDs.

The need for high track detection efficiency in the presence of high density motivates the decision to avoid combining hits into clusters. Because the tracks are so parallel, allowing or disallowing shared hits does not directly affect the efficiency of "real" track reconstruction, but the number of "fake" track candidates formed from improper combinations of "real" hits varies non-linearly with the track density, and the choice of number of shared hits to allow is critical in controlling the effect of combinatoric fakes.

The first section below is a brief outline of the track-finding algorithm. The next section is dedicated to a quantitative discussion of the track reconstruction challenges implicit in the theoretical spectrum plotted in Fig. 7.1. Results of Monte Carlo studies are presented and compared to conditions encountered in the August 1996 data, with particular attention paid to the issue of systematic variations in performance as a function of track density. It is concluded that using raw hits with no "clustering" and allowing no more than one hit to be shared between two tracks

provides the best performance at all encountered track densities in the Monte Carlo simulation, and that all indications are this should be true for the "real data" as well.

7.2 Track reconstruction algorithm

The original CCD data acquisition code and standard track reconstruction code for E-144 was written by Eric Prebys. For the most part, this code has remained unchanged, with only slight modifications made by this author. In brief, the data acquisition system scans each digitized frame from the CCD and constructs a list containing the coordinates of each pixel whose signal exceeds a set threshold above its pedestal reading; each such entry is called a "hit." This list of hits is written to disk and can be analyzed and re-analyzed off line as desired. The task of the track reconstruction software is to find the set of tracks that best matches the set of hits recorded for each event.

Given suitable alignment data (as described below), the track reconstruction proceeds as follows:

1. Optionally, gather hits which are close together into "clusters."
2. Convert CCD coordinate positions of clusters (or hits) to "points" in a common three dimensional coordinate system.
3. Construct track candidates using the hits in the three back planes. The exact procedure followed depends on the number of allowed shared points. Track candidates are subjected a χ^2 cut and/or cut on the maximum distance of the fitted track from its points. Optimizations of the implied loops over points are made by exploiting a known ordering of the point lists and the known cuts.
4. Optionally, attempt to construct four-hit track candidates by associating hits in the front CCD with three-hit tracks already found.
5. Optionally, cull the list of track candidates to eliminate unacceptable sharing of points between tracks. (This culling is always performed except when performing special tests to debug the algorithm.)
6. For each final track, calculate the "available nearby pixel ratio" by examining a box 5 pixels high and 15 pixels wide centered around each of the track's rear hits, counting the number of un-hit pixels outside of any hit pixel's clustering radius, and taking the ratio of the number of such pixels to the maximum number of such pixels possible in the limit of zero noise and infinitesimal flux. This "available nearby pixel ratio" is useful in studying the effect of hit density on track detection efficiency.

Track data is written to a data summary file in n-tuple format, along with event and run number data that allows association of track information with other E-144 detectors.

7.3 Momentum calculation and depixelization

The momentum of a track can be calculated from the horizontal angle of the track θ_x and the transverse kick of the magnet p_{kick}:

$$p_{\text{track}}(\theta_x) = \frac{p_{\text{kick}}}{|\theta_x|}. \tag{7.1}$$

In single-arm mode, only the rear three CCDs of each arm of the spectrometer are used. The 22.5-μm pixel size and 100-mm separation of the planes results in spikes in the θ_x and $p_{\rm track}(\theta_x)$ distributions. It would be possible to remove these spikes by randomly shifting "hit" positions ± 0.5 pixels prior to track fitting, but this would degrade the momentum resolution.

It is also possible to calculate momentum from a track's fitted position in the rear CCDs, 2.5 m downstream of the magnet's centerplane. The effect of the 22.5-μm pixel size over a 2.5-m lever arm is an order of magnitude smaller than the effect on the track angle data. Given the average position of all tracks at the magnet's centerplane $\bar{x}_{\rm MAG0}$, and a given track's position $x_{\rm CCD}$ at a distance $z_{\rm CCD\text{-}MAG} \cong 2.5$ m downstream, the momentum is

$$p_{\rm track}(x_{\rm MAG0}) = \frac{p_{\rm kick} z_{\rm CCD\text{-}MAG}}{|x_{\rm CCD} - \bar{x}_{\rm MAG0}|}. \tag{7.2}$$

This is equivalent to estimating the true track angle using the intercept position in the CCD and assuming the track originates at $\bar{x}_{\rm MAG0}$.

The momentum resolution of the track position method is limited by the beam divergence of 10 μrad, which translates into a position spread of 380 μm (rms) due to the 38 meter separation of IP1 and the CCDs. (Away from the Compton endpoint, the angular divergence of the backscattered photon beam is greater, leading to an even greater rms spread of track origins in the magnet.) The momentum resolution using the position method is actually lower than the resolution of the track angle method.

Note that the dominant sources of error in the two methods are uncorrelated. A weighted average of the results of the two methods can provide the highest resolution for track momentum with no spikes due to pixelization and no additional randomization required:

$$p_{\rm track}(x_{\rm MAG0}, \theta_x) = \frac{p_{\rm kick}}{\left|\frac{3}{4}\theta_x + \frac{1}{4}(x_{\rm CCD} - \bar{x}_{\rm MAG0})/z_{\rm CCD\text{-}MAG}\right|}. \tag{7.3}$$

In essence, the quantity $((x_{\rm CCD} - \bar{x}_{\rm MAG0})/z_{\rm CCD\text{-}MAG} - \theta_x)/4$ is used instead of a software random number generator to "depixelize" the track angles.

After all alignment and track reconstruction was completed, a slight variation with magnet setting was noticed in the apparent momentum of the kinematic edges, as shown in Fig. 7.2. This was interpreted as arising from either an error in the measurement of the magnet saturation curve or an overall systematic rotation of the CCD planes. Figure 7.3 shows the results of applying the Kolmogorov-Smirnov consistency test[50] to track momentum data from two different magnet settings as a function of a systematic shift in the momentum. Clearly, a correction is required.

As discussed previously[Chapter 6], the absolute position and rotation of the two arms in space is not known with as much precision as the relative position of the CCDs within a given arm. The correction was therefore implemented by increasing the magnitude of all track horizontal angles 0.36 mrad, assuming that the source of the variation of edge position with magnet setting described above is due to overall misalignment of the spectrometer arms. Subsequent detailed study of the horizontal distribution of hits within each CCD also supports this interpretation, as discussed previously.

Figure 7.4 shows the consistency of momentum distributions drawn from two arms of the spectrometer at the nominal magnet setting: no angle correction between arms is required.

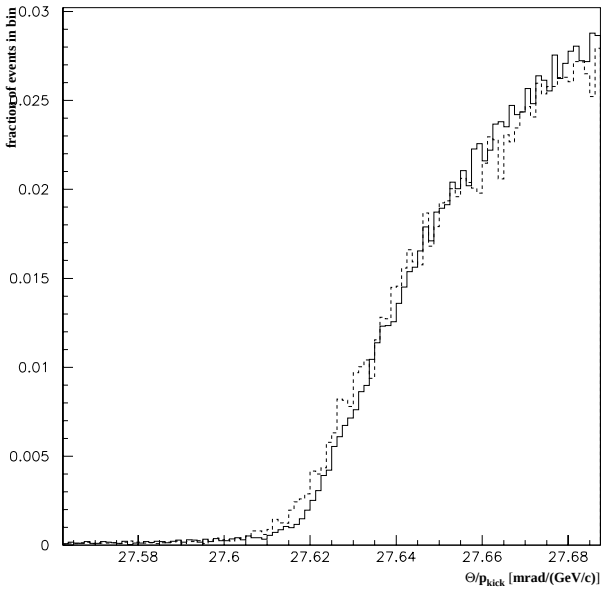

Figure 7.2: Distribution of reconstructed track angle divided by magnet kick, for two different magnet settings: 247.5 MeV/c (solid) and 219 MeV/c (dashed).

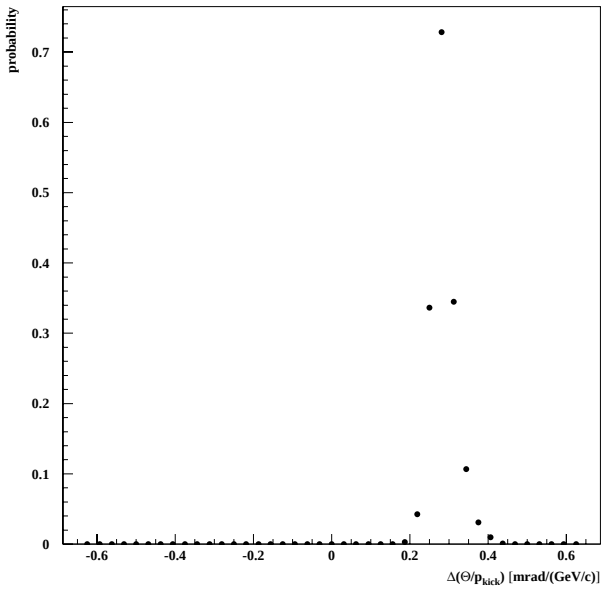

Figure 7.3: Probability of the reconstructed track momenta at the two different magnet settings 219 MeV/c and 247.5 MeV/c being drawn the same distribution, as a function of systematic shift in track angle.

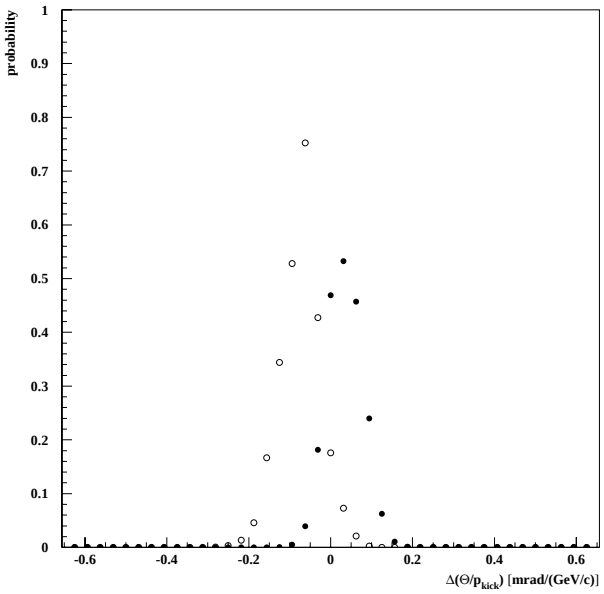

Figure 7.4: Probability of the momentum distributions from different arms being the same, as a function of systematic relative horizontal angle correction. Solid circles are from thick foil data at 247.5 MeV/c kick, open circles are from thick foil data at 219 MeV/c kick.

7.4 Quantitative discussion of a simulated example case

The simulated spectrum includes tracks from either arm of the spectrometer. There are approximately 24 tracks/arm in the 26 to 29 GeV/c momentum range in the example case of Fig. 7.1, with an approximately linear fall to zero density at the kinematic edge, so the track momentum density in a single arm is about 16 tracks/GeV/c near 26 GeV/c. The linear spatial density of hits in a CCD at distance z from the magnet is

$$\frac{dN}{dx} = \left(\frac{dx}{dp}\right)^{-1}\frac{dN}{dp} = \frac{p_{\rm track}^2}{z p_{\rm kick}}\frac{dN}{dp} \tag{7.4}$$

$$\simeq (26\ \text{GeV/c})^2\ (2.4\ \text{m})^{-1}\ (0.2475\ \text{GeV/c})^{-1}\ (16\ \text{tracks/GeV}) \tag{7.5}$$

$$= 18\ \text{tracks/mm},$$

where $p_{\rm kick}$ is the transverse kick of the magnet (usually 247.5 MeV/c) and $z \simeq 2.4$ m for the first CCD of the rear group of three. The CCD pixels are 0.0225 mm across, so this track linear density corresponds to a respectable 0.4 tracks per linear pixel.

The vertical angular divergence of the beam and the angular divergence introduced by the Bethe-Heitler pair creation process expand the track distribution vertically and reduce the track density per square pixel somewhat, but not enough to eliminate all concern regarding track density effects. With a 30 μm (rms) waist at IP1 and a beam vertical geometric emittance less than 3×10^{-11} meter-radians, the divergence was under 1 μrad. Given a 38 m separation between IP1 and the CCD's, the contribution to the vertical position dispersion from the emittance was less

than 38 μm, which is less than 2 pixel sizes. The contribution from the Bethe-Heitler kick at the converter foil is more significant: $\theta_{\text{B-H}} \simeq mc/p = 0.511/26000 \simeq 20\mu$rad, over 3.4 m, gives a 67 μm contribution to the rms vertical spread. Adding the emittance and Bethe-Heitler contributions in quadrature gives 77 μm. In the data actually acquired, the minimum vertical rms spread of the tracks in the CCDs is about twice this number. Using the observed number, the density per unit area is half as high as otherwise expected, but it is still high:

$$\frac{dN}{dA} = \frac{1}{\sqrt{2\pi}\sigma_y}\frac{dN}{dx} \tag{7.6}$$

$$\simeq 0.3989\ (150\ \mu\text{m})^{-1}\ (22.5\ \mu\text{m/pixel})\ (0.4\ \text{tracks/linear pixel}) \tag{7.7}$$

$$= 0.024\ \text{tracks/square pixel.}$$

This track density is high enough to motivate the use of single-pixel hits in the track reconstruction, as opposed to combining all hits within an appropriately centered block of m pixels into a single cluster, due to the high probability of cluster sharing if m is much larger than 1. The probability of a cluster centered on a track in a given CCD being hit by one or more other tracks is $1-(1-p)^m$, where p is the density of hits per square pixel. In the example case, $p \simeq 0.024$, so the probability of cluster sharing is 45% for a 5x5 block, 20% for a 3x3 block, but only 2% for a single pixel. Furthermore, the total angular divergence of the tracks multiplied by the interplane spacing of the back CCDs is less than one pixel size, so that any two "real" tracks sharing a cluster in one plane will share clusters in all three planes.

The decision whether to allow sharing of hits, and the determination of how many shared hits to allow per track, is determined by the effects of "fake" combinations of tracks. The number of "fakes" varies non-linearly with the track density. This effect can not be easily quantified using a single example, but it can be described. If all possible combinations of tracks are allowed, too many fakes get into the passband of the final signal cuts; however, disallowing all shared hits may have the effect of reducing the reconstruction efficiency for real tracks, since it may happen that a "fake" that would be rejected in the final analysis has a lower χ^2 than the track actually associated with a hit, due to the pixelization of the CCDs. The Monte Carlo simulation of the CCD system, described in the next section, is the best tool for analyzing this issue quantitatively.

7.5 Features of the combinatoric background

Fig. 7.5 shows the observed distribution of the positions of tracks at the center plane of the magnet. This distribution is well described by a sum of two gaussians: a narrow gaussian with rms ~ 0.6 mm and a wider gaussian with rms ~ 6 mm. The narrow gaussian width is consistent with the expected track angle resolution of the system; the wide gaussian width is consistent with the expected distribution of combinatoric backgrounds.

Figure 7.6 shows the geometry of combinatoric "fake" tracks constructed from "real" hits. Three nearby tracks create hits which can be combined in multiple ways. Most incorrect combinations do not point back into the origin of "true" tracks, unless they share two hits with a "true" track, in which case one track is accepted and one rejected as described above.

Figures 7.7 and 7.8 show the distribution of magnet plane intercept positions for shallow angle (high momentum) tracks. They are also well described by two gaussians, but in this case, the combinatoric background is displaced away from the center position of the magnet. This is consistent with the idea that the combinatoric backgrounds are dominated by hits intercepting the

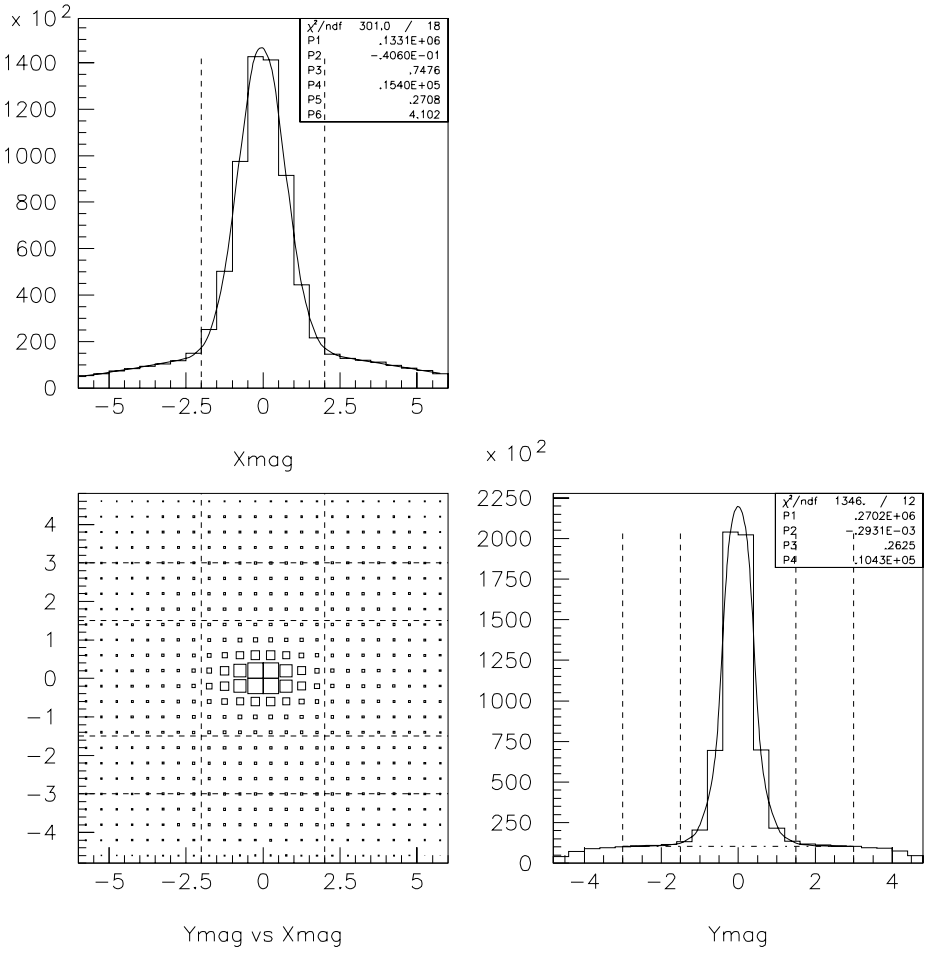

Figure 7.5: The two-dimensional distribution of the track positions at the center plane of the magnet. Box size is proportional to number of tracks in the bin.

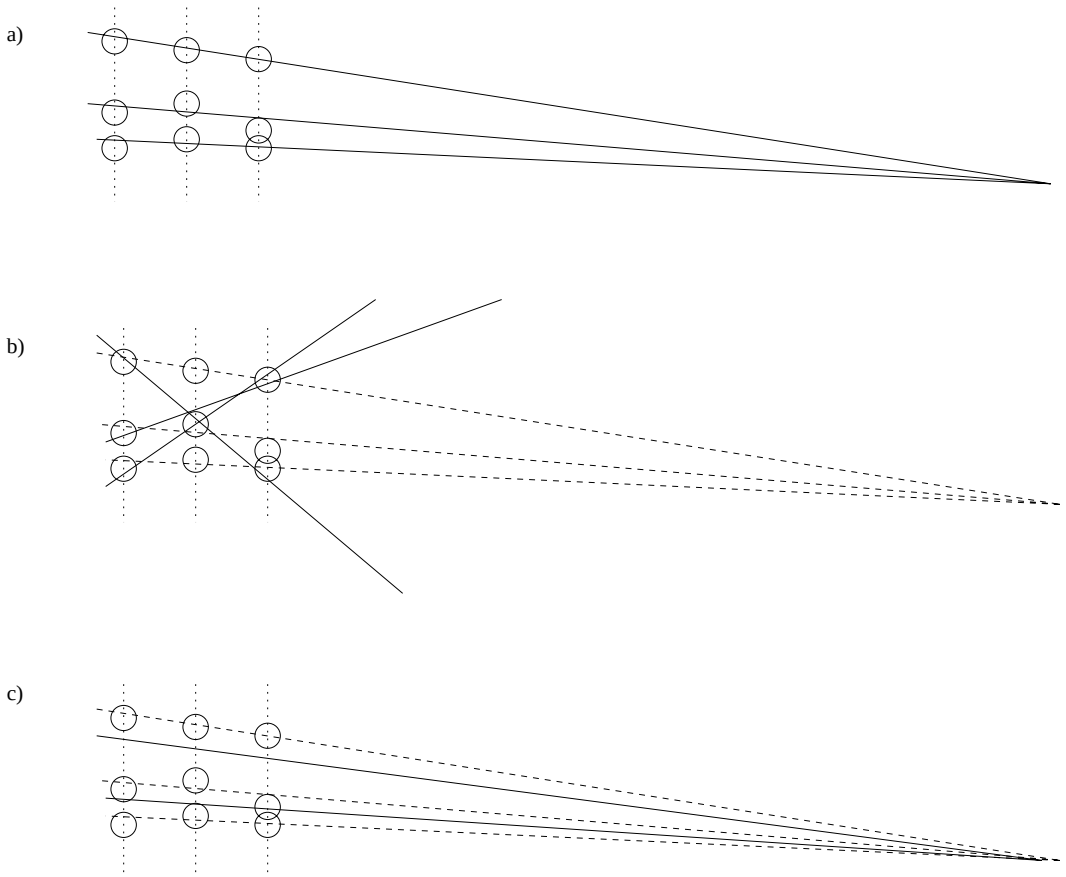

Figure 7.6: Hypothetical examples of "fake" (combinatoric background) tracks constructed from hits from "real" tracks. Three real tracks and their associated hits are shown in (a). In (b) are shown examples of combinatoric tracks contributing to the wide gaussian distribution in the XMAG/YMAG plane; two of the combinatoric "fakes" share no more than one hit with a lower-chi-squared track. Examples of "fake" tracks pointing into the narrow "signal" peak in the XMAG/YMAG plane are shown in (c), all of which share two hits with another track.

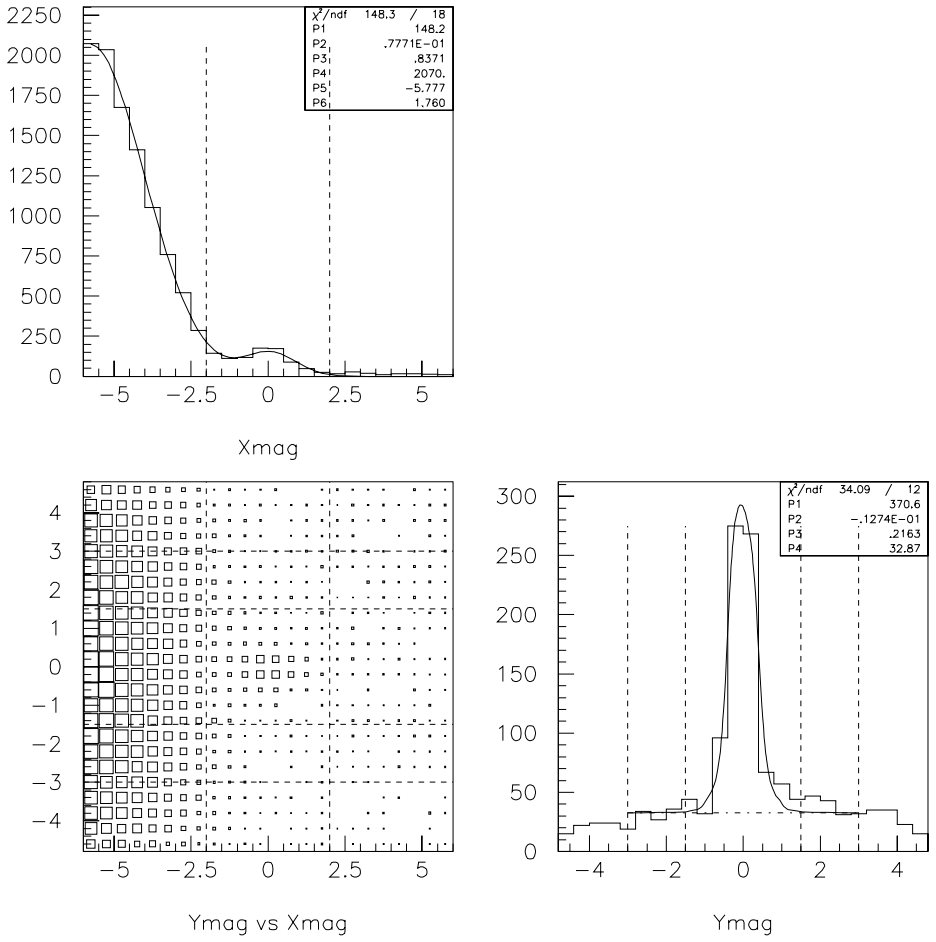

Figure 7.7: The distribution of the track positions at the center plane of the magnet, using only electron tracks with shallow angles (high momentum).

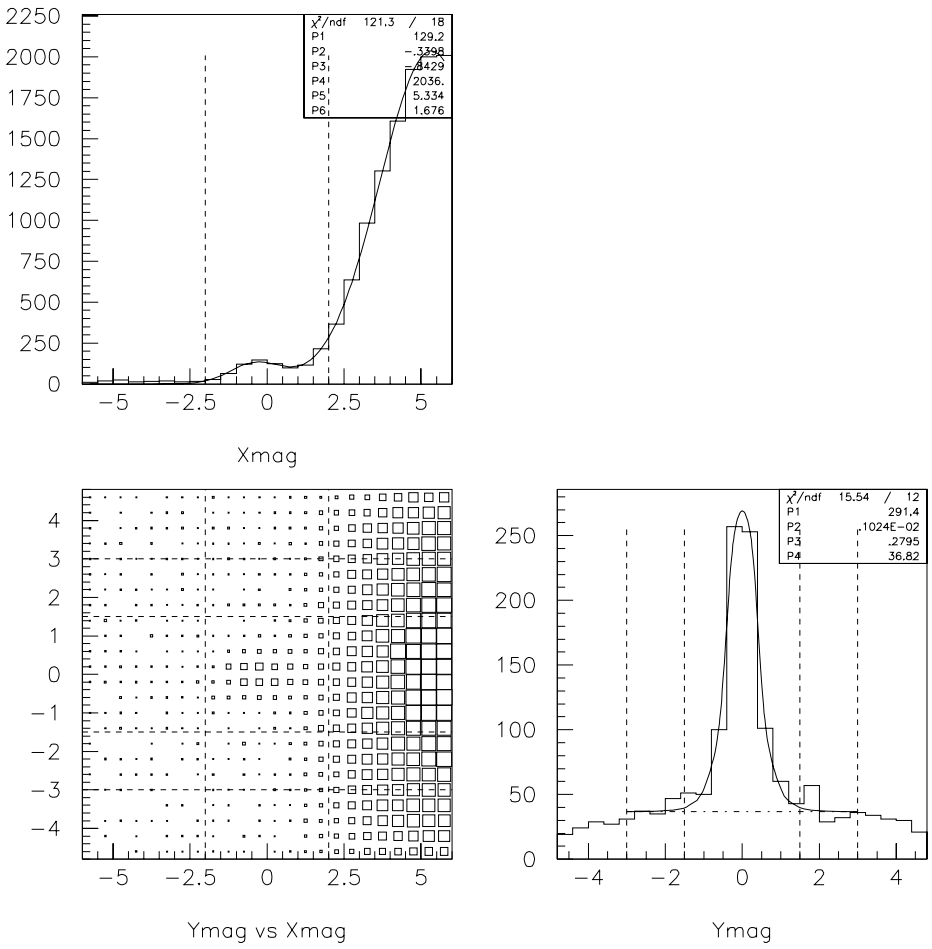

Figure 7.8: The distribution of the shallow-angle positron track positions at the center plane of the magnet.

most densely hit part of the CCDs, which in this case is the part of the CCDs intercepted by the particles with energies less than the Compton kinematic edge. Since only tracks with shallow angle are shown in the Figure, the "fakes" produced by hits in the CCD furthest from the beamline naturally tend to project back to a position away from the center of the magnet. This is fortunate since the number of "real" tracks at high momentum is so small.

In order to limit the number of "fake" tracks contaminating the signal, the analysis was confined to tracks whose horizontal position at the magnet centerplane, $x_{\rm mag}$, was within 2 mm of the centroid. The "signal" was taken to be the number of tracks whose vertical position at the magnet center plane, $y_{\rm mag}$, was within 1.5 mm of the centroid, and the number of tracks with $y_{\rm mag}$ between 1.5 mm and 3.0 mm of the centroid was used as an estimate of the background.

The size of the combinatoric background's gaussian distribution is significantly greater than the dimensions of the rectangular "signal" and "background" regions described above. For this reason, the number of combinatoric background tracks in the signal and background regions are about the same, and that number is proportional to $\exp(-(x_o(p)/\sigma)^2/2)$, where $x_o(p)$ is the offset of the background distribution for tracks with momentum p. This offset is expected to be similar to the difference between the position at the CCDs where an actual track of that momentum would be found and the position of the densely-populated region where "fakes" are generated. Transforming variables to be entirely in terms of momentum, and allowing for a background uniform in vertex position to account for the relatively small contribution of thermal noise and synchrotron light, the following form for the backgrounds is obtained:

$$\frac{dN_{\rm bg}}{dp} = Ae^{-(p-p_0)^2/\sigma_{p,\rm bg}^2/2} + B\left(\frac{p_0}{p}\right)^2, \tag{7.8}$$

where p is the momentum, and A, B, p_0, and $\sigma_{p,\rm bg}$ are all constants which are best empirically determined.

7.6 Results of track reconstruction of Monte Carlo data

The Monte Carlo simulation of the CCD system has been previously described [Chapter 6]. Application of the reconstruction procedure to the Monte Carlo data sets produces combinatoric backgrounds similar to those found in the real data. Figs. 7.9 and 7.10 show the distribution of track positions projected back the magnet centerplane for all tracks and for high-momentum positron tracks, generated in exactly the same way as Figs. 7.5 and 7.7. The primary difference is the narrower width of the $y_{\rm mag}$ distribution, due to the conservatively small angular divergence used in the Monte Carlo simulation. The Monte Carlo data set was generated using the "standard" (GLC4 table) parameters at $\eta_{\rm foc} = 0.31$. All events in this Monte Carlo data set had 7.5×10^6 backscattered photons.

The "available nearby pixel ratio" defined previously is plotted as a function of track momentum in Fig. 7.11, for the same Monte Carlo data set described previously, using two different options for the reconstruction algorithm: at left, no "cluster-finding" was used, and at right, hits were first combined into 3x3 clusters before reconstruction. The decrease in the available nearby pixel ratio reflects the increase in cluster size from 1 to 9 pixels.

Also plotted in Fig. 7.11 is the true efficiency of the reconstruction as a function of momentum for this data set, where the true efficiency is the ratio of the sideband-subtracted number of reconstructed tracks to the actual number of tracks in the Monte Carlo's input spectrum. Fig. 7.6

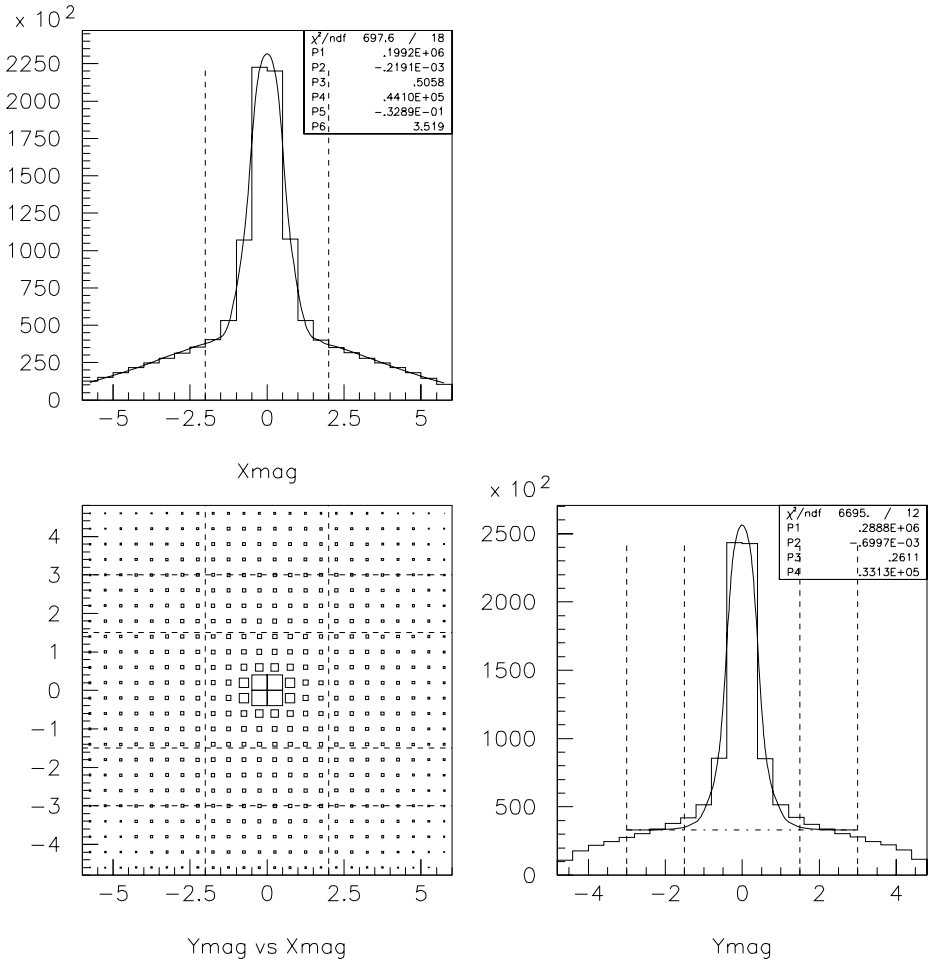

Figure 7.9: The two-dimensional distribution of the track positions at the center plane of the magnet, for Monte Carlo data.

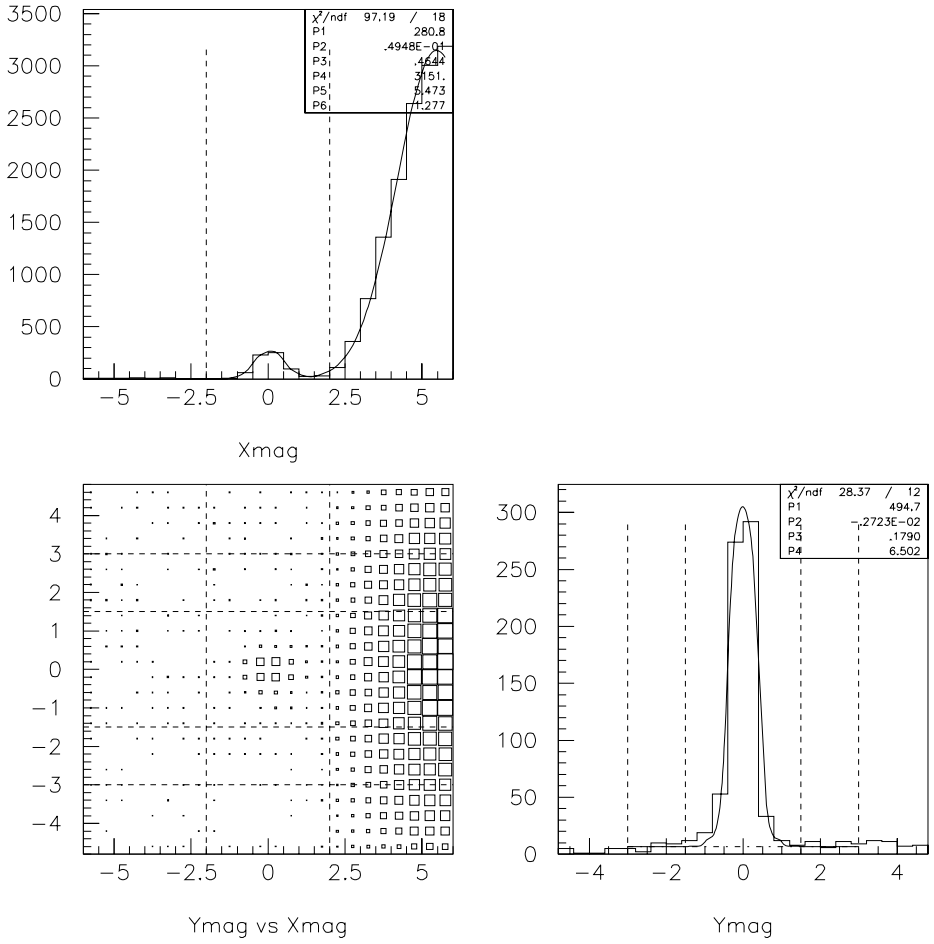

Figure 7.10: The two-dimensional distribution of the track positions at the center plane of the magnet, for Monte Carlo data, after selecting high momentum positron tracks.

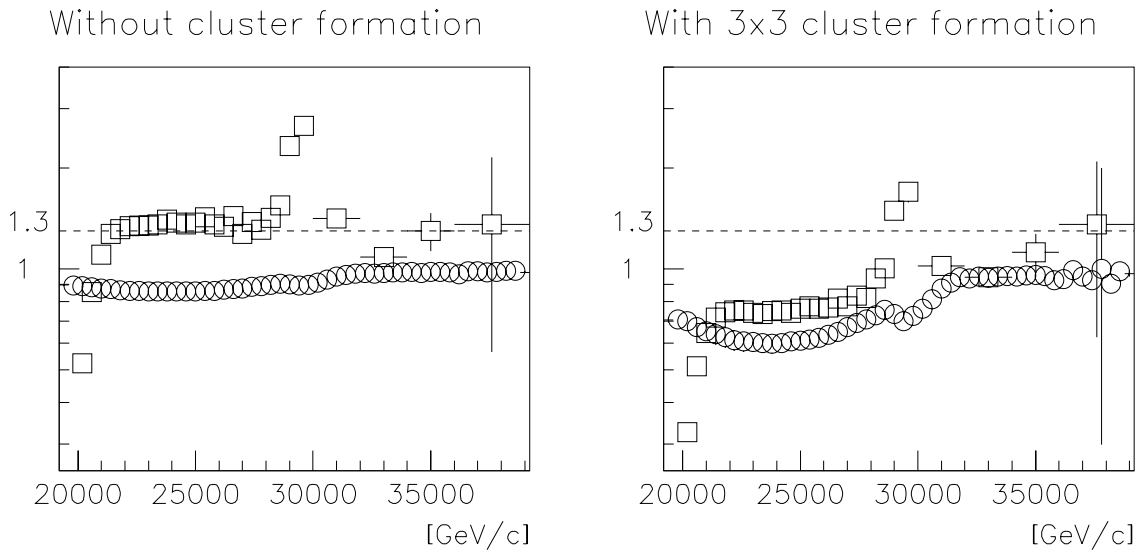

Figure 7.11: Available nearby pixel ratio (circles), and true track reconstruction efficiency (squares), as a function of track momentum, for Monte Carlo data. For the results plotted at left, a reconstruction algorithm with no "cluster-finding" was used, and at right, hits were first combined into 3x3 clusters before reconstruction.

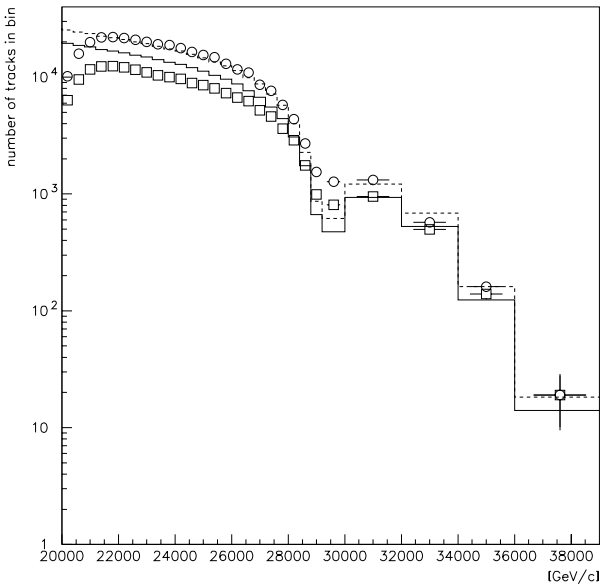

Figure 7.12: Track spectra input to the CCD system Monte Carlo (solid histogram), reconstructed without cluster formation (circles), and reconstructed with formation of 3x3 clusters (squares). The dashed histogram is simply 1.3 times the solid histogram, for reference. This figure was produced from the same Monte Carlo data set used to produce Fig. 7.11.

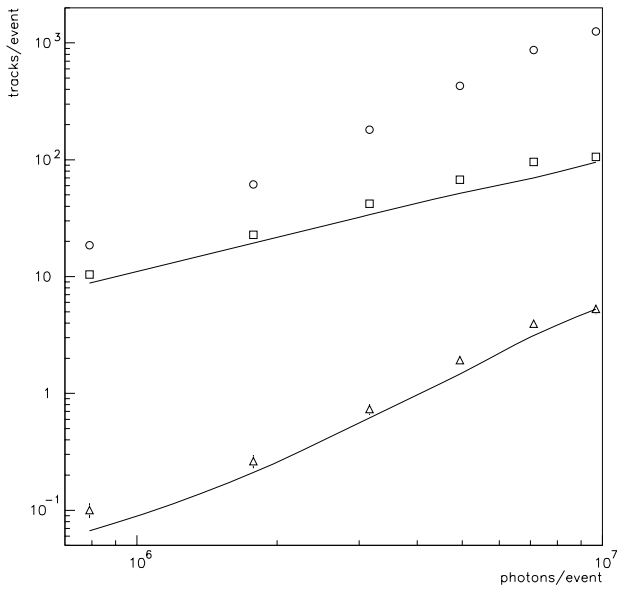

Figure 7.13: Variation with total number of photons of total number of reconstructed tracks (circles), sideband-subtracted number of tracks in an $n = 1$-dominated region (squares), and sideband-subtracted number of tracks in an $n = 2$-dominated region (triangles), for Monte Carlo data. Solid lines plot the input to the Monte Carlo from the theory.

shows the input and reconstructed track spectra. Note that the finite resolution of the spectrometer results in an increase in the number of tracks near the $n = 1$ edge, producing the spike in the "true efficiency" ratio, but the ratio of input to output does provide an accurate efficiency estimate away from the $n = 1$ edge. The results without clustering have an efficiency of 1.3, reflecting the "neighbor hit" phenomenon. Note that the $n = 1$ and $n = 2$ potions of the spectrum are affected equally. Grouping hits into clusters brings the efficiency to 1.0 in the low-density $n = 2$ region, but the efficiency in the $n = 1$ portion is reduced below 0.8. This discrepency in efficiency would be an intensity-dependant effect, seriously compromising the experiment were it to be allowed. No cluster finding has been used in any of the analysis of real data.

Monte Carlo data sets were generated for several different laser intensities. Figure 7.13 shows the variation versus total number of photons of the total number of reconstructed tracks (including combinatoric backgrounds); the number of reconstructed tracks in the $n = 1$-dominated signal region from 26 GeV/c to 29 GeV/c, after sideband subtraction; and the number of reconstructed tracks in the $n = 2$-dominated signal region from 30 GeV/c to 36 GeV/c, after sideband subtraction. Also plotted are the actual number of $n = 1$ and $n = 2$ tracks according to the input spectra given as input to the Monte Carlo. Note that the combinatoric background is large at high flux, and non-linear in the flux, and thus a potential threat to the non-linear Compton analysis. Happily, this threat can be squashed: the inband minus sideband signals are free of contamination.

The ratios of reconstructed to actual tracks are plotted as a function of total number of backscattered photons in Fig. 7.14.

The track reconstruction efficiency is quite good in the Monte Carlo simulation, over all track

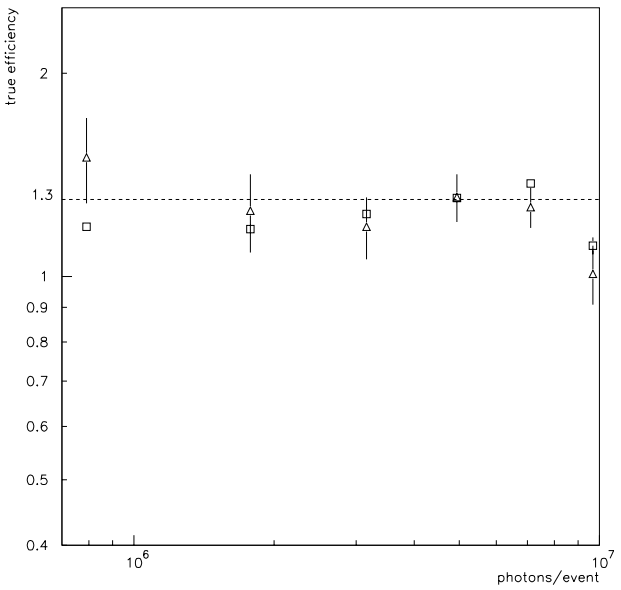

Figure 7.14: True efficiency of track reconstruction as a function of total number of photons, using Monte Carlo data sets as described in the text. Circles show average efficiency for $n = 1$ tracks, and squares show the average efficiency for $n = 2$ tracks, using momentum ranges described in the text.

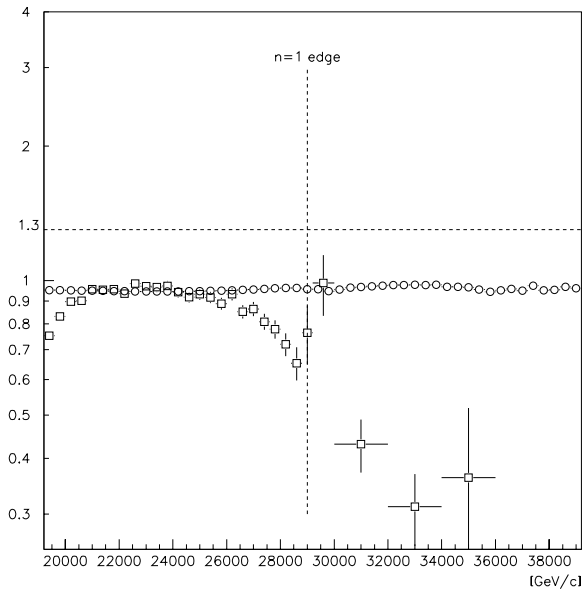

Figure 7.15: Available nearby pixel ratio (circles), and true track reconstruction efficiency (squares), as a function of track momentum, for real data from run 15296. In calculating the "true" efficiency, a reference spectrum with $\eta_{\rm foc}$ has been used, normalized by observed total number of photons; the actual value of η may be different, so no conclusions can be drawn regarding efficiency in regions where the $n = 2$ contribution is significant.

densities and track momenta of interest.

7.7 Track reconstruction performance with real data

The track density is lower in the real data than in the Monte Carlo. This can be seen by plotting the available nearby pixel ratio as a function of track momentum using only events having more than 7×10^6 backscattered photons. Such a plot appears in Fig. 7.15, superimposed on which is the Monte Carlo data plotted in Fig. 7.11. The lowest value of the available nearby pixel ratio is 0.95 in this data, in contrast to 0.90 in the Monte Carlo test. Since the tracks are less dense, the effects of tracks density are expected be less of a problem in the actual experiment than in the Monte Carlo.

Although the true reconstruction efficiency cannot be assessed in the $n = 2$ region of the spectrum without assuming the theory under test, the $n = 1$ Compton spectrum is very well known, and the true reconstruction efficiency can be calculated for that region; this is also plotted in Fig. 7.15, using a theoretical spectrum calculated for $\eta_{\rm foc} = 0.31$ as the reference. The efficiency is indeed near unity throughout most of the $n = 1$ region, while the low number of observed $n = 2$ tracks could simply be due to the actual value of η being lower than in the reference spectrum; note too that the $n = 2$ spectrum contributes significantly to part of the spectrum below the $n = 1$ edge, between 26 and 29 GeV/c.

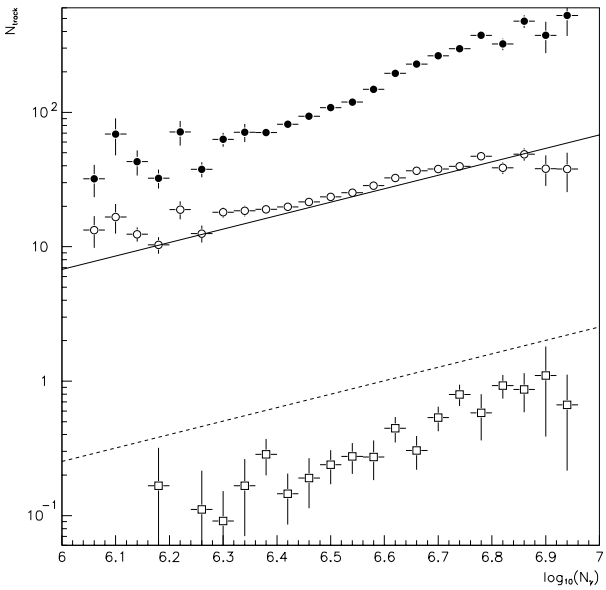

Figure 7.16: Variation with total number of photons of total number of reconstructed tracks (closed circles), sideband-subtracted number of tracks in an $n = 1$-dominated region (open circles), and sideband-subtracted number of tracks in an $n = 2$-dominated region (open squares), for real data from run 15296. A solid line shows the theoretical prediction for the $n = 1$ track rate using the known $n = 1$ Compton spectrum and converter foil thickness; a dashed line shows a prediction for the $n = 2$ track rate assuming the accepted theory, E-144 "standard" (GLC4) focal parameters, and $\eta_{\rm foc} = 0.31$.

One may also judge the smallness of high track density effects by the linearity of the number of reconstructed tracks as a function of total number of backscattered photons. This is shown in Fig. 7.16. The sideband-subtracted number of $n = 1$ tracks indeed varies quite linearly with the number of incident photons.

7.8 Conclusion

It is safe to conclude that the effects of high track density on combinatoric backgrounds and track reconstruction efficiency are less severe in the real data than in the Monte Carlo. Since the Monte Carlo shows excellent reconstruction properties, the real data should also be free of any significant combinatoric background contamination or density-dependent track efficiency variations. The linearity of the $n = 1$ track signal with the independently-measured flux of photons from the interaction point supports this conclusion.

Chapter 8

Results from the measurement of the nonlinear Compton photon spectrum

Synopsis

Results are presented for the spectrum of positrons and electrons arising from Bethe-Heitler conversion of photons from the interaction of 46.6-GeV electrons with the tightly-focused terawatt laser at the E-144 experiment. Particles with energies exceeding the ordinary Compton kinetic edge are clearly seen, with an endpoint consistent with a process in which two photons are absorbed and one emitted. However, the total number of such $n = 2$ particles is significantly less than calculated from theory for the experimental conditions as monitored by independent detectors. The observed spectrum can be fit to a theoretical curve if the laser intensity input to the theory is adjusted as a free parameter, but the fitted laser intensities thus found on various subsets of the data do not match the Cherenkov-monitor-based laser intensities. This is not due to polarization effects in electron multiple scattering, nor can it be due to coherent multiple photon emission, and rescaling of the Cherenkov monitors cannot simultaneously bring the electron and photon data into agreement. No adequate explanation for the result exists at present.

8.1 Introduction

Results from the E-144 electron and positron detectors have been extensively analyzed [1, 4, 51, 52, 49, 6], and found to be in agreement with the simulation to within experimental errors. These experimental errors are dominated primarily by the systematic accuracy and statistical precision of the laser focal parameter reconstruction, and secondarily by systematic and/or statistical limitations of the detectors.

Here, detailed results are presented from the CCD tracking spectrometer, which detected the positrons and electrons created by backscattered photons striking a thin aluminum target. Because of the high rates, this detector is less limited by statistics than the positron calorimeter; because of the absence of "splash" and showering in the detector, the systematic difficulties present in the electron calorimeter data are avoided. Additionally, the analysis of this data benefits from the cross-calibration of the Cherenkov monitors (as described in [Chapter 4]) which decreases the systematic uncertainty in the absolute scale of the laser intensity. Thus, this should be our best data yet.

Clear evidence is seen in the data for the presence of $n = 2$ non-linear Compton scattering. Surprisingly, the total rate is about half of what is expected given the observed recoil electron

yields. These results are presented and analyzed in detail below.

8.2 Results

Tracks were reconstructed and backgrounds estimated using the methods described in [Chapter 7]. Focal properties were estimated on every pulse, as described elsewhere[Chapter 5]. Using this data, the photon spectrum according to our simulation was calculated and redistributed according to the Bethe-Heitler spectrum, using the same formulae as used by EGS4[32]. The single-arm track spectrum thus calculated can be compared directly to the observed data, with an overall normalization to account for the net system efficiency as the only unknown.

Figure 8.1 shows the track spectrum integrated over all events (histogram), the simulation with fitted background and normalization (solid curve), and a scaled and shifted simulation with re-fitted background (curve with long dashes); in the latter curve, adjustments to the signal scale, background scale, background offset, and systematic horizontal angle correction (see Ref. [Chapter 7]) were optimized for maximum likelihood agreement of data to model. The parameters of the fits can be seen in Table 8.1. In both fits, a region of the spectrum near the $n = 1$ kinematic edge, from 28 GeV/c to 30 GeV/c, has been excluded from the fit. The background estimate from the sidebands (op. cit.) is also shown (histogram and curve with short dashes). Many features of the data match the theory very well: the $n = 1$ kinematic endpoint at 29.0 GeV/c momentum is clearly seen, there are hundreds of tracks above background in the $n = 2$ region, and the $n = 2$ kinematic endpoint at 35.8 GeV/c is obvious. However, there is a clear deficit in the number of tracks in the $n = 2$ region relative to the number in the $n = 1$ region.

The simulation was done for each event using focal parameter data from the Cherenkov monitor fits[Chapter 5]. The maximum-likelihood "scaled and shifted" curve in Fig. 8.1 uses the same simulation results. By careful study of Fig. 8.1, one may see that the maximum likelihood optimizer drastically adjusts the background scaling and offset in order to better fit the $n = 2$ portion of the spectrum: it actually multiplies the background estimate by a factor of 3.4. However, the results are not worth the effort: if the χ^2 is taken seriously (and it should be), then the probability of the observed track spectrum being drawn from the expected distribution according to the unfitted model is less than 1 in 3×10^{185}, whereas if the true distribution were that found by this maximum-likelihood fit, the probability of seeing the observed spectrum would be 1 in 4×10^{32}, an impressive, yet insufficient, 153 orders of magnitude improvement. These χ^2 ignore the region from 28 GeV/c to 30 GeV/c: if that region is included, the four-parameter fit cannot do nearly so well, and the χ^2 from both fits are dramatically higher. Furthermore, the multiplication of the background by a factor of 3.4 is non-physical, and results in a predicted spectrum significantly higher than the observed spectrum near the $n = 1$ edge, whereas the Monte Carlo simulation of the CCD system performance predicts that the observed signal should, if anything, be higher than the input spectrum. Thus, the scaling and shifting of signal and background is ill-advised.

Instead of using the Cherenkov-monitor-derived laser intensity information to calculate the theoretical spectra, it is possible to introduce focal area and/or pulse length as another free parameter. For such fits, the area and pulse length parameters are varied in some fixed relation that maintains the pulse energy constant. Trying many different intensities is not too computation-intensive because of Christian Bula's wonderful simulation look-up tables[26]. With focal intensity thus introduced as a free parameter, the observed spectra in the CCDs can be fit quite well; however, it is apparent from Fig. 8.1 that a significantly lower intensity must be required in order to fit the data.

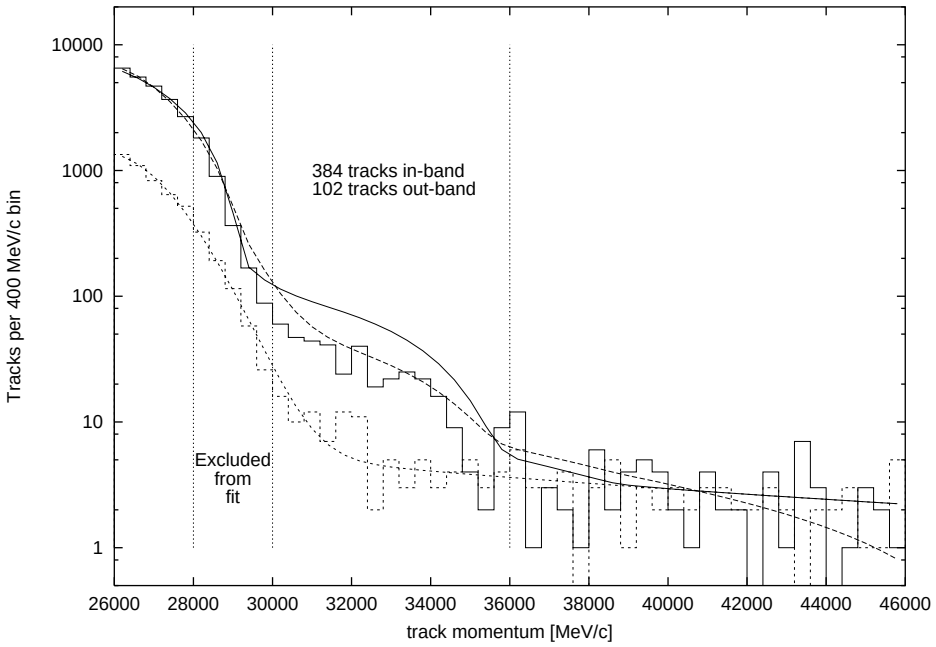

Figure 8.1: The track spectrum integrated over all events (histogram), simulation with fitted background (solid curve), scaled and shifted simulation with re-fitted background (curve with long dashes), and fitted background and estimated from sidebands (histogram and curve with short dashes). The $n = 1$ kinematic edge is located at 29 GeV/c; data from 28 GeV/c to 30 GeV/c is excluded from the fit to avoid any possibility of near-edge systematic effects.

Table 8.1: The parameters and χ^2 per degree of freedom (DOF) for each maximum-likelihood fit. By convention, "signal scale" equals 2.0 in the case of perfect efficiency in beam overlap at IP1 and in track detection in both arms of the spectrometer. Asymmetric errors for η indicate change required to reduce log-likelihood by 0.5; errors for other parameters are from the canonical covariance matrix of the fit, not including the covariance with η.

fit description	η	sig. scale	$\Delta\theta_{\mathrm{syst.}}$	background		χ^2/(DOF)
			[μrad]	scale	offset	/
Fig. 8.1, solid	—	1.23 ± 0.01	—	—	—	312/44
Fig. 8.1, dashed	—	0.50 ± 0.01	0 ± 6	3.52 ± 0.02	-6.84 ± 0.17	154/41
Fig. 8.2, very low	$0.080^{+0.026}_{-0.010}$	0.87 ± 0.02	-103 ± 19	—	—	33/42
Fig. 8.2, low	$0.156^{+0.012}_{-0.010}$	1.42 ± 0.02	-95 ± 16	—	—	58/42
Fig. 8.2, medium	$0.149^{+0.011}_{-0.011}$	1.53 ± 0.02	-96 ± 17	—	—	77/42
Fig. 8.2, high	$0.186^{+0.012}_{-0.011}$	1.61 ± 0.02	-124 ± 16	—	—	40/42
Fig. 8.2, very high	$0.183^{+0.018}_{-0.017}$	1.54 ± 0.03	-191 ± 22	—	—	57/42

Table 8.2: Definition of cuts in Cherenkov-monitor-based η made to separate data into subsets for analysis of dependence on laser intensity, along with the number of events, average values of η^2 and η^4, and fractional systematic effect of using a single value of η to characterize both $n = 1$ and $n = 2$ rates, for the subsets described in the text and Table 8.2. See Table 8.3 for more information.

descr.	$\eta_{\min}$	$\eta_{\max}$	no. events	$\langle\eta^2\rangle$	$\langle\eta^4\rangle^{1/2}$	$(\langle\eta^4\rangle - \langle\eta^2\rangle^2)^{1/2}/\langle\eta^2\rangle$
very low	0.15	0.20	136	.0329	.0332	13%
low	0.20	0.25	223	.0508	.0512	12%
medium	0.25	0.30	179	.0762	.0766	11%
high	0.30	0.35	173	.1047	.1051	9%
very high	0.35	0.40	83	.1404	.1408	8%

In order to study the relationship between CCD-derived laser intensities and Cherenkov-monitor-derived intensities, the data sample is split into five approximately-equal subsets according to the Cherenkov-monitor-based η measurement. Table 8.2 shows the range of Cherenkov-fit η accepted into each of the five subsets, along with several moments of the η distribution within each bin, and Table 8.3 shows additional statistical information for each subset, weighting the statistics by observed rates in various detectors. This statistical information is intended to convince the reader that systematic effects from the fluctuation of focal parameters in each data set should be small. Figure 8.2 shows the track spectrum from the five subsets, and Table 8.1 shows the parameters for each maximum-likelihood fit: signal scaling factor, systematic angle correction, background re-scaling, and offset. Note that, by convention, "signal scale" equals 2.0 in the case of perfect beam overlap at IP1 and perfect track detection and reconstruction in both arms of the spectrometer. Thus, a scaling factor of 1.5 implies an average value of 75% for the net system efficiency. Figure 8.3 plots the CCD-based η fit result versus the mean value of Cherenkov-fit-based η in each data subset. All of the above-mentioned figures and tables support the conclusion that the forward photon spectrum appears to be characterized by lower values of η than the electron spectrum measured by the Cherenkov monitors.

One might wonder if the apparent discrepancy between observation and theory (and/or between electrons and photons) could be a feature of the fitting algorithm or some other uninteresting systematic effect. To address this concern, it is helpful to look at something more "raw" than the fit-based η values plotted in Fig. 8.3. Two such "raw" signals in the CCDs are the background-subtracted count of tracks having momentum between 26 and 29 GeV/c, a measure of $n = 1$ rate, and the background-subtracted count of tracks with momentum between 30 and 35.6 GeV/c, a measure of $n = 2$ rate. (The region between 29 and 30 GeV/c is omitted in order to avoid any possibility of contamination from $n = 1$ signal.) These two counts will be called N1CCD and N2CCD, for convenience; they are plotted versus their corresponding Cherenkov monitor signals, NGAMMA and N2MO, in various combinations in Fig. 8.4. The expected linear relationship between the two "raw" $n = 1$ measurements is observed. As before, a net system efficiency of approximately 75% is deduced from the $n = 1$ data. A factor-of-two difference between the CCD and Cherenkov $n = 2$ data is observed in the "raw" data, which is consistent with the CCD spectrum fits giving values for η roughly $1/\sqrt{2}$ times the Cherenkov-monitor-based data.

The analysis can be further validated using the complete Monte Carlo simulation of the CCD system described in Ref. [Chapter 7]. This Monte Carlo code was used to generate simulated CCD

Table 8.3: The mean values of various collision parameters (from the Cherenkov monitor fits), with flat weighting and with weighting by various detector rates.

descr.	$\eta_{\min}$	$\eta_{\max}$	weighting	$\langle\eta\rangle$	$\langle A\rangle$ [μm^2]	$\langle\tau\rangle$ [psec]	$\langle E\rangle$ [mJ]	$\langle$(overlap factor)$\rangle$
very low	0.15	0.20	1	.181	43.4	2.17	322.	.516
			N_γ	.181	43.4	2.17	323.	.543
			N2MO	.183	43.2	2.16	325.	.548
			$n = 2$ tracks	.183	43.9	2.20	337.	.564
low	0.20	0.25	1	.225	37.3	1.88	375.	.583
			N_γ	.225	37.5	1.89	381.	.615
			N2MO	.227	37.9	1.92	396.	.622
			$n = 2$ tracks	.227	37.3	1.87	380.	.651
medium	0.25	0.30	1	.276	33.7	1.73	472.	.606
			N_γ	.276	34.1	1.76	484.	.659
			N2MO	.277	34.4	1.78	500.	.658
			$n = 2$ tracks	.275	34.2	1.77	485.	.640
high	0.30	0.35	1	.323	30.8	1.57	536.	.564
			N_γ	.323	31.3	1.60	556.	.616
			N2MO	.325	31.5	1.61	564.	.613
			$n = 2$ tracks	.323	31.3	1.61	556.	.611
very high	0.35	0.40	1	.374	27.8	1.39	574.	.573
			N_γ	.375	28.3	1.41	592.	.628
			N2MO	.376	28.3	1.42	597.	.630
			$n = 2$ tracks	.378	28.0	1.40	593.	.678

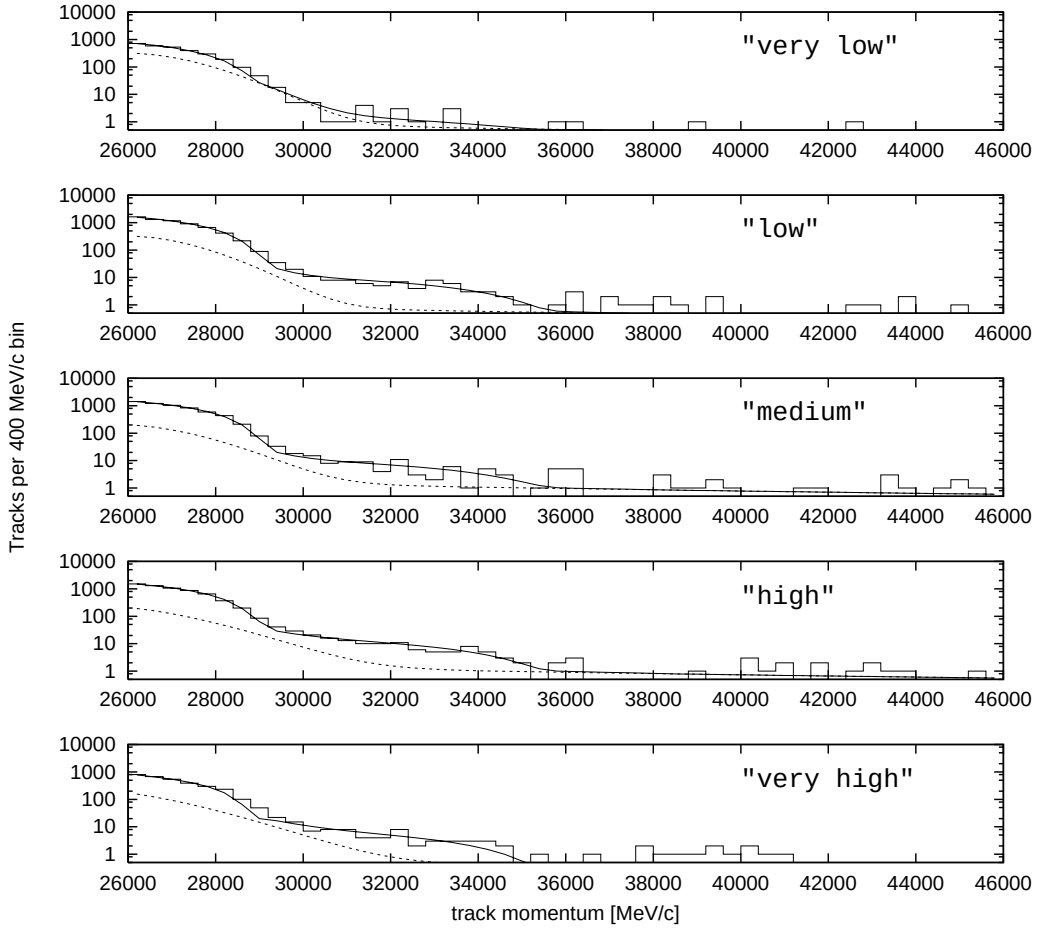

Figure 8.2: Track spectrum from five subsets of the CCD spectrometer data, described in Table 8.2.

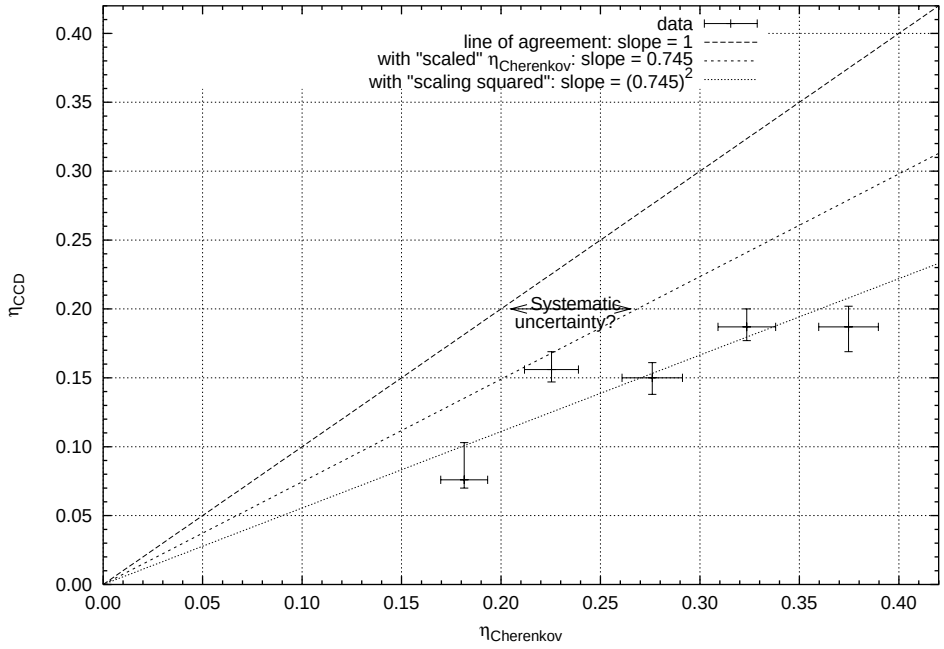

Figure 8.3: CCD-based η fit result versus the mean value of Cherenkov-fit-based η in each data subset: vertical error bars are from the CCD data fit, and horizontal error bars are the rms of the distribution of Cherenkov-fit η.

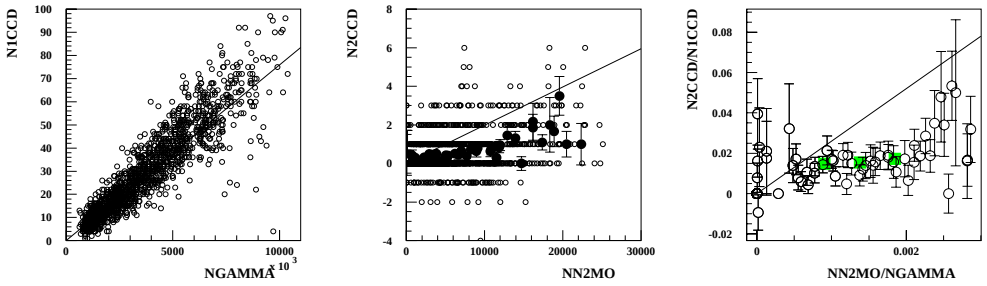

Figure 8.4: Comparison of "raw" $n = 1$ and $n = 2$ data from Cherenkov monitors and CCD track spectrometer: a) N1CCD vs. NGAMMA for each event; b) N2CCD vs. N2MO for each event (small points) and ratio of averages for subsamples of the data selected according to fitted N2MO/NGAMMA (larger circles with error bars); c) ratio of average N2CCD to average N1CCD plotted vs ratio of average N2MO to average NGAMMA, for small subsamples selected as above. Note that systematic effects push the ratio to artificially high values in the low-statistics bins at the extremes of N2MO/NGAMMA. In each graph, the curve is the model, assuming 75% net system efficiency as suggested by the fitted scale parameters in Table 8.1.

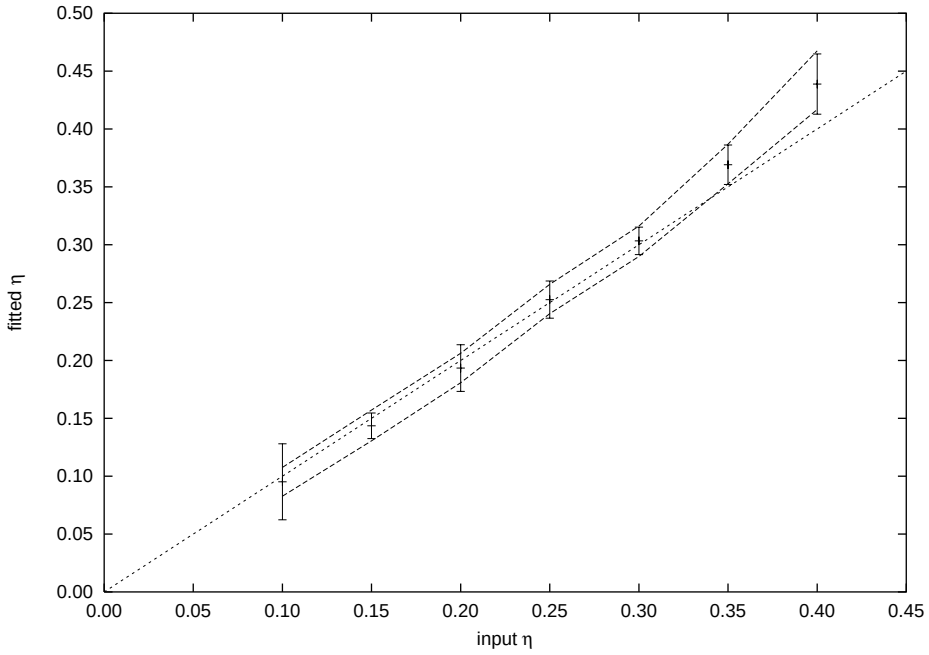

Figure 8.5: Results of analyzing Monte Carlo CCD data: fitted η versus the actual value of η input to the Monte Carlo, with vertical error bars indicating the rms of the fitted η from several independent Monte Carlo data sets, and dashed lines above and below indicating the average of the upper and lower asymmetric error bar 1-sigma limits. The dotted line has unity slope, indicating the ideal behavior for the reconstruction and fitting algorithm.

data files for collisions at $\eta = 0.10, 0.15, 0.20, 0.25, 0.30,$ and 0.35. Eleven to thirteen data sets were generated at each of these values of η. The same track reconstruction and analysis code was applied to these Monte Carlo data sets as was applied to the real data. The results of this analysis are shown in Table 8.4 and Fig. 8.5, which are analogous to Table 8.1 and Fig. 8.3 for the real data. There is no sign of any bias or internal inconsistency in the simulation or analysis procedures: agreement between theory and analyzed Monte Carlo data is excellent at low values of η, while at the very highest η, the track efficiency degradation in the densely-populated $n = 1$ region causes a small systematic *increase* in apparent η.

In [Chapter 5], the point is made that multiple sets of collision geometry parameters can lead to the same electron Cherenkov monitor readings. Notwithstanding, it is also suggested that there is effectively a one-to-one correspondence between Cherenkov monitor readings and the spectrum of forward photons, *i.e.*, the degeneracy among fitted laser focal parameters is largely unimportant, at least when the parameters are confined to the limits allowed by the Cherenkov monitor fit algorithm. Nevertheless, there are slight differences in how the two spectra vary with laser focal area, and one might wonder if a sufficiently large shifting of fitted focal parameters could bring the two particle spectra into agreement.

In fact, the "slight differences" are too small to exploit in this manner, as can be seen from the power law exponents in Table 8.5, adapted from [Chapter 5]. The area power-law exponents for the $n = 2$ electron and photon counts differ by only 0.2, so in order to account for the roughly factor-of-two discrepancy, the area would have to be increased by a factor of over 30. Assuming

Table 8.4: The average and rms of fit parameters and average and rms χ^2 (unnormalized, with 45 degrees of freedom) for the maximum-likelihood fits to the Monte Carlo CCD data. At least eleven simulations were performed at each value of η listed.

η input	η fitted	sig. scale	$\Delta\theta_{\text{syst.}}[\mu\text{rad}]$	b.g. scale	b.g. offset	χ^2
0.10	0.095 ± 0.033	1.79 ± 0.34	54 ± 61	46.94 ± 85.85	-0.04 ± 0.09	39 ± 13
0.15	0.144 ± 0.011	1.76 ± 0.16	56 ± 27	11.55 ± 5.76	-0.06 ± 0.10	33 ± 14
0.20	0.193 ± 0.020	1.72 ± 0.16	61 ± 38	5.85 ± 1.76	-0.01 ± 0.04	39 ± 14
0.25	0.253 ± 0.016	1.79 ± 0.16	37 ± 27	3.28 ± 0.85	-0.00 ± 0.01	31 ± 9
0.30	0.303 ± 0.012	1.70 ± 0.13	53 ± 52	2.67 ± 0.36	-0.00 ± 0.00	29 ± 6
0.35	0.369 ± 0.017	1.61 ± 0.15	87 ± 64	2.20 ± 0.23	0.01 ± 0.03	35 ± 12
0.40	0.439 ± 0.026	1.44 ± 0.21	130 ± 46	2.08 ± 0.23	0.01 ± 0.02	37 ± 9

Table 8.5: Fitted power law exponents for various interesting regions of the electron and photon spectra. A denotes the laser focal area, τ the laser pulse length, and E_l the total pulse energy, $E_l \propto A\tau\eta^2$.

	electrons			photons		positrons
	$n=1$	$n=2$	$n=3$	$n=1$	$n=2$	(all)
η^2 or E_l (fixed A, τ)	1.0	2.0	2.9	1.0	2.0	5.1
A (fixed η^2, τ)	1.0	1.5	1.7	1.0	1.3	2.0
τ (fixed η^2, A)	1.1	1.1	1.0	1.1	1.0	1.1
A (fixed τ, E_l)	0.0	-0.5	-1.2	0.0	-0.7	-3.1

for the sake of discussion that the power-law exponents can be extrapolated that far, then making such a radical change in area without a corresponding change in pulse energy would accomplish the factor-of-two change in the relative rates by reducing the absolute number of $n = 2$ tracks by a factor of 10 while reducing the absolute number of $n = 2$ electrons by a factor of 5. This would be totally incompatible with the absolute rates seen in the CCDs and Cherenkov monitors. Compensating by increasing the laser pulse energy over a factor of 2 would alleviate the problem in the $n = 2$ detectors, but the absolute $n = 1$ rate would then be a factor of two higher while the $n = 3$ electron rate would be a factor of 7 lower than observed. Thus, reconciling the observation with our simulation simply by adjusting the focal parameters is ruled out based on the particle detector data alone. Moreover, despite the lack of direct shot-to-shot measurements of focal area and pulse length, it is highly unlikely that the laser could be so mischaracterized that the laser energy could actually be two times higher than thought, and if the focus were really 30 times larger than believed, the light almost certainly could not have been recollimated and returned to the laser room[53].

In short, many plausible experimental explanations of the discrepancy have been ruled out, and the simplest interpretation of the irreconcilabilty of the forward photon and recoil electron data is that some aspect of the data is not in agreement with our model.

8.3 Possible reasons for discrepancy

The lack of agreement between simulation and observation is clear, but the source of the discrepancy is not self-evident. Having found an apparent disagreement between model and data, there is no

a priori reason to suppose that the cause must be a deficit in $n = 2$ photons for a given set of collision parameters. Because of the indirect way we estimate laser intensity, it could equally well be that we have an excess of $n = 2$ electrons.[1] This ambiguous state of affairs is distressing to the experimentalist, allowing one more degree of freedom for speculative explanations of the effect than can be addressed with the present data, and it is equally frustrating for the theorist, leaving twice as much open ground to cover; but, it can not be helped.

Before considering the possibility that the data could indicate a problem in the theory of quantum electrodynamics, there are two less dramatic possibilities to consider. The first possibility is that the simulation does not properly take into account all theoretically-expected features inherent in the processes included in our simulation. The second possibility is that some additional process, currently not included in our simulation but predicted under existing QED theory, is in fact more important than assumed and should be incorporated into our model.

At present, the only processes included in the simulation which are relevant to the electron and photon spectra are single and repeated nonlinear Compton scattering:

$$\left.\begin{array}{rcl} e + n(\text{laser}) & \rightarrow & e' + \gamma \\ e' + n(\text{laser}) & \rightarrow & e'' + \gamma' \\ e'' + n(\text{laser}) & \rightarrow & e''' + \gamma'' \\ \ldots \text{ et cetera } \ldots & & \end{array}\right\} \tag{8.1}$$

Currently, the simulation treats all electrons as unpolarized in the initial, intermediate, and final states. One might suppose that polarization of the intermediate electrons (e', e'',...) might be important, just as the effect of polarization of the backscattered photons was found to be important in the two-step pair production process [49]. This is a potential candidate for the first type of explanation.

For an example of a candidate explanation of the second type, consider the process

$$e + n(\text{laser}) \rightarrow e' + m\gamma, \tag{8.2}$$

i.e., coherent *emission* of m multiple photons by an electron in a strong wave. This is certainly allowed and presumably calculable under QED, but is not included in our simulation in any form.

Due to the relatively recent nature of these results, no other candidate explanations for the disagreement between observation and simulation have been advanced at present. The two candidates are discussed further in the next two subsections, where it is shown that electron polarization effects are likely to be important in some cases and should be included in the simulation as soon as possible, but that they cannot be the cause the observed effect in this data, while inclusion of coherent multiple emission is likely to be unimportant for the E-144 interaction geometry.

8.3.1 Intermediate electron polarization

The incident electrons have vertically-aligned spins (up or down, hence transverse to the beamline), and geometric considerations suggest that the scattering rate for such electrons should be equal

[1] If there were no independent measure of the total photon flux, then the disagreement seen could also be due to a deficit in the $n = 1$ electron rate or an excess in the $n = 1$ CCD track rate, as the apparently excellent agreement in the $n = 1$ rates is an artifact of the data-based method used to estimate the relative efficiency of the CCD and electron-counting systems. However, the forward photon Cherenkov monitor CCM1 was found to match the $n = 1$ electron monitors extremely well. An effect radically affecting the shape of the $n = 1$ CCD track spectrum while leaving the total flux unchanged can not be ruled out, but it seems unlikely.

to the unpolarized scattering rate; in any event, the polarization is flipped randomly from pulse to pulse, and the data is summed over all pulses. However, the photons radiated have a preferred polarization, and this affects the pair production rate.[49] The recoil electrons may also have a preferred polarization. The concern is that polarization of the recoil electrons may affect the likelihood of their rescattering. Noting that photons backscattered from a circularly-polarized laser can attain 100% polarization, it has been suggested that the effect on multiply-scattered electron rates could be as large as a factor of $2^{(m-1)}$ for the *circularly*-polarized laser for scattering near the kinematic endpoint,[54] where m is the number of incoherent Compton interactions undergone by the electron.

Below, three arguments are made that polarization of the recoil electrons is not the source of the discrepancy seen in this data. The first is a quantitative argument that the magnitude of the polarization effect cannot be large enough to explain the results. The second is a qualitative argument that the sign of the effect is wrong. The third is a geometric argument that, in the special case of a linearly-polarized laser, there is no rate dependence on the initial electron polarization, and therefore there can be no effect on multiple scattering rate no matter how polarized the recoil electrons may be.

For the quantitative argument, note that less than 50% of the rate in the $n = 2$ electron counter (N2MO) is contributed by multiple scattering. Therefore, even if the multiple scattering contribution were to double, the total rate could increase by only 50%. This is not sufficient to explain the discrepancy.

For the qualitative argument, consider that, in order for double $n = 1$ scattering to scatter an electron into N2MO, both scatters must occur near the kinematic endpoint, *i.e.*, both scatters must have a polar angle near 180° in the center of mass frame. When an electron backscatters in this way off a circularly polarized laser, it is most likely to radiate a photon of helicity opposite to the incident wave, after which the electron has the same helicity (opposite spin) as the radiated photon. This direction of spin is less favorable to further scattering near the endpoint, so that one should expect a *suppression* of the multiple scattering rate, not an enhancement as is required to explain the experimental results.

When scattering off of linearly polarized light, an electron will tend to radiate photons of polarization parallel to the incident light. However, there is no obvious reason why electrons of any particular polarization should be more or less likely to scatter. In fact, if we consider electrons in eigenstates of transverse spin operators, *e.g.*, spin up or down when the electron travels to the east and the laser travels west, with a vertically-polarized laser, then the two orthogonal states "spin up" and "spin down" are geometrically indistinguishable and therefore must have the same scattering probability. This equally applies to a transverse, horizontal basis for the spin. While it is true that the left and right helicity states of the electron ("spin east" and "spin west") are geometrically distinguishable, the helicity of the photons changes under parity while the polarization of the laser does not, and therefore parity symmetry of QED requires that they also have the same scattering rate. There may be (and is) an effect on the azimuthal distribution of the scattered electrons, of course, and the initial electron polarization doubtlessly affects the final electron polarization, but the total scattering rate and the distribution of final electron polar angles and momenta must be the same after integration over azimuthal angle. In short, when the laser is linearly polarized, the scattering rate is independent of the initial electron polarization, and the effect of polarization of the recoil electrons must be zero for the linearly polarized laser.

8.3.2 Coherent emission of multiple photons

Roughly speaking, emission of a second photon should introduce a factor of $\alpha \cong 1/137$ into the scattering rate. One might therefore expect the total contribution from the two-photon-emitted processes to the recoil electron spectrum to be about 0.7% of the one-photon-emitted processes. One consideration that might give us pause, however, is that multiple incoherent interactions (Eq. 8.1) constitute a subset of the more general multiple-emitted-photon processes, and they contribute a large fraction of the total scattering rate. For example, the double-scattering process contributes about 50% of the scattered electrons seen between the $n = 1$ and $n = 2$ kinematic edges. The reason for this is that the incoherent double-scattering yield grows as the square of the time it spends in the field, rather than simply linearly. A better (or at least, different) way to estimate the contribution of coherent multiple photon emission might be to compare the coherence length of the scattering process to the total depth of the laser focus as seen by the electron.

For Compton scattering, the coherence length is roughly equal to the wavelength of the laser light. The effective thickness of the focal spot along the electron's path is $\sqrt{A}/\sin\theta \simeq 20$ μm, where $A \simeq 30$ μm^2 is the focal area and $\theta = 17^\circ$ is the collision angle. In traversing this path, the electron will encounter approximately 70 cycles of the electromagnetic wave. Thus, the coherent contribution would be roughly 1/70 times the 50% contribution of the incoherent multiple contribution in the $n = 2$ region, giving the same estimate of $1/\alpha \simeq 0.7\%$ as from the simple vertex-counting argument. This means that the vertex-counting argument is valid in its application to the coherent multiple emission processes, and that a simulation including only incoherent multiple scattering is accurate enough for this particular interaction geometry.

Note that if the collision angle were 170 degrees, the electron would encounter less than one full cycle as it traversed the focal spot. In this case, coherent effects would presumably be significant compared to incoherent multiple scattering. This may have significance to anyone contemplating a laser acceleration experiment. However, since η varies with collision angle as $1 + \cos\theta$, the value of η would be over two orders of magnitude less, all other things being equal, so "hard" nonlinear scattering effects may not be of concern at all in such a case.

8.3.3 Other ideas

Given the way focal parameters are estimated from Cherenkov monitors and the way Cherenkov monitors are calibrated from ECAL, it's possible that a factor of 2 or 3 recalibration of ECAL would lower eta and solve all problems in the e- and photon data. However, the eta estimates would then be way too low for the number of positrons we see. This could be partly offset by revising assumptions of whether area or pulse length (or both) are the dominant contributors to intensity fluctuations, at some cost to the "too perfect" slope of positron rate versus fitted intensity. There are awkward limits on both what can be known and what can be invented about the focal parameters, which are discussed in more detail in [Chapter 5].

It was suggested[55] that a different laser polarization (*e.g.*, elliptical) could produce this sort of effect. Slightly higher η would balance the e+ polarization effect for the $n = 5$ (6 photons in the initial state) pair production[49], while the $n = 2$ coherent process could be suppressed, and multiple scattering might be mostly unaffected. Perhaps the presumed linear polarization could be ellipticized by an unanticipated effect in the doubling crystal or windows. Elliptical polarization cannot be treated in either of the current simulations, but the effect of circular polarization of the laser was investigated simply by substituting the circular polarization look-up tables for the linear polarization look-up tables. Slightly higher values of η were indeed fitted for both the CCD

and Cherenkov monitors, but they were higher by nearly the same factor, so the deficit was not resolved.

8.4 Conclusions

There is clear evidence for the occurrence of the emission of photons corresponding to the backscattering of the second harmonic of the incident wave. The statistical significance of this conclusion is huge: simply from the count of tracks between 30 GeV/c and 36 GeV/c (Fig. 8.1), where only 102 ± 10 background tracks are expected, one may calculate that a statistical fluctuation of 29 standard deviations would be required to generate the observed 384 tracks. However, if one accepts the reliability of the track detection system and reconstruction code [Chapter 7] and the validity of the Cherenkov monitor calibration [Chapter 4] and associated focal parameter estimation [Chapter 5] procedure, then there can be no question that the data indicate a relative deficit in the total number of $n = 2$ photons, or equivalently, an excess in the $n = 2$ and $n = 3$ electron rates, as compared with our model.

Appendix A

E144 ECAL "Best Aperture" Reconstruction

Synopsis

A method is described for choosing the coefficients of the reconstruction matrix R such that the reconstructed signal $R \cdot X$ represents the integral of the incident energy over a sharply defined region at the face of the ECAL, independant of the shape of the spectrum.

A reasonable way to "reconstruct" the spectrum incident on the face of the ECAL would be to assume some shape $f(y)$ for the distribution of energy within each row (e.g. exponential in y), calculate the coefficients of the matrix $M[f]$ which describes the transformation from total incident energy in each row to observed readout after showering, and then find the reconstruction matrix R such that $R \cdot M[f] = I$. This is *not* the method described in this paper.

The "best aperture" reconstruction ignores the matrix $M[f]$. The matrix R is chosen such that the reconstructed signal responds only to points on the incident spectrum which lie within a sharply defined region on the face of the ECAL. This is done using empircally measured shower and splash coefficients, obtained from test beam and XT scan data, respectively. In theory, the method is capable of finding R matrices such that the response of the reconstructed signal to energy incident more than 3 mm outside of the nominal aperture is less than 1%, while the response to splash is typically 9 GeV/$10^6\gamma$ or less, after taking into account the ECAL energy resolution and possible errors in parameterization of empirically measured splash and shower profiles.

A.1 Basis of method

The response of the calorimeter to a delta-function beam incident at position y (and horizontally centered) can be written as a column vector $X_j(y)$, where j labels the readout channels of the ECAL. The readout of the ECAL when presented with a spectrum $f(y)$ is therefore

$$\mathcal{X}_j[f] = \int dy f(y) X_j(y)$$

When the reconstruction matrix R is applied to the observed readout, the reconstructed signal is given by

$$\mathcal{R}_i[f] = \int dy f(y) R_{ij} X_j(y)$$

where i labels the reconstructed signals, and a sum over j is implied. The functions $g_i(y) = R_{ij}X_j(y)$ are therefore the Green's functions for the response of reconstructed signal i to the differential bit of the spectrum incident at y. It can be regarded as the aperture or acceptance function of the reconstructed channel.

In order to obtain a measurement of the total flux in a region $y_{i-1} < y < y_i$, it is not necessary to know the function $f(y)$, provided we can adjust the coefficients R_{ij} such that $g_i(y)$ is zero for y outside the region and one for y within the region. While this condition can not be achieved precisely, it can be approached to a high degree of accuracy, depending on how well the shower spreading coefficients are known and what other conditions (such as insensitivity to the "splash" backgrounds) are imposed.

A.2 Calculation of R

To find the reconstruction matrix which most closely meets the requirements, a chi-squared measure is defined, and then minimized with respect to the elements of the reconstruction matrix R_{ij}. To make the problem easier, discrete values of y are used. Call these values y_k. Define $X_{ik} = X_i(y_k)$. Define $D_{ik} = g_i^{\text{desired}}(y_k)$. Define σ_{ik} as the tolerance allowed for the deviation between the achieved and desired aperture function. Then the chi-squared measure is

$$\chi^2 = \sum_i \sum_k \frac{(R_{ij}X_{jk} - D_{ik})^2}{\sigma_{ik}^2}$$

An additional condition on R is that the response of the reconstructed signal be essentially immune to the backgrounds due to "splash". If the shape of the splash is known, i.e. the splash in readout channel j is given by a known constant coefficient V_j times another number which changes event-to-event but is the same for all channels, then what is desired is that $R_{ij}V_j = 0$ for all i. Define the tolerance on this condition as σ'. The new chi-squared for the R_{ij} calculation is

$$\chi^2 = \sum_i \sum_k \frac{(R_{ij}X_{jk} - D_{ik})^2}{\sigma_{ik}^2} + \frac{(R_{ij}V_j)^2}{\sigma'^2}$$

The R_{ij} which minimize this chi-squared can easily be found. The minimization condition is

$$\frac{1}{2}\frac{\partial \chi^2}{\partial R_{ij}} = R_{ij'}\left(\frac{V_{j'}V_j}{\sigma'^2} + \frac{X_{j'k}X_{jk}}{\sigma_{ik}^2}\right) - \frac{D_{ik}X_{jk}}{\sigma_{ik}^2} = 0$$

which is solved by

$$\begin{aligned} \mathbf{A}\vec{R}_i &= \mathbf{B}_i \\ \vec{R}_i &= \mathbf{A}^{-1}\mathbf{B}_i \end{aligned}$$

where

$$\begin{aligned} \mathbf{A} = [A_{jj'}] &= \left[\frac{V_{j'}V_j}{\sigma'^2} + \frac{X_{j'k}X_{jk}}{\sigma_{ik}^2}\right] \\ \mathbf{B}_i = [B_{ij}] &= \left[\frac{D_{ik}X_{jk}}{\sigma_{ik}^2}\right] \end{aligned}$$

The above equations give an R matrix whose predicted aperture function is extremely close to that desired, typically within a hundredth of a percent or better. However, the elements of R obtained are extremely large (e.g. $\pm 10^4$) and change drastically with slight changes in the shower or splash profile. The "reconstructed signal" using such an R will be utter nonsense unless the actual response of the ECAL matches the theoretical model given by $X_j(y)$ and V_j to a very high degree of precision and accuracy.

To obtain an R matrix which is robust under the expected range of error of X_j and V_j, the chi-squared can be expanded to sum over the range of the input parameters. The R matrix so obtained should have much the same aperture function for any shower or splash profile within the expected error. The final chi-squared is therefore

$$\chi^2 = \underbrace{\sum}_{\text{systematic errors in shower and splash parameterization}} \underbrace{\sum}_{\text{readout electronic noise in ECAL}} \sum_i \sum_k \frac{(R_{ij}X_{jk} - D_{ik})^2}{\sigma_{ik}^2} + \frac{(R_{ij}V_j)^2}{\sigma'^2}$$

Minimizing the above chi-squared leads to much more reasonable values for the R matrix elements. Channels with low signal and high noise get low coefficients, while the readout channels with the highest coefficient for a given reconstructed channel are all in the row whose signal is being reconstructed.

A.3 Implementation

1 mm intervals were used for all results shown in this paper. The ideal aperture function values D_{ik} were taken to be 1 inside the ideal aperture, 0 outside the aperture, and 0.5 if the position y_k was exactly on the boundary between two channels. A tolerance of 2% was used for points below the ideal aperture, 10% for points within the aperture, and $1\%/2^{\Delta r}$ for points Δr rows above the top of the ideal aperture.

The splash coefficients are dependant on the Y position of the ECAL. Splash coefficients (in GeV per 10^3 gammas) at various ECAL Y positions were taken from fits to XT scans. Each channel of the ECAL was studied separately. The data for a given channel was plotted versus the ECAL Y position and then fitted to a standard form. Error estimates are made for each parameter of the splash fit, and these errors are used in the calculation of the best R matrix as described in the previous section. A file containing the parameters and their errors is generated; a separate file is made for green and IR data.

The shower functions $X_j(\delta y)$ used were taken from data provided by Kostya Shmakov. The same response to a delta-function beam is assumed for all ECAL Y positions.

The coefficients in the R matrix vary with Y position, due to the change in the splash profile. However, for any Y position, the predicted response of the reconstructed signal to energy incident more than 3 mm outside of the nominal aperture is less than 1%, while the response to splash is typically 9 GeV/$10^6\gamma$ or less.

Figure A.1 shows the predicted aperture functions for 4 reconstructed rows.

A.4 Addenda

A few more things have been added to the basic algorithm described above:

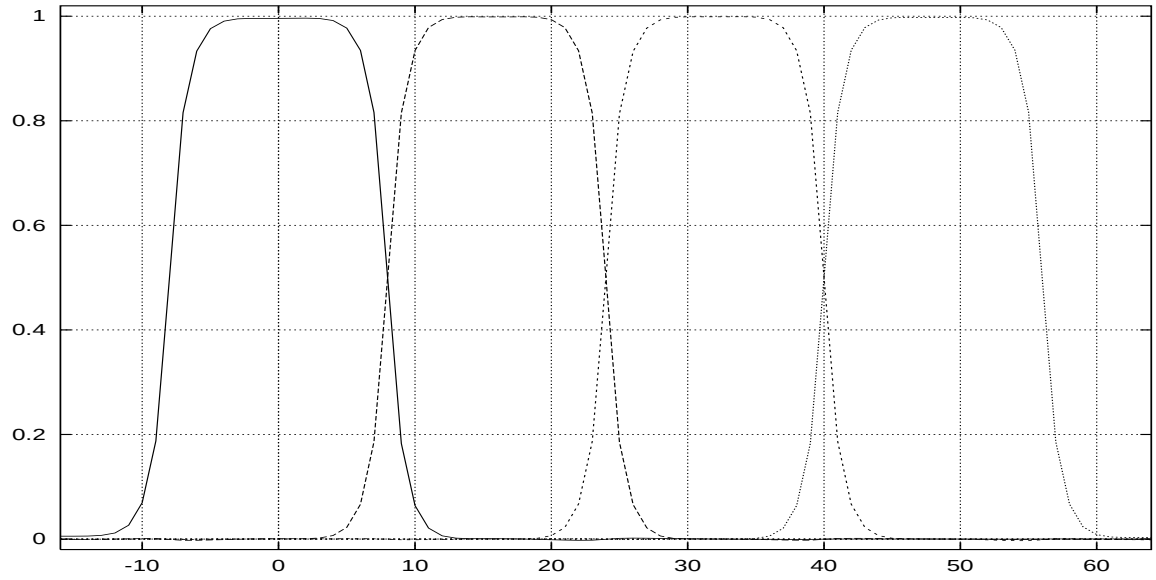

Figure A.1: Aperture functions for 4 reconstructed rows. The ordinate is the position of the "test beam" in mm.

- A second splash vector, the "downsplash" has been introduced to address concerns about the anomalous ECAL signal at low ECAL position and high gamma flux. The χ^2 attempts to produce a reconstruction which meets all of the original requirements *and* has zero response to a "downsplash-shaped" signal. This is done by adding another term of the form $(R_{ij}V_j)^2/\sigma'^2$ to the chi-squared.

- The code now calculates the theoretical response to the splash and downsplash, with error estimates accounting for the calculated input parameter errors, for each reconstructed signal. In other words, the code predicts the amount of splash or downsplash that "bleeds" into the reconstructed face-of-the-ECAL signal. This is done simply by applying the final R matrix to a raw signal vector containing only signal due to the given type of splash; the error is calculated by summing the square of each coefficient times the expected variance of the corresponding raw signal.

- Two additional rows in the R matrix give coefficients for reconstructing signals which respond only to splash or downsplash (respectively) and not at all to electrons hitting the face of the ECAL. This is done by setting the "desired" response to any "face-of-the-ECAL" signal to zero for all positions k, and replacing $(R_{ij}V_j)^2/\sigma'^2$ with $(R_{ij}V_j - 1)^2/\sigma'^2$ in the chi-squared. These splash estimators are used for checking the validity of the calorimeter model (e.g. by plotting the splash reconstruction versus CCM1) and for scaling the "splash-bleed" numbers to estimate the overall corruption of the reconstruction by splash or downsplash.

Appendix B

Nonlinear luminosity integrals and application to overlap scans

Synopsis

An approach is described for rapid calculation of the rates of non-linear and multiple scattering processes in the case where the geometry of the field producing the non-linear effect changes significantly over the size scales of the scattered beam pulse.

B.1 Non-linear luminosities and cross-sections

In this section, the non-linear cross-section is defined so that the scattering rate per unit time per unit volume for an n photon Compton process in the electron's rest frame is

$$W^{(n)} = \rho_e^{\mathrm{ERF}} \rho_l^{\mathrm{ERF}} (\eta^2)^{n-1} \sigma(n, \eta^2) \tag{B.1}$$

where ERF denotes the electron rest frame, ρ_e^{ERF} is the density of electrons as seen in ERF, ρ_l^{ERF} is the density of the photons of the external E-M field as seen in ERF, η is the Lorentz-invariant intensity parameter, and $\sigma(n, \eta^2)$ is the cross-section of the n photon process. The factor of $(\eta^2)^{n-1}$ is included so that the cross-section approaches a finite, non-zero constant as $\eta^2 \to 0$. The dimensionless scalar $(\eta^2)^{n-1}$ is used instead of raising ρ_l^{ERF} to the n^{th} power in order to keep the relativistic transformation of $\sigma(n, \eta^2)$ as simple as possible. The cross-section so defined has units of area for any n.

Note that there is a simple, covariant relationship between the Lorentz-invariant parameter η^2 and the density of photons in a given reference frame times the wavelength of the photons in that frame:

$$\eta^2 \;=\; \frac{3}{4\pi\alpha} \sigma_T \lambda_l \rho_l \;\cong\; 32.7 \sigma_T \lambda_l \rho_l \tag{B.2}$$

where α is the fine structure constant, σ_T is the Thomson cross-section, and λ_l is the wavelength of the light. Therefore, eq. B.1 simply reflects the leading-order behavior of the non-linear process: the rate scales as ρ_l^n. The non-linear cross section could therefore be defined as the coefficient of $\rho_e^{\mathrm{ERF}} (\rho_l^{\mathrm{ERF}})^n \lambda_l^{n-1}$, in which case the units of the "cross-section" would work out as (area)^n. This would appear superficially similar to the multiple-scattering "cross-section" defined in the next session, but that would really be more confusing than illuminating, since the procedure for getting

a scattering rate from a cross-section is quite different in the two cases. We will stick with the cross-section as defined by B.1.

Equation B.1 can be made Lorentz-invariant by expressing the densities in ERF in terms of the product of a 4-vector current and the 4-velocity of the electrons.

$$\begin{aligned} \rho_e^{\mathrm{ERF}} &= J_e^\mu u_{e\mu} \\ \rho_l^{\mathrm{ERF}} &= J_l^\mu u_{e\mu} \end{aligned} \tag{B.3}$$

This leads to the manifestly Lorentz-covariant expression

$$W^{(n)} = (J_e^\mu u_{e\mu})(J_l^\nu u_{e\nu})(\eta^2)^{n-1}\sigma(n,\eta^2) \tag{B.4}$$

The scattering rate per unit 4-volume is invariant under Lorentz transformations, so the cross-section as defined by the above equation is also Lorentz invariant. "Cross-section" always means the geometric cross-section as seen in the electron's rest frame. In the "lab" frame, where the electrons move along the $+\hat{e}_3$ direction and the photons move along the $-\sin\theta\hat{e}_1 - \cos\theta\hat{e}_3$ direction (*e.g.* $\theta = 0$ is a head-on collision), we have

$$\begin{aligned} u_e &= \frac{1}{\sqrt{1-\beta^2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \beta \end{bmatrix} \\ J_l &= \rho_l^{\mathrm{Lab}} \begin{bmatrix} 1 \\ -\sin\theta \\ 0 \\ -\cos\theta \end{bmatrix} \\ J_e &= \rho_e^{\mathrm{Lab}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \beta \end{bmatrix} \\ W^{(n)} &= \rho_e^{\mathrm{Lab}}\rho_l^{\mathrm{Lab}}(\eta^2)^{n-1}(1+\beta\cos\theta)\sigma(n,\eta^2) \end{aligned} \tag{B.5}$$

Thus, the cross-section so defined behaves like the geometric cross-section for a linear process. The total number of scatters is, of course, given by

$$\begin{aligned} N(n) &= \int d^3x dt W^{(n)}(\vec{x},t) \\ &= \int d^3x dt \rho_e(\vec{x},t)\rho_l(\vec{x},t)(\eta^2(\vec{x},t))^{n-1} \\ &\quad \times(1+\beta\cos\theta)\sigma(n,\eta^2(\vec{x},t)) \end{aligned} \tag{B.6}$$

It may happen that the cross-section does not vary greatly over the range of η^2 of interest in a given problem. If so, then it is useful to define the luminosity as

$$\mathcal{L}(n) \equiv \int d^3x dt \rho_e(\vec{x},t)\rho_l(\vec{x},t)(\eta^2(\vec{x},t))^{n-1}(1+\beta\cos\theta) \tag{B.7}$$

$$N(n) \cong \mathcal{L}(n)\sigma(n) \tag{B.8}$$

This allows the "physics" contained in $\sigma(n)$ to be separated from the experimental parameters that affect $\mathcal{L}(n)$. This physics describes the leading-order behavior for the process, *i.e.* the $(\eta^2)^n$ dependence of the rate. Unfortunately, it is a hard fact of life that the higher-order behavior of $\sigma(n, \eta^2(\vec{x}, t))$ for large η^2 will always be convolved with the experimental configuration, due to the dependence of η^2 on position and time. However, note that if the cross-section can be expanded as a power series in η^2, then the total scattering rate can be found more precisely using $\mathcal{L}(n)$ for higher n:

$$\begin{aligned} \sigma(n, \eta^2) &= \sigma_0(n) + \sigma_1(n)\eta^2 + \sigma_2(n)\eta^4 + \cdots \\ N(n) &= \mathcal{L}(n)\sigma_0(n) + \mathcal{L}(n+1)\sigma_1(n) + \mathcal{L}(n+2)\sigma_2(n) + \cdots \end{aligned}$$

Another use for $\mathcal{L}(n)$ is to calculate a weighted average of the parameter η^2. If an n^{th} order process has a probability proportional to $(\eta^2)^n$, then the mean value of η^2 at points at which the process occurs (*i.e.* weighted by the probability of the scattering process) is given by

$$\left\langle \eta^2 \right\rangle_n = \frac{\mathcal{L}(n+1)}{\mathcal{L}(n)} \tag{B.9}$$

B.2 Differential cross-section for multiple-scattering

The scattering process is described by a differential cross section of the form $\frac{d\sigma}{du}(n, \eta^2; u_i, u_f)$, where u_i and u_f are the initial and final states of the scattered particle, η^2 is the intensity parameter when and where the scattering occurs, and n describes the process, *i.e.* it is the number of photons involved in a multi-photon Compton event. The number of single scatters into the final state u_f is given by eq. B.6:

$$\begin{aligned} \frac{dN}{du}(n; u_i, u_f) &= \int dt d^3x \rho_e(\vec{x}, t)\rho_l(\vec{x}, t)(\eta^2(\vec{x}, t))^{n-1} \\ &\quad \times (1 + \beta\cos\theta)\tfrac{d\sigma}{du}(n, \eta^2(\vec{x}, t); u_i, u_f) \\ &= \int dt d^3x \rho_e(\vec{x} + \vec{v}t, t)\rho_l(\vec{x} + \vec{v}t, t)(\eta^2(\vec{x} + \vec{v}t, t))^{n-1} \\ &\quad \times (1 + \beta\cos\theta)\tfrac{d\sigma}{du}(n, \eta^2(\vec{x} + \vec{v}t, t); u_i, u_f) \\ &= \int d^3x dt \rho_{e0}(\vec{x})\rho_l(\vec{x} + \vec{v}t, t)(\eta^2(\vec{x} + \vec{v}t, t))^{n-1} \\ &\quad \times (1 + \beta\cos\theta)\tfrac{d\sigma}{du}(n, \eta^2(\vec{x} + \vec{v}t, t); u_i, u_f) \end{aligned} \tag{B.10}$$

In the second form of the equation, the variable of integration $\vec{x}$ has been shifted by $\vec{v}t$, as we are free to do since the integral is over all space and time, and the order of time and space integrals can be freely exchanged. The vector $\vec{v}$ may be taken to be the electron velocity. Then $\vec{x}$ becomes the position of the electron at $t = 0$, *i.e.* it is the relative position of the electron within the electron bunch distribution. The last form of the equation simply substitutes $\rho_e(\vec{x} + \vec{v}t, t) = \rho_{e0}(\vec{x})$, defining $\rho_{e0}(\vec{x})$ to be the spatial distribution of the electron bunch at time $t = 0$. In contrast with the strongly-focussed laser field, it is assumed that the electron bunch does not change shape significantly as it crosses the interaction region.

If the particle scatters twice, it scatters first into an intermediate state u_1, then into the final stat u_f. The density of electrons for the second scattering process must be found by integrating

the scattering probability for the first process up to the time of the second scattering event. The number of double-scatters is thus given by

$$\begin{aligned}
\frac{dN}{du}(\{n_1,n_2\};u_i,u_f) = & \\
&\int du_1 \int d^3x \rho_{e0}(\vec{x}) \int_{t_2=-\infty}^{t_2=+\infty} dt_2 \rho_l(\vec{x}+\vec{v}t_2,t_2)(\eta^2(\vec{x}+\vec{v}t_2,t_2))^{n_2-1} \\
&\quad \times(1+\beta\cos\theta)\tfrac{d\sigma}{du}(n_2,\eta^2(\vec{x}+\vec{v}t_2,t_2);u_1,u_f) \\
&\quad \times \int_{t_1=-\infty}^{t_1=t_2} dt_1 \rho_l(\vec{x}+\vec{v}t_1,t_1)(\eta^2(\vec{x}+\vec{v}t_1,t_1))^{n_1-1} \\
&\quad \times(1+\beta\cos\theta)\tfrac{d\sigma}{du}(n_1,\eta^2(\vec{x}+\vec{v}t_1,t_1);u_i,u_1) \\
= & \int d^3x \rho_{e0}(\vec{x}) \int_{t_2=-\infty}^{t_2=+\infty} dt_2 \int_{t_1=-\infty}^{t_1=t_2} dt_1 (1+\beta\cos\theta)^2 \\
&\quad \times\rho_l(\vec{x}+\vec{v}t_2,t_2)(\eta^2(\vec{x}+\vec{v}t_2,t_2))^{n_2-1} \\
&\quad \times\rho_l(\vec{x}+\vec{v}t_1,t_1)(\eta^2(\vec{x}+\vec{v}t_1,t_1))^{n_1-1} \\
&\quad \times\tfrac{d\sigma}{du}(\{n_1,n_2\},\{\eta^2(\vec{x}+\vec{v}t_1,t_1),\eta^2(\vec{x}+\vec{v}t_2,t_2)\};u_i,u_f)
\end{aligned} \tag{B.11}$$

where the double-scattering "cross-section" is defined as

$$\tfrac{d\sigma}{du}(\{n_1,n_2\},\{\eta_1^2,\eta_2^2\};u_i,u_f) \equiv \int du_1 \tfrac{d\sigma}{du}(n_1,\eta_1^2;u_i,u_1)\tfrac{d\sigma}{du}(n_2,\eta_2^2;u_1,u_f) \tag{B.12}$$

The double-scattering cross-section has units of area squared, and a factor of $(1+\beta\cos\theta)^2$ appears in the rate calculation. In the more general case of scattering m times through processes $\{n_1,n_2,\ldots n_m\}$, we have

$$\begin{aligned}
\tfrac{d\sigma}{du}(\{n_j\},\{\eta_j^2\};u_i,u_f) = & \int du_1 \ldots du_{m-1} \prod_{j=2}^{j=m-1} \tfrac{d\sigma}{du}(n_j,\eta_j^2;u_{j-1},u_j) \\
& \times\tfrac{d\sigma}{du}(n_1,\eta_1^2;u_i,u_1)\tfrac{d\sigma}{du}(n_m,\eta_m^2;u_{m-1},u_f)
\end{aligned} \tag{B.13}$$

$$\begin{aligned}
&\frac{dN}{du}(\{n_j\};u_i,u_f) = \\
&\quad \int d^3x \rho_{e0}(\vec{x}) \underbrace{\int dt_m \cdots \int dt_1}_{t_{j-1}<t_j} \tfrac{d\sigma}{du}(\{n_j\},\{\eta_j^2\};u_i,u_f) \prod_{j=1}^{j=m} \rho_{lj}(\eta_j^2)^{n_j-1}
\end{aligned} \tag{B.14}$$

where η_j^2 and ρ_{lj} are short-hand for $\eta^2(\vec{x}+\vec{v}t_j,t_j)$ and $\rho_l(\vec{x}+\vec{v}t_j,t_j)$, respectively.

As with single-scattering, if $\tfrac{d\sigma}{du}(\{n_j\},\{\eta_j^2\};u_i,u_f)$ does not change appreciably over the range of η^2 of interest, then one can approximately describe the process by defining a luminosity function.

$$\frac{dN}{du}(\{n_j\};u_i,u_f) \cong \tfrac{d\sigma}{du}(\{n_j\};u_i,u_f)\mathcal{L}(\{n_j\}) \tag{B.15}$$

$$\mathcal{L}(\{n_j\}) \equiv \int d^3x \rho_{e0}(\vec{x}) \underbrace{\int dt_m \cdots \int dt_1}_{t_{j-1}<t_j} (1+\beta\cos\theta)^m \prod_{j=1}^{j=m} \rho_{lj}(\eta_j^2)^{n_j-1} \tag{B.16}$$

The m-scattering "cross-section" has units of (area)^m, so the m-scattering "luminosity" has units of (area)$^{-m}$. The m-scattering cross-section in principle contains only the "physics" of the leading-order behavior of the process, but again any higher-order intensity-dependence of the cross-section will be convolved with parameters of the experiment prior to observation.

B.3 Derivation of luminosity for non-linear and multiple scattering processes as a function of relative beam overlap

Calculating the total number of scatters or the "luminosity" as defined above is, in principle, as simple as inserting the functional forms for $\rho_{e0}(\vec{x})$ and $\rho_l(\vec{x'},t)$ into equation B.14 or B.16 above, then performing the integrals. If the number of scatters is to be calculated exactly in the case where the higher-order, non-linear intensity dependence of the generalized cross-section is significant, one must also insert the appropriate cross-section into equation B.14.

In practice, a scattering rate prediction for a given set of experimental parameters is most easily obtained by either performing the integrals numerically or by tracking particles through the interaction region in a Monte-Carlo code. However, when fitting the "X-T scan data" from the experiment, MINUIT or any other fit optimization package will require repeated predictions for various different parameters for each point in the scan. If there are 2000 samples in a given scan, then 2000 predictions have to be generated just to calculate a single chi-squared or likelihood value; MINUIT can easily call the chi-squared or likelihood function over 500 times given the number of parameters to fit. Therefore, even if our Monte-Carlo or numerical integration code could be optimized such that it took only a minute to run, it could take almost two years to fit a single X-T scan. Although the execution time might be reduced by using more sophisticated techniques (or more powerful computers), it is obviously highly desirable to have an analytic expression for the generalized luminosity in a closed form.

Although with this experiment we have pushed η^2 to values approaching (and occasionally exceeding) 1, most of the volume in which interactions occur has much lower values of η^2, and so the observed rates tend to be dominated by processes at lower η^2. Therefore, we still expect to be able to extract the leading-order coefficients (*i.e.* the cross-sections) of the scattering processes, and certainly the experimental parameters such as beam sizes and timing overlaps, using the assumption of an essentially constant generalized cross-section, as defined above.

As the peak value of η^2 is increased, we expect a systematic change of our "leading order" cross section as determined by the fit under this model. This is caused by the higher order non-linearities in the cross-section, and as mentioned before it will not be possible to completely deconvolve that behavior from parameters specific to our experiment. Comparison at that level can only be done directly between Monte-Carlo and experiment, or perhaps between "fitted" Monte-Carlo and "fitted" experiment. However, by isolating the leading-order behavior from the specifics of the electron and laser beam distributions in space and time, it is hoped that the higher-order behavior can be regarded as somehow "less convolved" with the uninteresting mundanities of the experiment.

The crossing of the beams in the $\hat{z} \times \hat{x}$ plane of two coordinate systems is shown in figure B.1.

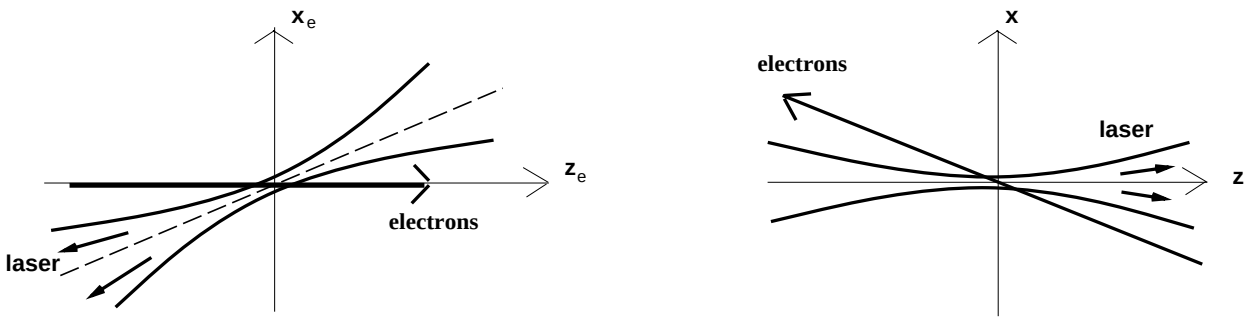

Figure B.1: Coordinate systems: electron-aligned (left) and laser-aligned (right)

The distribution of the electrons $\rho_{e0}(\vec{x_0})$ can most easily be described in the coordinate system on the left, while the shape of the laser field $\rho_l(\vec{x},t)$ can be most easily described in the coordinate system on the right.

In the "laser-aligned" coordinate system, the density of photons at point $\vec{x}=(x,y,z)$ and time t is given by

$$\rho_l(x,y,z,t) = \frac{N_l \exp(-\frac{1}{2}\frac{(z-t)^2}{\sigma_{lz}^2} - \frac{1}{2}\frac{(x/\sigma_{lx})^2+(y/\sigma_{ly})^2}{1+(z/z_R)^2})}{(2\pi)^{3/2}\sigma_{lx}\sigma_{ly}\sigma_{lz}(1+(z/z_R)^2)} \tag{B.17}$$

where σ_{lx} and σ_{ly} are the RMS widths of the spot at the focus, σ_{lz} is the RMS bunch length, and z_R is the effective Rayleigh length of the focus, known as $\beta_\star$ to accelerator physicists. Table B.1 summarizes few useful definitions and relations taken from Meyerhoffer's technical notes on Gaussian, non-diffraction limited laser focussing [Meyerhoffer, 1994 and 1995] and adapted to my notation.

In the "electron-aligned" coordinates, a Gaussian electron bunch will be described simply by

$$\rho_{e0}(\vec{x_b}) = \frac{N_e\sqrt{\det(\mathbf{V_{eb}}^{-1})}}{(2\pi)^{3/2}} \exp(-\frac{1}{2}(\vec{x_b}-\vec{x_{b0}})^T \mathbf{V_{eb}}^{-1}(\vec{x_b}-\vec{x_{b0}})) \tag{B.18}$$

where in the simplest case of a beam with no skew, the covariance matrix $\mathbf{V_{eb}}$ is simply

$$\mathbf{V_{eb}}^{-1} = \begin{bmatrix} 1/\sigma_{ex}^2 & 0 & 0 \\ 0 & 1/\sigma_{ey}^2 & 0 \\ 0 & 0 & 1/\sigma_{ez}^2 \end{bmatrix} \tag{B.19}$$

The vector $\vec{x_{b0}}$ is the location of the centroid of the electron bunch at time $t=0$. The subscript b on $\vec{x_b}, \vec{x_{b0}}, \mathbf{V_{eb}}^{-1}$ indicate the "electron-aligned" coordinates.

Since the laser's geometry is more complicated than the that of the electron bunch, the calculation will be performed using the "laser-aligned" coordinate system shown on the right of figure B.1, at least initially. It is easy to perform a passive rotation of $\mathbf{V_e}$ to obtain

$$\rho_{e0}(\vec{x}) = \frac{N_e\sqrt{\det(\mathbf{V_e}^{-1})}}{(2\pi)^{3/2}} \exp(-\frac{1}{2}(\vec{x}-\vec{x_0})^T \mathbf{V_e}^{-1}(\vec{x}-\vec{x_0})) \tag{B.20}$$

$$\mathbf{V_e}^{-1} = \mathbf{R}(\theta)\mathbf{V_e'}^{-1}\mathbf{R}(-\theta)$$

Table B.1: Some relations between various optics parameters

$$
\begin{aligned}
\lambda &\equiv \text{wavelength of the laser} \\
w_0 &\equiv 1/e^2 \text{ radius of intensity at focussing optic} \\
\sigma_{lx}, \sigma_{ly} &\equiv \text{RMS width } (1/\sqrt{e} \text{ radius) of intensity at focus} \\
A_{\text{eff}} &\equiv 2\pi\sigma_{lx}\sigma_{ly} \\
\sigma_0 &\equiv \text{diffraction-limited RMS width of intensity at focus} \\
A_0 &\equiv 2\pi\sigma_0^2 \\
\alpha &\equiv A_{\text{eff}}/A_0 \\
z_{R0} &\equiv \text{diffraction-limited Rayleigh length} \\
z_R &\equiv \text{effective Rayleigh length as used in equation B.17} \\
f &\equiv \text{focal length of the focussing optic} \\
f_\# &\equiv f/(2w_0) \\
A_0 &= \frac{2}{\pi}(f_\#\lambda)^2 \\
z_{R0} &= \frac{4}{\pi}f_\#^2\lambda \;=\; \sqrt{\frac{8}{\pi}A_0}\, f_\# \\
z_R &= \sqrt{\alpha}\, z_{R0} \;=\; \sqrt{\frac{8}{\pi}A_{\text{eff}}}\, f_\#
\end{aligned}
$$

$$= \begin{bmatrix} \cos^2\theta\sigma_{ex}^{-2} + \sin^2\theta\sigma_{ez}^{-2} & 0 & \cos\theta\sin\theta(\sigma_{ex}^{-2} - \sigma_{ez}^{-2}) \\ 0 & \sigma_{ey}^{-2} & 0 \\ \cos\theta\sin\theta(\sigma_{ex}^{-2} - \sigma_{ez}^{-2}) & 0 & \sin^2\theta\sigma_{ex}^{-2} + \cos^2\theta\sigma_{ez}^{-2} \end{bmatrix}$$

The most complicated features of the integrand in the luminosity integral (eq. B.16) arise from the "z" dependance of the laser intensity given the short Rayleigh length of the focus. The argument to the ρ_l and η^2 functions is $\vec{x} + \vec{v}t_j$. Therefore, as written, we would have to substitute $z \to z - \beta\cos\theta\, t_j$ and $x \to x + \beta\cos\theta\, t_j$ into the right hand side of equation B.17, and it would be messy to evaluate the integrals over x, z, and $t_1, t_2, ...t_m$. A change of variable is in order. Define

$$z'_j = z - \beta\cos\theta\, t_j \tag{B.21}$$

Then equation B.16 becomes

$$\mathcal{L} = \int d^3x \rho_{e0}(\vec{x}) \underbrace{\int dz'_m \cdots \int dz'_1}_{z'_{j-1} < z'_j} \left(\frac{1+\beta\cos\theta}{\beta\cos\theta}\right)^m \prod_{j=1}^{j=m} \rho_{lj}(\eta_{lj}^2)^{n_j - 1} \tag{B.22}$$

where

$$\rho_{lj} = \rho_l\left(x + \tan\theta(z - z'_j), y, z'_j, \frac{z - z'_j}{\beta\cos\theta}\right) \tag{B.23}$$

and η_{lj}^2 is related to ρ_{lj} through equation B.2.

Now the integrand is simply a Gaussian in x, y, and z, and something more complicated with respect to $z'_1 \ldots z'_m$. Note that since the limits of the z'_j do not depend on (x, y, z), the Gaussian integrals over those three variables could be performed first. However, this would simply delay the inevitable reckoning with the non-Gaussian integrals; also, integration over (x, y, z) couples the z'_j integrals, and makes it hard to tell which features of the luminosity profile are due to the laser's shape and which are due to the size of the electrons. So it makes sense to attack the integrals over z'_j now.

We can write the integrand for each integral over z'_j as

$$\begin{aligned} L_j &= \left(\frac{1+\beta\cos\theta}{\beta\cos\theta}\right)\rho_{lj}(\eta_{lj}^2)^{n_j-1} \\ &= \left(1 + \frac{1}{\beta\cos\theta}\right)\left(\frac{3\sigma_T\lambda_l}{4\pi\alpha}\right)^{n_j-1}\left(\frac{N_l}{(2\pi)^{3/2}\sigma_{lx}\sigma_{ly}\sigma_{lz}}\right)^{n_j} \\ &\quad \times \frac{1}{\left(1 + (z'_j/z_R)^2\right)^{n_j}} \exp(-\frac{1}{2}n_j\chi_j^2) \end{aligned} \tag{B.24}$$

where

$$\begin{aligned} \chi_j^2 &= \frac{\left(z'_j - (z + x/\tan\theta)\right)^2}{(\sigma_{lx}/\tan\theta)^2(1 + (z'_j/z_R)^2)} + \frac{\left(z'_j - z/(1+\beta\cos\theta)\right)^2}{\sigma_{lz}^2(\beta\cos\theta)^2/(1+\beta\cos\theta)^2} \\ &\quad + \frac{y^2}{\sigma_{ly}^2(1 + (z'_j/z_R)^2)} \end{aligned} \tag{B.25}$$

Table B.2: Range of parameters for E144 experiment

Parameter	Minimum (μm)	Maximum (μm)
σ_{lx}, σ_{ly}	1	4
z_R	23	95
σ_{ex}, σ_{ey}	30	130
σ_{lz}	120	360
σ_{ez}	600	1100
x_0	-300	300
y_0	-150	150
z_0	-2000	2000

The term in the exponent χ_j^2 is approximately quadratic in z_j' over any range of z_j' that is small compared to z_R, so $\exp(-\chi_j^2/2)$ is a Gaussian with width $(\sigma_{lx}/\tan\theta)\sqrt{1+(z_j'/z_R)^2}$. This width is in fact small compared to z_R provided $\left|z_j'\right|/z_R \ll \tan\theta z_R/\sigma_{lx} = 4\tan\theta f_\#$. Therefore, the integrand is a narrow Gaussian times a slowly varying function, and so the integral can be well approximated as the integral of the Gaussian times the value of the function at the center of the Gaussian, provided the above condition holds.

In the paragraph above, it was assumed that the first term in equation B.25 dominates over the second, which is always the case if the above condition holds and $\sigma_l z > z_R$. This means that the center of the Gaussian is simply $z + x/\tan\theta$ in this case.

The range of size and position parameters of the E144 experiment is shown in table B.2. The f-number for the E144 laser optics is 6, and the crossing angle is 17 degrees, so the condition for the width of the Gaussian to be small compared to z_R is $|z + x/\tan\theta|/z_R \ll 7.3$. Unfortunately, in the worst case $|z + x/\tan\theta|$ can be 3000 μm, while z_R can be as small as 23 μm. Fortunately, the function $1/(1+(z'/z_R)^2)^n$ varies very slowly with z' when $|z'|/z_R \gg 1$. The correct condition for our approximation to be valid is that the fractional change of the "slowly-varying" function over the width of the Gaussian be small:

$$\begin{aligned}
\left|\frac{\partial}{\partial z'}\left(1+(z'/z_R)^2\right)^{-n}\right| \frac{(\sigma_{lx}/\tan\theta)\sqrt{1+(z_j'/z_R)^2}}{(1+(z'/z_R)^2)^{-n}} & \\
= \; n\left(1+(z'/z_R)^2\right)^{-1/2} \frac{2z'\sigma_{lx}}{z_R^2\tan\theta} & \;\ll\; 1 \\
\frac{z'}{\sqrt{z'^2+z_R^2}} \;\ll\; \frac{2}{n}\tan\theta f_\# & \qquad \text{(B.26)}
\end{aligned}$$

As long as n is not too large, the approximation should be good over all values of z_j', provided we use the correct width and center for the Gaussian. In fact, the width at large values of z' is smaller than $(\sigma_{lx}/\tan\theta)\sqrt{1+(z_j'/z_R)^2}$, so the approximation is even more accurate than would be indicated by (B.26) above. The width can be calculated exactly without any trouble: the inverse square of the width is

$$\sigma_{z'}^{-2} = \sigma_{lx}^{-2}\tan^2\theta/\left(1+(z'/z_R)^2\right) + \sigma_{lz}^{-2}\left(1+1/(\beta\cos\theta)\right)^2 \qquad \text{(B.27)}$$

At z' very far from 0, the correct center of the Gaussian is not $z + x/\tan\theta$, but rather $z/(1 + \beta\cos\theta)$. However, the fractional difference of the "slowly varying function" $\left(1 + (z'/z_R)^2\right)^n$ is small between these two values of z' at large values of z'. Therefore, we can use $z + x/\tan\theta$ as the value to "plug in to" the slowly varying function $\left(1 + (z'/z_R)^2\right)^n$.

Finally, there is the issue of the limits on the integrals over z'_j in equation B.22. We are approximating the integrand as a simple Gaussian times a constant, and each independant Gaussian in z'_j is centered around the same z'_0. In the space of the z'_j coordinates, the center of the m-dimensional Gaussian is on the symmetry axis of the limits of integration. The integrand is completely symmetric about that point. Therefore, the value of the m-dimensional integral with the limits $z_{j-1} < z_j$ is simply the integral over all space divided by the symmetry factor $(m!)$.

One final interesting fact: the value $z + x/\tan\theta$ is equal to $x_b/\sin\theta$, where x_b is the "x" position in the electron-aligned coordinates shown in figure B.1 and used in equation B.18. This is convenient, since after integrating over the z'_j, the next integrals to perform will fold in the electron beam distribution.

After performing all of the integrals over z'_j, we again obtain a pseudo-Gaussian form. It is useful here to assume $\sigma_{lz} \gg \sigma_{lx}$.

$$\mathcal{L} = \int dx_b\, dy_b\, dz_b\, \overline{\mathcal{N}} \exp(-\frac{1}{2}\chi^2) \tag{B.28}$$

with

$$\begin{aligned}
\overline{\mathcal{N}} &= \frac{N_e}{(2\pi)^{3/2}\sigma_{ex}\sigma_{ey}\sigma_{ez}} \left(\frac{N_l}{(2\pi)^{3/2}\sigma_{lx}\sigma_{ly}\sigma_{lz}}\right)^N \left(\frac{3\sigma_T\lambda_l}{4\pi\alpha}\right)^{N-m} \\
&\quad \times \left((1 + \frac{1}{\beta\cos\theta})\frac{\sigma_{lx}}{\tan\theta}\right)^m \left(1 + \left(\frac{x_b}{z_R\sin\theta}\right)^2\right)^{m/2-N} \\
&\quad \times \frac{(2\pi)^{m/2}}{m!(\prod n_j)^{1/2}} \qquad \text{(B.29)} \\
\chi^2 &= \frac{(z_b - (\beta\csc\theta + \cot\theta)x_b)^2}{\beta^2\sigma_{lz}^2/N} + \frac{y_b^2}{\sigma_{ly}^2\left(1 + \left(\frac{x_b}{z_R\sin\theta}\right)^2\right)/N} \\
&\quad + \frac{(x_b - x_0)^2}{\sigma_{ex}^2} + \frac{(y_b - y_0)^2}{\sigma_{ey}^2} + \frac{(z_b - z_0)^2}{\sigma_{ez}^2} \qquad \text{(B.30)}
\end{aligned}$$

and $N \equiv \sum n_j$. The subscript b indicates the "electron-aligned" coordinates.

In fact, this integrand is precisely Gaussian in z_b and y_b, so those integrals are easily performed. The final integral over x_b will be the most complicated yet, so that will be extracted into a separate function. The result can be written in the following form:

$$\mathcal{L} = \bar{\mathcal{N}}_l \exp(-\frac{1}{2}C)\, f(x_0 - x_s,\, \sigma_x,\, z_R \sin\theta,\, 1 + N\sigma_{ey}^2/\sigma_{ly}^2,\, N - (m+1)/2) \tag{B.31}$$

where the f function is defined as

$$f(x, \sigma, x_R, \gamma, n) \equiv \int dx' \frac{\exp\left(-\frac{1}{2}(x-x')^2/\sigma^2\right)}{(1+(x'/x_R)^2)^n \left(\gamma + (x'/x_R)^2\right)^{1/2}} \tag{B.32}$$

In the case $\sigma_{lx} << \sigma_{ex}$, the various constants are

$$x_s = \frac{\sin^2\theta(N\sigma_{ez}^2 + \beta^2\sigma_{lz}^2)x_0 + \sin\theta(\beta+\cos\theta)\sigma_{ex}^2 z_0}{\sin^2\theta(N\sigma_{ez}^2 + \beta^2\sigma_{lz}^2) + N(\beta+\cos\theta)^2\sigma_{ex}^2} \tag{B.33}$$

$$\begin{aligned}\overline{\mathcal{N}_l} &= \frac{N_e}{(2\pi)^{3/2}\sigma_{ex}\sigma_{ey}\sigma_{ez}} \left(\frac{N_l}{(2\pi)^{3/2}\sigma_{lx}\sigma_{ly}\sigma_{lz}}\right)^N \\ &\quad \times \frac{(2\pi)^{m/2+1}(\sigma_{lx}/\tan\theta)^m \sigma_z \sigma_{ey}}{m!(\prod n_j)^{1/2}} \\ &\quad \times \left(\frac{3\sigma_T\lambda_l}{4\pi\alpha}\right)^{N-m} \left(1 + \frac{1}{\beta\cos\theta}\right)^m \end{aligned} \tag{B.34}$$

$$\sigma_z = \frac{\sigma_{ez}\sigma_{lz}}{\sqrt{\sigma_{lz}^2 + N\sigma_{ez}^2}} \tag{B.35}$$

$$\sigma_x = \frac{\sigma_{ex}\sqrt{\sigma_{lz}^2 + N\sigma_{ez}^2}}{\sqrt{\sigma_{lz}^2 + N\sigma_{ez}^2 + N\tan^2\theta\,\sigma_{ex}^2}} \tag{B.36}$$

$$C = \frac{y_0^2}{\Sigma_y^2} + \frac{(z_0 - \frac{\beta+\cos\theta}{\sin\theta}x_0)^2}{\Sigma_z^2} \tag{B.37}$$

$$\Sigma_y^2 = \sigma_{ey}^2 + \left(1 + \frac{x^2}{(z_R\sin\theta)^2}\right)\sigma_{ly}^2/N \tag{B.38}$$

$$\Sigma_z^2 = \sigma_{ez}^2 + \beta^2\sigma_{lz}^2/N + \left(\frac{\beta+\cos\theta}{\sin\theta}\right)^2 \sigma_e x^2 \tag{B.39}$$

The factor $\exp(-C/2)$ clearly shows the Gaussian shape of the luminosity function as y_0 and z_0 are scanned, as well as the x_0 dependance of the center of the timing curve. The width Σ_y of the y_0 scan is exactly what one would guess intuitively. The width of the timing scan Σ_z is also what one would expect, given the correlation between x and z. The horizontal size of the laser σ_{lx} does not appear in the timing scan width due to the assumption of $\sigma_{lx} \ll \sigma_{ex}$ above, which is okay since σ_{lx} is so small compared to σ_{ez} at any place of interest, even many Rayleigh lengths away from focus. Note that the timing scan width is slightly different for different order processes, since the RMS width of ρ_l^N is smaller than the width of ρ_l, by a factor of $\sqrt{N}$.

The f function gives the x_0 dependance of the luminosity. The integral which defines it is a little more tricky to evaluate than our previous "pseudo-Gaussian" integrals, since the width of the Gaussian factor in the integrand $\sigma_x \approx \sigma_{ex}$ is actually larger than the foreshortened Rayleigh parameter $x_R \equiv z_R \sin\theta$. When the center of the Gaussian factor is far from zero, the integral is still well approximated by the area of the Gaussian times the value of $(1+(x/x_R)^2)^{-n}$, as in the "pseudo-Gaussian" approximations made above. However, when the center of the wide Gaussian factor is near zero, it is the Gaussian that is nearly constant, while the other factor is highly peaked; in this case, the area under the $(1+(x/x_R)^2)^{-n}$ term is a more accurate approximation. (Special handling is required for the case where $n = 0$.) Luckily, it is not necessary to integrate over the f function. It can be evaluated numerically or approximated as an appropriately weighted average of the two limiting cases described above. Empirically, I have found that

$$f \approx \frac{f_{x\ll\sigma}\exp(-x^2/(2\sigma^2)) + f_{x\gg\sigma}\exp(-2\sigma^2/x^2)}{\exp(-x^2/(2\sigma^2)) + \exp(-2\sigma^2/x^2)} \tag{B.40}$$

very closely approximates the result of numerical integration over a wide range of parameters.

Note that different processes can *in principle* be partly distinguished by their different behavior as x_0 is scanned. The integral that defines the f function has $(1+x^2/x_R^2)$ raised to the $N-(m+1)/2$ power in the denominator, where N is the sum of all the n_j in the scattering process, and m is the number of scatters. In fact, while more than one combination of scattering steps can lead to the same value of $N-(m+1)/2$, only one process at a given N can have the value $N-1$, and that is the single-scattering process, $m = 1$. Multiple-scattering processes with the same N will have a lower value of $N-(m+1)/2$, so their x_0 dependance will be broader. This should make it possible to extract the cross-section of the "coherent" n-photon scattering process even though the "incoherent" processes scatter electrons into the same part of the spectrum. Whether this is practical to do with E144 data already taken has not yet been determined, although initial results are encouraging. The fit seems stable, not susceptible to false minima or large correlations between parameters; furthermore, it does seem to be able to properly distinguish between double $n = 1$ scattering and single $n = 2$ scattering with Monte-Carlo data.

The general form given by equation B.31 along with the definitions B.32 through B.40 will allow us to fit X-T scan data in a physically meaningful way. This will give us some overlap-based measurement of the beam parameters, as well as a proper estimate of the scattering rate at best overlap. Better yet, it "should" be possible to extract the invariant leading-order cross-sections for coherent n-photon Compton scattering.

Appendix C

Volkof's solution

Synopsis

A derivation is given of the Volkof solution to the Dirac equation in the presence of a plane wave. This derivation proceeds from the first-order Dirac equation, in contrast to other treatments that proceed from the second order equation.

C.1 Introduction

Let us assume, for the sake of argument, that we are infinitely good at mathematics, such that we can immediately find all possible solutions to any given, well-posed mathematical problem. Then quantum electrodynamics (QED) provides a complete and wholly transparent description of the interactions of electrons, positrons, and light. For QED tells us that for any given set of external boundary conditions (such as the states of an electron beam and a laser beam prior to entering an interaction region), we need merely solve for the electron-positron wavefunction ψ and the electromagnetic (EM) 4-potential field A_μ which minimizes the action S, defined as

$$S = \int \mathcal{L}_{\text{QED}}\, d^4x, \tag{C.1}$$

where

$$\mathcal{L}_{\text{QED}} \equiv \bar{\psi}(i\partial_\mu\gamma^\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu, \tag{C.2}$$

and $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor. This least-action principle can be transformed into the following set of differential equations (by application of the Euler-Lagrange equations and Noether's theorem):

$$\begin{array}{lcccll}
(\hat{p}_\mu\gamma^\mu - eA_\mu\gamma^\mu - m)\psi & = & 0 & & & \text{(Dirac's equation)} \\
\partial_\mu F^{\mu\nu} & = & e\bar{\psi}\gamma^\nu\psi & = & j^\nu & \text{(Maxwell's inhomogeneous equations)} \\
\partial_\mu j^\mu & = & 0 & & & \text{(current conservation)} \\
\partial_\mu \epsilon^{\mu\nu}{}_{\alpha\beta}F^{\alpha\beta} & = & 0 & & & \text{(Maxwell's homogeneous equations),}
\end{array}$$

where $\hat{p}_\mu \equiv i\partial_\mu$, and j^ν is the 4-vector current density. It is in solving these equations for arbitrary boundary conditions that we are required to be infinitely good at math. In general, exact solutions to these equations have been found only for special cases. The usual approach is to start with the solution for some simple case (e.g., propagation through free space with no sources present) and use perturbation theory to find the solution to the actual problem to adaquate accuracy.

C.2 Volkof's solution

Volkof found a solution to Dirac's equation which allows the magnitude of an externally-applied EM field A_{ext} to be arbitrarily large but requires it to be a plane wave:

$$A_{\text{ext}}(x) = A_{\text{ext}}(\phi) \tag{C.3}$$

$$\phi \equiv (kx), \tag{C.4}$$

where A_{ext}, k, and x are 4-vectors, and $(kx) \equiv k^\nu x_\nu$. Requiring the EM wave to be a plane wave reduces the hard parts of the problem to single-dimensional differential equations which can be integrated directly. Note that $A(\phi)$ does not necessarily have to be a periodic function of ϕ, let alone a sinusoid: the A function could represent incoherent thermal radiation, for example, as long as it propagates along the direction k. Furthermore, k does not necessarily have to be the wavevector of the wave even A does happen to represent a coherent and narrowband wave. All physical results will be invariant with respect to an overall scaling of k.

The Volkof solution also requires that the 4-potential satisfy the condition

$$(kA_{\text{ext}}) = 0. \tag{C.5}$$

This implies the Lorentz gauge and a sourcefree interaction region, and also eliminates a component of a physically-meaningless constant term in the potential.

Volkof's solution to Dirac's equation for an electron in a plane EM wave has the form

$$\psi_p(x) = e^{-i(px)} F(\phi), \tag{C.6}$$

where p is a constant 4-vector, which will turn out to be the 4-momentum of the electron prior to entering the interaction region ($A_{\text{ext}} \overset{t\to-\infty}{\longrightarrow} 0$), and $(p)^2 = m^2$. Inserting this trial solution into Dirac's equation, one obtains

$$(\not{p} - e\not{A}_{\text{ext}} - m)F + i\not{k}F' = 0, \tag{C.7}$$

where $\not{p} \equiv p^\mu \gamma_m u$. For a wave travelling at the speed of light, $\not{k}\not{k} = (kk) = 0$, and so

$$\not{k}(\not{p} - e\not{A}_{\text{ext}} - m)F = 0. \tag{C.8}$$

This equation can be satisfied by a solution of the form

$$F = (1 + \frac{e}{2(kp)}\not{k}\not{A}_{\text{ext}})ue^{iS}, \tag{C.9}$$

where u is a Dirac spinor and S is a function of ϕ for which we will solve. Inserting this hindsight-inspired "trial solution" into equation C.8, making use of the identity $\not{k}\not{A}\not{k}\not{A} = -\not{A}\not{k}\not{k}\not{A} = 0$, and remembering that p^μ is just some constant 4-vector and not an operator,

$$\begin{aligned}
\not{k}(\not{p} - e\not{A}_{\text{ext}} - m)F &= \left(\not{k}\not{p} - e\not{k}\not{A} - m\not{k} + \frac{e}{2(kp)}\not{k}\not{p}\not{k}\not{A}\right) ue^{iS} \\
&= \left(\not{k}\not{p} - e\not{k}\not{A} - m\not{k} + \frac{e}{2(kp)}(2(kp) - \not{p}\not{k})\not{k}\not{A}\right) ue^{iS} \\
&= (\not{k}\not{p} - m\not{k})\, ue^{iS} \\
&= (2(kp) - \not{p}\not{k} - m\not{k})\, ue^{iS} \;=\; 0 \\
u &= \frac{(\not{p} + m)\not{k}}{2(kp)} u
\end{aligned}$$

$$(\not{p} - m)u \;=\; \frac{((p)^2 - m^2)\not{k}}{2(kp)}u \;=\; 0. \tag{C.10}$$

Therefore, the form of the wavefunction given in equations C.6 and C.9 requires the spinor u to be simply the usual spinor for momentum p. Now substituting back into equation C.7 and solving for S,

$$S'(\phi) \;=\; \frac{e^2}{2(kp)}(A_{\rm ext}(\phi))^2 - \frac{e(pA_{\rm ext}(\phi))}{(kp)} \tag{C.11}$$

$$S(\phi) \;=\; \int_{-\infty}^{\phi} d\phi'\, S'(\phi') \tag{C.12}$$

The complete Volkof solution is

$$\psi_{{\rm V};p}(x) = \left(1 + \frac{e\not{k}\not{A}_{\rm ext}}{2(kp)}\right) u e^{-i(px)+iS(\phi)}. \tag{C.13}$$

The average momentum of the electron in the strong EM wave, the so-called "quasi-momentum," is

$$\begin{aligned} q^\mu \equiv \langle \bar{\psi}_{\rm V}\hat{p}^\mu\psi_{\rm V}\rangle \;&=\; p^\mu - \langle S'\rangle k^\mu + ik^\mu \frac{e}{2(kp)}\bar{u}\not{k}\langle\not{A}'\rangle u \\ &=\; p^\mu + \frac{m^2}{2(kp)}\eta^2 k^\mu, \end{aligned} \tag{C.14}$$

where

$$\eta^2 \;\equiv\; \left\langle -e^2(A_{\rm ext}(\phi))^2 + 2e(pA_{\rm ext}(\phi)) + \frac{ie}{4}F_{\lambda\nu}\bar{u}\sigma^{\lambda\nu}u \right\rangle. \tag{C.15}$$

If the time-averaged value of $A_{\rm ext}(\phi)$ is zero, then

$$\eta^2 \;=\; -e^2\left\langle (A_{\rm ext}(\phi))^2 \right\rangle. \tag{C.16}$$

Note that when the time-averaged value of $A_{\rm ext}$ is zero, the invariant η^2 depends only on the properties of the wave, i.e., the wave's intensity and wavelength. All processes occuring in the strong EM wave will have rates that depend on η^2; nonlinearity increases with increasing η^2.

The square of the quasi-momentum is called the "effective mass":

$$m_*^2 \equiv (q)^2 = (1+\eta^2)m^2. \tag{C.17}$$

The physical significance of the quasi-momentum q^μ and the effective mass m_* is that the kinematics of scattering involve q^μ rather than p^μ, and that a center-of-mass energy of at least $2m_*$ is needed to produce e^+e^- pairs in photon-photon interactions.

References

[1] C. Bula *et al.*, Phys. Rev. Lett. **76**, 3116 (1996).

[2] Theofilos Kotseroglou, "Observation of Nonlinear Compton Scattering", Ph.D. Thesis, University of Rochester UR-1459 (1996).

[3] K. Shmakov, Ph.D. thesis, University of Tennessee (1997).

[4] D. Burke *et al.*, Phys. Rev. Lett. **79**, 1626 (1997).

[5] S. Boege, Ph.D. thesis, University of Rochester (1996).

[6] T. Koffas, Ph.D. thesis, University of Rochester (1998).

[7] N.D. Sengupta, Bull. Math. Soc. (Calcutta) **44**, 175 (1952).

[8] H.R. Reiss, J. Math. Phys. **3**, 59 (1962); Phys. Rev. Lett. **26**, 1072 (1971).

[9] A.I. Nikishov and V.I. Ritus, Soviet Physics JETP **19**, 529 (1964).

[10] N.B. Narozhnyi, A.I. Nikishov and V.I. Ritus, Soviet Physics JETP **20**, 622 (1965).

[11] V.B. Berestetskii, E.M. Lifshitz and L.P. Pitaevskii, *Quantum Electrodynamics*, 2nd ed. (Pergamon Press, New York, 1982), secs. 40 and 101.

[12] C.I. Moore, J.P. Knauer, and D.D. Meyerhofer, Phys. Rev. Lett. **74**, 2439 (1995).

[13] A. Belkacem *et al.*, Phys. Lett. **B177**, 211 (1986); Phys. Lett. **B206**, 561 (1988); R. Medenwalt et al., Phys. Lett. **B227**, 483 (1989).

[14] D.M. Volkov, Zeits. f. Phys. **94**, 250 (1935).

[15] A.I. Nikishov and V.I. Ritus, Soviet Physics JETP **25**, 1135 (1967).

[16] Y.S. Tsai, Phys. Rev. D **48**, 96 (1993).

[17] A.I. Nikishov and V.I. Ritus, Soviet Physics JETP **20**, 757 (1965).

[18] C. Bamber *et al.*, "0.5 Hz, Phase-Stabilized Terawatt Laser System with Nd:Glass Slab Amplifier for Nonlinear QED Experiments", University of Rochester Report UR-1428, submitted to Laser Physics.

[19] V. Balakin *et al.*, Phys. Rev. Lett. **74**, 2479 (1995).

[20] M. Berndt, *et al.*, *Final Focus Test Beam: Project Design Report*, SLAC report 376 (1991).

[21] Robert Holtzapple, "Longitudinal dynamics at the Stanford Linear Collider", Ph.D. Thesis, Stanford University; SLAC-0487 (1996).

[22] E. Prebys, Possible Readout for the E-144 Calorimeters. Internal Note. (1992).

[23] S.H. Rokni, *et al.*, SLAC publication SLAC-PUB-6784-REV, 1996; submitted to Health Phys.

[24] C. Field, Nucl. Instrum. Methods A **360**, 467 (1995).

[25] T. Kotseroglou *et al.*, "Picosecond Timing of Terawatt Laser Pulses with the SLAC 46 GeV Electron Beam", SLAC-PUB-7130; Nucl. Instr. and Meth. A **383**, 309 (1996).

[26] C. Bula, "NUMINT", E-144 internal pub. (1997).

[27] P. Chen and R.B. Palmer, in Proceedings of the Advanced Accelerator Concepts Workshop, Port Jefferson, NY, 888 (1992).

[28] I.F. Ginzburg, *et al.*, JETP Lett. **34**, 491 (1981); Pisma Zh. Eksp. Teor. Fiz. **34**, 514 (1981).

[29] V. Telnov, Nucl. Instr. Meth. A **294**, 72 (1990).

[30] The NLC Design Group, "Zeroth-Order Design Report for the Next Linear Collider, Volume I", SLAC-474/LBNL-5424/UCRL-ID-124161/UC-414 (1996).

[31] R.C. Field, "Monitor calibration", in *Proceedings of the [Informal] E-144 Collaboration Meeting, SLAC, January, 1997*, E-144 internal pub., p. 29 (1997).

[32] W.R. Nelson *et al.*, "The Electron - Gamma Shower (EGS4) Code System," SLAC-Report-265, Dec. 1985.

[33] D.D. Meyerhofer, "Summary of Gaussian focusing," E-144 technical note (1994).

[34] D.D. Meyerhofer, "Non-diffraction limited focusing," E-144 technical note (1995).

[35] H.J. Larson, *Introduction to Probability Theory and Statistical Inference*, third edition, Wiley, New York (1982).

[36] C. Field, "The wire scanner system of the Final Focus Test Beam", SLAC-PUB-6717 (1994); Nucl. Instrum. Meth. **A360**, 467 (1995).

[37] R.M. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996).

[38] J.D. Jackson, *Classical Electrodynamics*, 2nd ed. (J. Wiley, New York, 1975), p. 695.

[39] K.T. McDonald, "Features of the CCD Spectrometer, VII: Gamma-Line Configuration for Dec. 1995", E-144 internal note.

[40] F.W.Aitchley and R.O. Dobbs, *Geological Investigation of the Stanford Two-Mile Linear Accelerator Site*, Stanford (1960).

[41] R.B. Neal, *et al.*, *The Stanford Two-Mile Accelerator*, Stanford (1968).

[42] G. Bowden, *et al.*, "Precision magnet movers for the Final Focus Test Beam," SLAC-PUB-95-6132 (1995); Nucl. Instrum. Meth. **A368**, 579 (1996).

[43] G. Horton-Smith, "Calibration of the E-144 readout of toroids I6 and I7", E-144 internal memo (1995).

[44] D. Walz, private communication.

[45] L. Hendrickson, *et al.*, "Generalized fast feedback system in the SLC", SLAC-PUB-5683 (1991); also in *Proceedings of the 1991 International Conference on Accelerators and Large Experimental Physics Control Systems at Tsukuba, Japan*, KEK (1991).

[46] M. Woods, *et al.*, "Measurement of the longitudinal polarization of the Final Focus Test Beam," E-144 internal report (1994).

[47] M. Breidenbach, *et al.*, "SLC Hardware Manual", SLAC Controls department internal pub. (1988).

[48] G. Horton-Smith, "Chi-squared of March '95 data", in *Proceedings of the [Informal] E-144 Collaboration Meeting, SLAC, September 9, 1996*, E-144 internal pub., p. 68 (1996).

[49] K. Shmakov, Ph.D. thesis, U. Tennessee (1997).

[50] W.H. Press *et al.*, *Numerical Recipes*, 2nd ed., (Cambridge Univ. Press, 1996).

[51] T. Kotseroglou, Ph.D. thesis, U. Rochester (1996).

[52] S. Boege, Ph.D. thesis, U. Rochester (1996).

[53] D. Meyerhoffer, private communication (1997).

[54] K. Shmakov, private communication (1998).

[55] D. Burke, private communication (1998).

Internal references:

[Chapter 1] G. Horton-Smith, "Study of Nonlinear QED in Intense Laser-High Energy Electron Collisions", PhD thesis, chap. 1, Stanford (1998).

[Chapter 2] G. Horton-Smith, "MCSCAT: A Monte-Carlo Particle Scattering Simulation", E-144 internal note (1998).

[Chapter 3] G. Horton-Smith, "Electron beam diagnostics and controls, and more", E-144 internal note (1998).

[Chapter 4] G. Horton-Smith, "Cross-calibration of E-144 Cherenkov monitors with the electron calorimeter", E-144 internal note (1998).

[Chapter 5] G. Horton-Smith, "Estimation of laser focal properties using Cherenkov monitor data", E-144 internal note (1998).

[Chapter 6] G. Horton-Smith, "Features of the E-144 CCD Tracking Spectrometer and Monte-Carlo Simulation", E-144 internal note (1998).

[Chapter 7] G. Horton-Smith, "CCD Track Reconstruction for E-144", E-144 internal note (1998).

[Chapter 8] G. Horton-Smith, "Results from the measurement of the nonlinear Compton photons", E-144 internal note (1998).

[Appendix A] G. Horton-Smith, "E144 ECAL 'Best Aperture' Reconstruction", E-144 internal note (1998).

[Appendix B] G. Horton-Smith, "Nonlinear luminosity integrals and application to overlap scans", E-144 internal note (1998).