

PHOTOPRODUCTION OF ETA MESONS FROM DEUTERIUM AT  $0^\circ$  AND  $180^\circ$   
WITH PHOTON ENERGY FROM 725 MEV TO 1225 MEV

Thesis by  
Steven J. Yellin

In Partial Fulfillment of the Requirements  
For the Degree of  
Doctor of Philosophy

California Institute of Technology  
Pasadena, California

1971

(Submitted February 26, 1971)

## ACKNOWLEDGMENTS

This experiment was suggested and supervised by C. A. Heusch, who rescued me from theoretical physics. It was performed jointly with W. A. McNeely, Jr., who began the construction and calibration of the Cherenkov counters before I joined him. The two of us share responsibility for many achievements and blunders. Charles Prescott helped in the running, and more importantly, he introduced us to the computer used for monitoring this experiment. Such people as Steven Cheng, Robert Kline, Bruce Winstein, Kirk McDonald, Ernest Jan, Lily Yeh and Leon Rochester also helped run this experiment.

The synchrotron crew under Paul Van Ligten and the synchrotron operators led by Al Neubieser did much of the labor necessary to keep this experiment alive, and they made the laboratory a pleasant place in which to work. Special thanks for their extra efforts are due to Bill Friedler, Earle Emery, Dick Wileman and Walter Nilsson.

Personal financial support was provided by the National Science Foundation and the California Institute of Technology. The synchrotron laboratory itself was financed by the Atomic Energy Commission.

## ABSTRACT

In an experiment conducted at the Caltech 1.5 GeV electron synchrotron, we measured the cross section for photoproduction of eta mesons from deuterium. The measurement was performed by detecting and measuring the energies of both photons from the  $\eta \rightarrow 2\gamma$  decay mode using two totally absorbing lead glass Cherenkov counters. These counters were placed symmetrically about the photon beam line in order to best detect eta mesons photoproduced along the beam line. By varying the placement of the Cherenkov counters and by varying the synchrotron end-point energy, we were able to obtain information on the forward and backward cross section for photoproduction over an energy range from about 725 MeV to about 1225 MeV.

Within the framework of the impulse approximation, we expressed our results as a sum of the differential cross section for eta photoproduction from protons plus that from neutrons. The unfolding of the cross section was performed by finding the function that was the smoothest (using a well-defined measure of smoothness) function that fit the data to within a given degree of accuracy (as measured by  $\chi^2$ ).

The results show that for photon energy below 950 MeV, the cross section for photoproduction from neutrons is about equal in the forward and backward directions, and is about the same as the cross section for photoproduction from protons. Above 950 MeV, a significant angular asymmetry is seen, and the neutron cross section shows signs of reaching a minimum, then rising. These results are inter-

preted in terms of contributing isobars.

## TABLE OF CONTENTS

<u>PART</u>	<u>TITLE</u>	<u>PAGE</u>
I.	INTRODUCTION	
	A. Symmetry Considerations	1
	B. Resonances and Exchanges	6
II.	EXPERIMENTAL METHOD	
	A. General Considerations	17
	B. The Primary Beam	23
	C. The Deuterium Target	27
	D. The Cherenkov Counters	29
	E. Electronics	39
III.	DATA ANALYSIS	
	A. Raw Data	52
	B. Accidental Background	56
	C. Background Subtraction	65
	D. Corrections to the Yield	81
	E. Extracting the Cross Section from the Yields	84
IV.	RESULTS	93
V.	CONCLUSIONS	103
VI.	APPENDICES	
	A. Gain Matching	107
	B. Monte Carlo Method	112
	C. Optimizing Resolution	115

<u>PART</u>	<u>TITLE</u>	<u>PAGE</u>
	D. Inside the Black Box	123
	E. Maximum Likelihood Method	134
	F. Synchrotron End-point Correction	145
	G. Kinematics	149
	H. Details of the Cherenkov Counters	167
	I. Error Analysis	172
VII.	REFERENCES	182

## LIST OF FIGURES

<u>FIGURE</u>	<u>TITLE</u>	<u>PAGE</u>
1.	Experimental Layout	21
2.	Condensing Deuterium Target	28
3.	Cherenkov Counter Calibration	37
4.	Spectrum of Radioactive Source	38
5.	Electronics Block Diagram	49
6.	Oscilloscope Displays	50
7.	Timing Distribution	57
8.	Average Slewing Correction	61
9.	Calibration of Time Digitizer	62
10.	Distribution of Pulse-heights for a Typical Run	70
11.	Division of the i-j Plane	72
12.	Computed Background for a Typical Run	76
13.	Pulse-height Distribution Minus the Estimated Background	77
14.	Our Best Estimate of the Foreground	78
15.	Pulse-height Distribution Minus the Combined Estimated Foreground and Background	79
16.	Differential Cross Section for Eta Photopro- duction from Deuterium. No Threshold Factor, $\chi^2 = 20$	99
17.	Differential Cross Section for Eta Photo-	100

<u>FIGURE</u>	<u>TITLE</u>	<u>PAGE</u>
	production from Deuterium. Phase Space Threshold Assumed. $\chi^2 = 20$	
18.	Differential Cross Section for Eta Photo- production from Deuterium. Phase Space Threshold Assumed. $\chi^2 = 60$	101
19.	Apparent Number of Photoelectrons as a Function of Shower Energy	111
20.	Effect of Varying Synchrotron End-point	148
21.	Impulse Approximation	163
22.	Form Factor for Hulthen Function	164
23.	Hulthen Wave Function	165
24.	Relationship Between Form Factor and Counter Half-angle	166
25.	Cherenkov Counter	169
26.	Phototube Placement	171

## LIST OF TABLES

<u>TABLE</u>	<u>TITLE</u>	<u>PAGE</u>
I.1	SU(3) Derived Relationships for Resonances	7
II.1	Optimal Counter Half-angles, $\theta$ , for Detection of Etas Produced by Photons of Energy K	22
II.2	Ion Chamber Calibration Constant	26
II.3	Circuits Used in This Experiment	51
III.1	Settings for Which Data Were Collected	54
IV.1	Cross Section from Optimizing Resolution	102
VI.1	Summary of Lead Glass Characteristics	170
VI.2	Tests of Background Fitting Procedure	181

To my mother and sister

## I. INTRODUCTION

## A. Symmetry Considerations

At the time we decided to perform this experiment on  $\gamma + d \rightarrow \eta + n + p$ , much information on the reaction  $\gamma + p \rightarrow \eta + p$  below 1.5 GeV photon energy was available or about to become available (see especially reference 22 and references therein). The cross section had been measured at a few angles to within about a ten percent accuracy from slightly above threshold to 1.45 GeV with a density of about one measurement every 15 MeV. A measurement of the polarization of the recoil proton<sup>(22)</sup> was being analyzed. Information was also becoming available on cross sections for eta photoproduction from protons at energies higher than the 1.5 GeV to which we were limited at Caltech. But no measurements had yet been made on eta photoproduction from neutrons. Moreover, work had already begun on a  $\gamma + p \rightarrow \eta + p$  experiment with kinematics such that the recoil proton could not be detected. Merely replacing the hydrogen target with a deuterium target allowed us to examine eta photoproduction from neutrons. As it turned out, a group at Frascati was also working on  $\gamma + n \rightarrow \eta + n$ .<sup>(23)</sup> Their measurements gave cross sections at two energies around 850 MeV and at center-of-mass angles  $-.4 < \cos(\theta) < .6$ . We proposed to map out the cross section for the same reaction, but at  $\cos(\theta) \approx \pm 1$  and over the energy range from threshold to as high an energy as we could successfully reach with the methods, equipment and time available. The reader deserves an explanation of how such information can be valuable.

We begin by writing down a general expression for photoproduction to lowest order in the electromagnetic interaction. Consider the reaction

$$\gamma + a \rightarrow b.$$

The amplitude for this reaction to lowest order is of the form

$$\text{Amplitude} \propto e^\mu \langle b | j_\mu(\vec{k}) | a \rangle,$$

where  $e^\mu$  is the polarization four-vector of the photon (we use the notation  $x^\mu y_\mu \equiv \vec{x} \cdot \vec{y} \equiv x_\alpha y_\alpha - \vec{x} \cdot \vec{y}$ ),  $j_\mu(\vec{k})$  is the electromagnetic current operator and  $\vec{k}$  is the four vector photon momentum. Such matrix elements are non-zero only if the isospin of  $b$  differs from that of  $a$  by zero or one. From this observation, it follows that  $j$  can be made up of no more than an isoscalar component plus an isovector component. If we write  $j_\mu = j_\mu^V + j_\mu^S \equiv \text{isovector} + \text{isoscalar}$ , then it is a matter of simple group theory to obtain certain relations for photoproduction of pi mesons from nucleons.<sup>(24)</sup> Following Walker,<sup>(18)</sup> we use the notation  $A^{V3}$  for the part of the amplitude corresponding to  $\pi N$  final states with  $I=3/2$ ,  $A^{V1}$  for the part corresponding to  $I=1/2$  final states from the  $j^V$  part of the current, and  $A^S$  for the  $I=1/2$  final states from  $j^S$ .

Then

$$\gamma + p \rightarrow \pi^+ + n: \quad A^+ = \sqrt{1/3} A^{V3} - \sqrt{2/3} (A^{V1} - A^S),$$

$$\gamma + p \rightarrow \pi^0 + p: \quad A^0 = \sqrt{2/3} A^{V3} + \sqrt{1/3} (A^{V1} - A^S),$$

I A.1

$$\begin{aligned}
\gamma + n \rightarrow \pi^- + p: \quad A^- &= \sqrt{1/3} A^{V3} - \sqrt{2/3} (A^{V1} + A^S), \\
\gamma + n \rightarrow \pi^0 + n: \quad A^{0n} &= \sqrt{2/3} A^{V3} + \sqrt{1/3} (A^{V1} + A^S).
\end{aligned}
\tag{I A.1}$$

We write down the corresponding equations for photoproduction of etas.

$$\begin{aligned}
\gamma + p \rightarrow \eta + p: \quad A^{\eta+} &= A^{S\eta} - A^{V\eta}, \\
\gamma + n \rightarrow \eta + n: \quad A^{\eta 0} &= A^{S\eta} + A^{V\eta}.
\end{aligned}
\tag{I A.2}$$

Photoproduction of pions was historically studied first, and has been and will continue to be known better than eta photoproduction. But although eta photoproduction experiments are harder, in some respects they are more rewarding. The special simplicity of equations I A.2 arises from the fact that the final state can only be isospin 1/2, rather than both 1/2 and 3/2. In the case of pion photoproduction, in order to separate I=1/2 from I=3/2, two photoproduction amplitudes must be known, and amplitudes are usually hard to extract from experimental yields, which represent probabilities. The separation is performed automatically when the final state is an eta-nucleon system.

Let us now extend our discussion of symmetries to SU(3) (see for example reference 37). If we take the Fourier transform of  $j_0$  and integrate over all space we get the electric charge, Q, which in the language of SU(3) satisfies  $Q = I_z + Y/2 \equiv$  z-component of the isospin + half the hypercharge. This operator,  $Q = j_0(0)$ , is the U-spin zero member of an SU(3) octet of operators. If we assume  $j_\mu(\vec{k})$  also has U=0 (regardless of whether or not it is entirely octet) then

electromagnetic transitions cannot change  $U$ . As Lipkin<sup>(26)</sup> first pointed out, the member of the  $\underline{10}^*$  (anti-decuplet) isodoublet corresponding to the proton has  $U$ -spin =  $3/2$ , so cannot be photoproduced from the  $U=1/2$  proton. The member of the  $\underline{10}^*$  isodoublet corresponding to the neutron has  $U=1$ , and can be photoproduced. One reason, then, for our interest in photoproduction of  $I=1/2$  states from neutrons is the possibility of seeing evidence of a  $\underline{10}^*$ . Remember that the intermediate states of our experiment will all be  $I=1/2$  non-strange baryons. A  $\underline{10}^*$ , or for that matter anything other than an octet, would be surprising in our experiment because in any three-quark model of the baryons,  $I=1/2$  non-strange baryons must be members of octets.

If we move a little further along the path toward current algebra, we can suppose that the electromagnetic current is not only a  $U=0$  operator, but is a member of an  $SU(3)$  octet. We now can decompose the photoproduction amplitudes with respect to  $SU(3)$  representations in the same way used in the derivation of equations I A.1 and I A.2, but first we restrict ourselves to the case of resonances in the intermediate state. The intermediate states of photoproduction from nucleons must be limited to members of the possible irreducible representations of

$$\underline{8} \times \underline{8} = \underline{1} + \underline{8}^d + \underline{8}^f + \underline{10} + \underline{10}^* + \underline{27}.$$

The restriction to strangeness = zero, isospin =  $1/2$  excludes the  $\underline{1}$  and  $\underline{10}$  from consideration. Then simple manipulation of  $SU(3)$  Clebsch-Gordon coefficients at the photon vertex relates  $A^S$  to  $A^{V1}$  and relates  $A^{S\eta}$  to  $A^{V\eta}$ . Similar manipulation at the decay vertex

relates  $A^{V\eta}$  to  $A^{V1}$  and  $A^{S\eta}$  to  $A^S$ . The results, of course, depend on the representation to which the resonance belongs. Table I.1 summarizes the results.  $\alpha_\gamma/(1-\alpha_\gamma)$  is the D to F ratio at the photon vertex when the intermediate state is in an octet, and  $\alpha/(1-\alpha)$  is the D to F ratio at the decay vertex.  $\alpha_\gamma$  and  $\alpha$  can be expected to vary from one resonance to another. When an octet resonance is pure isovector (as Walker<sup>(18)</sup> obtains for the  $S'_{11}(1535)$ , the  $D'_{13}(1520)$  and the  $F_{15}(1688)$ ) then it means  $\alpha_\gamma \approx .75$ . When an octet resonance is observed in photoproduction of pi mesons, but not in photoproduction of etas, then for that resonance  $\alpha \approx .75$ .

For those components of the amplitude that are in a  $10^*$  or  $27$ , the ratios given in Table I.1 are correct regardless of whether or not a resonance dominates. Then equations I A.1 and I A.2 can have their  $10^*$  and  $27$  parts separated from their octet parts:

$$\begin{aligned}
 A^+ &= \sqrt{1/3} A^{V3} - \sqrt{2/3} (A_8^{V1} - A_8^S + 2A_{27}) , \\
 A^0 &= \sqrt{2/3} A^{V3} + \sqrt{1/3} (A_8^{V1} - A_8^S + 2A_{27}) , \\
 A^- &= \sqrt{1/3} A^{V3} - \sqrt{2/3} (A_8^{V1} + A_8^S + A_{10} - A_{27}) , \\
 A^{\text{on}} &= \sqrt{2/3} A^{V3} + \sqrt{1/3} (A_8^{V1} + A_8^S + A_{10} - A_{27}) , \\
 A^{\eta+} &= A_8^{S\eta} - A_8^{V\eta} + 6A_{27} , \\
 A^{\eta0} &= A_8^{S\eta} + A_8^{V\eta} + A_{10} + 3A_{27} .
 \end{aligned}
 \tag{I A.3}$$

For each octet resonance, the components of the octet amplitudes corresponding to that resonance still satisfy the relationships given in Table I.1. From equations I A.3 we see, as pointed out earlier,

that the  $\underline{10}^*$  does not contribute to photoproduction from protons. We also see that the  $\underline{27}$  component contributes more to eta photoproduction than to pi photoproduction and contributes more to photoproduction from protons than from neutrons.

### B. Resonances and Exchanges

Some of the intermediate states that we might expect in eta photoproduction are (from the Review of Particle Properties)<sup>(25)</sup>  $P'_{11}(1470)$ ,  $D'_{13}(1520)$ ,  $S'_{11}(1535)$ ,  $D_{15}(1670)$ ,  $F_{15}(1688)$ ,  $S''_{11}(1700)$ , and  $P''_{11}(1780)$ . Other particles are expected to be too high in mass to greatly affect this experiment. We will discuss these resonances one by one.

$P'_{11}(1470)$ , the Roper resonance, has not yet been shown to appear in photoproduction. It is probably below threshold for our reaction, but its mass is known very poorly. It was once suspected, in fact, that the large cross section near threshold is caused by the Roper. Although it now seems pretty certain that the  $S'_{11}$  dominates near threshold, the tail of the Roper may be able to influence the cross section even if the  $P'_{11}$  is below threshold. The effect of the Roper resonance must be suppressed by the angular momentum barrier. As Levi Setti<sup>(32)</sup> points out, the effect of the angular momentum barrier can be described qualitatively as follows: Maximum angular momentum allowable in outgoing state = center-of-mass momentum of an outgoing particle x maximum distance over which the strong interactions can act (one or two fermis). From this point of view, we expect the

TABLE I.1

SU(3) Derived Relationships for Resonances

Ratio	Representation of Resonance		
	$\underline{8}$	$\underline{10}^*$	$\underline{27}$
$\frac{A^5}{A^{VI}}$	$-\frac{3-4\alpha}{3}\gamma$	1	-3
$\frac{A^{S\eta}}{A^{V\eta}}$			
$\frac{A^{S\eta}}{A^S}$	$-\frac{3-4\alpha}{3}$		
$\frac{A^{V\eta}}{A^{VI}}$			

tail of the Roper to start having an effect when the photon lab energy exceeds 750 MeV (so that the eta has c.m. momentum of over 134 MeV/c and the center-of-mass energy is over 1513 MeV). The interference of the  $P'_{11}$  with the  $S'_{11}$  would cause a  $\cos(\theta)$  term in the center-of-mass cross section. Such a term is not observed in  $\gamma + p \rightarrow \eta + p$  (as of this writing) up to a photon energy of 865 MeV. See, for example, reference 21. If the Roper resonance were in an  $SU(3) \underline{10}^*$ , this absence of the  $\cos(\theta)$  term would be easily explained by U-spin conservation, as discussed in the previous section. But according to the quark model as used by Faiman and Hendry<sup>(28)</sup> (and later by Copley, Karl and Obryk),<sup>(29)</sup> the  $P'_{11}$  should be produced even less frequently from neutrons than from protons. If we see evidence of a Roper resonance in photoproduction from neutrons, it will at least contradict the version of the quark model used by Copley, Karl and Obryk. Such evidence would even suggest that the  $P'_{11}$  is in a  $\underline{10}^*$ , contradicting all models which make baryons from three quarks.

We now turn to a discussion of the  $D'_{13}(1520)$ , observed prominently in elastic pion-nucleon scattering (the "second resonance") and in  $\gamma + N \rightarrow N + \pi$ . Although the angular momentum barrier should strongly suppress the  $D'_{13}$  in  $\pi^- + p \rightarrow \eta + n$ , a  $D'_{13}$  term can be seen there as an interference with the nearby  $S'_{11}$ . Such interference produces a substantial  $\cos^2(\theta)$  term in the reaction  $\pi^- + p \rightarrow \eta + n$  at pion kinetic energy of 655 and 704 MeV.<sup>(20)</sup> These energies correspond to 805 and 855 MeV in eta photoproduction. But an experiment conducted at Orsay<sup>(21)</sup> found little or no  $\cos^2(\theta)$  dependence in  $\gamma + p \rightarrow \eta + p$  when the photon energy varied from 750 MeV to 865 MeV.

In order to see whether such an interference term is to be expected in  $\gamma + n \rightarrow \eta + n$ , we must try to understand why none was seen with a proton target.

Suppose an experiment is performed in which two unpolarized particles collide to produce one or more outgoing particles. We have in mind a cross section measurement or a measurement of the polarization of a recoil particle. In the center-of-mass, the results of the experiment must be independent of the azimuthal angle about the axis of the two unpolarized incoming particles (the z-axis). It follows that the interference between states of differing  $|J_z|$  cannot contribute to the results of such an experiment. For the  $D'_{13}$ ,  $|J_z| = 1/2$  or  $3/2$ . But it is known from partial wave analyses of  $\gamma + p \rightarrow \pi + N$ <sup>(17, 18, 19)</sup> that in photoproduction from protons the  $D'_{13}$  is produced primarily in a  $|J_z| = 3/2$  state. It follows that even if the  $D'_{13}$  contributes as large a fraction to the amplitude for  $\gamma + p \rightarrow \eta + p$  as it does to that for  $\pi^- + p \rightarrow \eta + n$ , in the photoproduction case the  $D'_{13}$  will not interfere with the  $J=1/2$  states to produce a  $\cos^2(\theta)$  term. Similarly it will not contribute to the recoil proton polarization. One way of detecting the  $D'_{13}$  in  $\gamma + p \rightarrow \eta + p$  would be to measure the polarized photon asymmetry or the polarized target asymmetry. In either case, only the interference between differing  $|J_z|$  contributes to the asymmetry.

At this point in our reasoning, the question is whether  $|J_z| = 1/2$  can be expected when the  $D'_{13}$  is photoproduced from neutrons. According to Copley, Karl and Obryk,<sup>(29)</sup> the answer is yes. From their use of a quark model, the cancellation of the  $|J_z| = 1/2$  that occurs in photo-

production of the  $D'_{13}^+$  should not occur in photoproduction of the  $D'_{13}^0$ . But consider, on the other hand, the argument of Bietti.<sup>(30)</sup> He concludes from current algebra considerations that

$$\text{a) } \frac{E_{2-}^{V1}}{M_{2-}^{V1}} = 2.7, \text{ and}$$

$$\text{b) } E_{2-}^S \approx 0.$$

$E_{2-}^{V1}$  and  $M_{2-}^{V1}$  are parts of the electric dipole and magnetic quadrupole moments that contribute to  $A^{V1}$  in equation I A.1.  $E_{2-}^S$  is a part of the electric dipole that contributes to  $A^S$ . Subscript "n±" denotes a part of the amplitude corresponding to a state of angular momentum  $j = n \pm \frac{1}{2}$  and parity  $-(-1)^n$ . The  $D'_{13}$  is then contained in "2-" multipoles. From equation I A.1, the statement that  $D'_{13}$  is not photo-produced with  $|J_z| = 1/2$  is equivalent to the statement that there is no component of  $A^{V1} - A^S$  corresponding to such a state, which implies that

$$3 = \frac{E_{2-}^{V1} - E_{2-}^S}{M_{2-}^{V1} - M_{2-}^S} \approx \frac{E_{2-}^{V1}}{M_{2-}^{V1} - M_{2-}^S} = \frac{2.7}{1 - M_{2-}^S / M_{2-}^{V1}}$$

(see Walker's paper<sup>(18)</sup> for definitions allowing one to relate the condition  $|J_z| = 3/2$  to relations between multipoles). Thus both  $E_{2-}^S$  and  $M_{2-}^S$  are small, and  $D'_{13}$  is produced primarily through the isovector current. From equation I A.1 it follows that if  $A^S$  is zero for production of the  $D'_{13}$ , then the  $D'_{13}$  should be produced with just as small a fraction of  $|J_z| = 1/2$  from a neutron target as from a

proton target. Our confidence in Bietti's work is encouraged by the phase shift analysis of Walker,<sup>(18)</sup> who finds that the  $D'_{13}$  is indeed produced primarily by isovector current and that the  $|J_z| = 1/2$  component of the  $D'_{13}$  is small for both proton and neutron targets. Copley, Karl, and Obryk recognize the contradiction between Walker's results and their theory, but point to more recent experimental evidence from DESY contradicting Walker's conclusions.<sup>(31)</sup> Walker himself admits that his fits based on photoproduction from neutrons "must be regarded as tentative."<sup>(18)</sup> Measurement of the  $\cos^2(\theta)$  term in eta photoproduction from neutrons would supply valuable evidence for resolving the above discussed theoretical and experimental contradictions. The Frascati experiment<sup>(23)</sup> in which the  $\gamma + n \rightarrow \eta + n$  cross section was measured at center-of-mass angles  $-.4 < \cos(\theta) < +.6$ , combined with our results at  $\cos(\theta) \approx \pm 1$ , can give a crude estimate of the  $\cos^2(\theta)$  term at energies appropriate for observing  $S'_{11} - D'_{13}$  interference.

Next we discuss the  $S'_{11}$ . Both the reactions  $\pi^- + p \rightarrow \eta + n$  and  $\gamma + p \rightarrow \eta + p$  show a strong enhancement in the cross section near threshold. In the former reaction, references 33 and 34 made it clear that the S-wave dominates. It is plausible, then, that the similar behavior of  $\gamma + p \rightarrow \eta + p$  near threshold is also explained by  $S'_{11}$  dominance. This view is supported by the phase shift analyses of references 35 and 36, where it is found that a dominant  $P'_{11}$  would be inconsistent with the data. See also Appendix F.

Over a rather small energy range, the Frascati group finds approximately the same cross section for  $\gamma + n \rightarrow \eta + n$  as for

$\gamma + p \rightarrow \eta + p$ . Suppose our more complete explanation of the energy range near threshold supports the belief that near threshold  $A^{\eta^0}$  is dominated by an  $S'_{11}$  of about the same magnitude as appears in  $A^{\eta^+}$ . Then by the symmetry considerations of the previous section, the  $S'_{11}$  is verified to be in an SU(3) octet. It is hard to say much about the value of  $\alpha$  for the  $S'_{11}$  because the effect of phase space depends so strongly on the poorly known  $S'_{11}$  mass and width. But if  $\alpha$  were near .75, we pointed out in the previous section that the amplitude for decay into the eta-nucleon system would be small compared to that into the pi-nucleon system. Because the phase space for the eta-nucleon system is so much smaller than that for the pi-nucleon system,  $\alpha$  near .75 would imply decay almost exclusively into a pi-nucleon state. The large  $S'_{11}$  enhancement near threshold of  $\gamma + p \rightarrow \eta + p$  and of  $\pi^- + p \rightarrow \eta + n$  excludes this possibility, and it is even likely that the majority of the  $S'_{11}$  decays are into an eta-nucleon state (see the baryon references given in the Review of Particle Properties<sup>(25)</sup>). As mentioned previously, Walker<sup>(18)</sup> tentatively finds that  $A^S/A^{V1}$  is small, so that for this resonance  $\alpha_\gamma$  is near .75.

Recall that for pion photoproduction off protons it is well established that for the  $D'_{13}$  the  $|J_z|=1/2$  amplitudes nearly vanish. From the point of view of Copley, Karl and Obryk, this fact is a consequence of the near cancellation of two quantities. In the approximation that the  $S'_{11}$  is produced at the same energy as the  $D'_{13}$ , exact cancellation for the  $D'_{13}$  would imply that the ratio of isoscalar to isovector contribution to the  $S'_{11}$  is

$$\frac{A^S}{A^{V1}} = \frac{A^{S\eta}}{A^{V\eta}} = \frac{\tan \theta + 2}{\tan \theta - 16} ,$$

where  $\theta$  is the mixing angle between the  $S_{11}$  with total quark spin  $1/2$  and the  $S_{11}$  with total quark spin  $3/2$ . The quark model experts have not yet been able to decide what the mixing angle should be. Faiman and Hendry<sup>(28)</sup> suggested  $35^\circ$ ; Copley, Karl and Obryk<sup>(29)</sup> preferred  $70^\circ$ . In either case,  $A^{S\eta}/A^{V\eta}$  is small.

From what has been said so far, both quark model and experiment agree that the  $S'_{11}$  is produced primarily in an isovector interaction. But in that case, we would expect  $\gamma + d \not\rightarrow \eta + d$  near threshold because deuterons are isoscalars. Anderson and Prepost,<sup>(38)</sup> however, find that the cross section for  $\gamma + d \rightarrow \eta + d$  is not zero, and in fact is large enough to be consistent with production entirely by an isoscalar interaction. This result favors an  $S_{11}$  mixing angle near  $90^\circ$ , and contradicts Walker. Furthermore, as Faiman and Hendry<sup>(28)</sup> point out,  $\theta$  near  $90^\circ$  implies the existence of an unacceptably large  $S''_{11}$  (1700).

In our experiment, not only do we hope to improve on the accuracy of the Frascati results, but we may have a handle on the relative sign between  $A^{\eta^+}$  and  $A^{\eta^0}$  (if the sign is positive, the reaction is mainly isoscalar; if the sign is negative, the reaction is mainly isovector). The eta can be photoproduced from either the proton or the neutron in the deuteron. But the final states for the two processes contain the same particles, so they can interfere. Constructive interference would indicate positive relative sign between  $A^{\eta^+}$  and  $A^{\eta^0}$ , and destructive interference would indicate negative relative sign. We will later

analyze this experiment in terms of the impulse approximation, but we can already see qualitatively that the interference should appear only in forward eta photoproduction. For backward photoproduction, the nucleon from which the eta was photoproduced would be moving rapidly down the beam line while the spectator would be moving relatively slowly in the lab. The final states with the proton as spectator could then be distinguished in principle from those with the neutron as spectator, and the two kinds of final states would not interfere. We will therefore look for evidence of an enhancement or depression in the forward cross section.

We have completed our discussion of the possible (known) intermediate states near threshold. The  $D_{15}(1670)$  lies in the upper part of the energy range covered by this experiment. Its absence from photoproduction from protons was first predicted by Moorhouse<sup>(40)</sup> as a consequence of a quark model. Ecklund and Walker<sup>(39)</sup> did find possible evidence for a  $D_{15}$  in positive pion photoproduction from protons, but its amplitude was small. We can expect a larger amplitude for pion photoproduction from neutrons, but it is possible that the D to F ratio for the  $D_{15}$  is such as to prevent its decay into the eta-nucleon channel. This effect occurs for the  $F_{15}$ .

The  $F_{15}(1688)$  is not observed in  $\gamma + p \rightarrow \eta + p$ , but it is seen in  $\gamma + p \rightarrow \pi^0 + p$ . Heusch, Prescott and Dashen<sup>(43)</sup> used SU(3) arguments corresponding to certain entries in Table I.1 to conclude from these facts that the  $F_{15}$  must be in an octet and must have  $.5 \lesssim \alpha \lesssim 1.0$ . For our purposes, we need only note that the  $F_{15}^0$  is photoproduced from neutrons far less frequently than the  $F_{15}^+$  is photoproduced from protons,<sup>(53)</sup>

contradicting the result of reference 18, which had the  $F_{15}$  photoproduced primarily by an isovector interaction. Because isospin conservation at the decay vertex implies that the amplitude for  $F_{15}^0 \rightarrow \eta + n$  is equal to that for  $F_{15}^+ \rightarrow \eta + p$ , the  $F_{15}$  should be even harder to see in eta photoproduction from neutrons than it is in eta photoproduction from protons.

The  $S_{11}''$  (1700) may appear in our experiment as an angular asymmetry caused by its interference with the  $P_{11}''$  (1780). In the framework of the quark model, Moorhouse<sup>(40)</sup> has predicted that if the mixing angle of the  $S_{11}$  resonances is small, the  $S_{11}''$  (1700) will be absent in photoproduction from protons and present in photoproduction from neutrons.

The  $P_{11}''$  (1780) is seen clearly in the reaction  $\pi^- + p \rightarrow \eta + n$  as a bump in the cross section and as an interference between its low energy tail and the high energy tail of the  $S_{11}'$  (1535).<sup>(20)</sup> This resonance seems to appear (somewhat less clearly) in  $\gamma + p \rightarrow \eta + p$ ,<sup>(42)</sup> so the angular asymmetry from such an interference could appear in our data.

We conclude this part with a brief discussion of the possible t-channel contributions. Charge conjugation limits us to such states as the  $\rho^0$ ,  $\omega$ ,  $\phi$  and  $B^0$  (1235). Assuming the usual  $\omega$ - $\phi$  mixing angle ( $\sin^2(\theta) = 1/3$ ) and assuming the electromagnetic current transforms under SU(3) like a U = zero member of an octet, the vector meson-photon-pseudoscalar meson coupling constants can be related:

$$g_{\phi\pi\gamma} = 0 \quad ,$$

$$g_{\omega\pi\gamma} = 3 g_{\rho\pi\gamma} \quad ,$$

$$g_{\rho\eta\gamma} = \sqrt{3} g_{\rho\pi\gamma} \quad ,$$

$$g_{\phi\eta\gamma} = -4/\sqrt{6} g_{\rho\pi\gamma} \quad ,$$

$$g_{\omega\eta\gamma} = \frac{1}{\sqrt{3}} g_{\rho\pi\gamma} \quad ,$$

where we have neglected  $\eta$ - $\eta^0$  mixing. Since the  $\phi$  doesn't couple strongly to the nucleon,  $\rho$  exchange is expected to dominate the eta photoproduction exchanges, and it contributes only to  $A^{V\eta}$ . Dar and Weisskopf<sup>(41)</sup> argue more carefully, come to a similar conclusion (ignoring the B(1235)), and successfully predict the high energy behavior of eta photoproduction from protons. Similar success should be anticipated when their methods are applied to  $\gamma + n \rightarrow \eta + n$ . Our experiment, however, covers only energies within the resonance dominated region. Although t-channel exchanges can produce interference effects at our energies, we will say no more about this topic because we don't believe we can reliably estimate such effects.

## II. EXPERIMENTAL METHOD

## A. General Considerations

Because it was our intention to learn about photoproduction of neutrons in the forward and backward directions, we were faced with certain serious experimental difficulties. The non-existence of free neutron targets required the use of nuclei with more than one nucleon. The deuteron is the simplest such nucleus, and has the special virtue of being loosely bound. But even in the deuteron there is considerable Fermi motion of the individual nucleons. The effect of this Fermi motion was, of course, a decrease in our kinematical resolution.

Another problem associated with photoproduction of etas from neutrons is the detection of the recoil particle. Since we were especially concerned with forward and backward angles, the recoil nucleon tended to be going down the beam line where it was very difficult to separate from the general background. A spectrometer is useless when the recoil particle is a neutron, and even for protons a spectrometer can't help if the recoil proton is moving too slowly to leave the target, as can occur in forward photoproduction.

In this experiment we ignored the recoil nucleon and concentrated on detecting the eta from its  $\eta \rightarrow 2\gamma$  mode of decay. If we could measure exactly the energy and direction of each decay photon, we would thereby be measuring the four-momentum of each such photon, and if the two photons came from a particle decay, we would know the four-momentum of the particle. The invariant mass of the two photon system would determine whether or not the event was an eta decay.

In practice it is difficult to accurately measure both the energy and direction of photons in the energy range associated with this experiment. We found that the accuracy of our measurement was barely adequate for determination of the number of etas among the events detected at a given experimental setting, and was not adequate for determining the momentum of individual etas.

Figure 1 shows the layout of the experiment. A beam of bremsstrahlung photons with endpoint energy somewhere between 700 and 1250 MeV was collimated, scraped, and magnetically swept clean of charged particles before passing through a liquid deuterium target. Some of the photons participated in nuclear reactions, including photoproduction of etas, but most passed through the target unaffected, and the beam energy was absorbed in an ion chamber and in the lead and cement that shielded the rest of the apparatus from the ion chamber.

In order to detect etas through their two photon decay mode, each photon had to pass through lead apertures, then through a veto counter, then go into a totally absorbing Cherenkov counter. The veto counter served to eliminate the background from charged particles. It also served to veto showers associated with photons converting in the lead aperture. The two Cherenkov counters with their vetoes, apertures, and shielding (both against radiation and magnetism) were placed symmetrically on either side of the beam upon rails on a platform. The summed signal from each Cherenkov counter was pulse height analyzed and the time between the signals from the two counters

was measured. The resulting three numbers were stored on magnetic tape by a PDP-5 on-line computer.

For a given eta momentum, the position of the counters on the rails was chosen so as to maximize the counting rate. Some useful kinematic facts for determining the appropriate setting and for separating etas from the background follow.

Let  $\theta$  be the half-angle between the two decay photons, and let  $\frac{dn}{d\theta}$  be the density of photon pairs for a given eta velocity,  $\beta$ .  $\gamma^{-2} = 1 - \beta^2$ . Then<sup>(1)</sup>

$$\frac{dn}{d\theta} = \frac{\cos \theta}{2\beta\gamma \sin^2 \theta \sqrt{\gamma^2 \sin^2 \theta - 1}}$$

When the etas we wished to detect were expected to have velocity  $\beta$ , the photon detectors were separated by half-angle such that  $\frac{dn}{d\theta}$  was large -- i.e.,  $\sin \theta = 1/\gamma$ , or  $\cos \theta = \beta$ . The photon detectors were symmetrically placed because we were interested in photoproduction along the beam line.

Call  $E_1$  and  $E_2$  the energies of the two photons detected.

The background contamination from detection of two photons from such processes as

$$\gamma + Z \rightarrow \pi^0 + Z$$

$$\gamma + Z \rightarrow \pi^0 + \gamma + Z$$

$$\gamma + Z \rightarrow 2\gamma + Z$$

$$\gamma + Z \rightarrow 2\pi^0 + Z \rightarrow 4\gamma + Z$$

$$\gamma + Z \rightarrow e^+ + e^- + Z,$$

with bremsstrahlung of the electron and positron, limited severely

our ability to obtain the eta yield at each setting. Etas were, however, separable from the background by virtue of the equation

$$m = 2 \sin \theta \sqrt{E_1 E_2}$$

where "m" is the invariant mass of the photon pair, and is the eta mass when the event comes from eta decay. Details of the background subtraction will be covered later in the section on data analysis.

The physical layout of the experiment was such as to permit  $\theta$  to be varied from below  $35^\circ$  to about  $102^\circ$ . If we neglect Fermi motion, for each incoming photon energy,  $k$ , there is associated an eta velocity  $\beta$  for forward eta production, and with that velocity is associated an optimal  $\theta = \arccos(\beta)$ . Similarly there is an optimal  $\theta$  for backward photoproduction. Table II.1 shows the optimal  $\theta$  as a function of  $k$  for forward and backward produced etas.

As we collected data, we saw that the background became large compared to the eta signal for the more extreme forward and backward angles physically allowed by our layout, so we collected data only for  $\theta$  values between  $35^\circ$  and  $97^\circ$ .

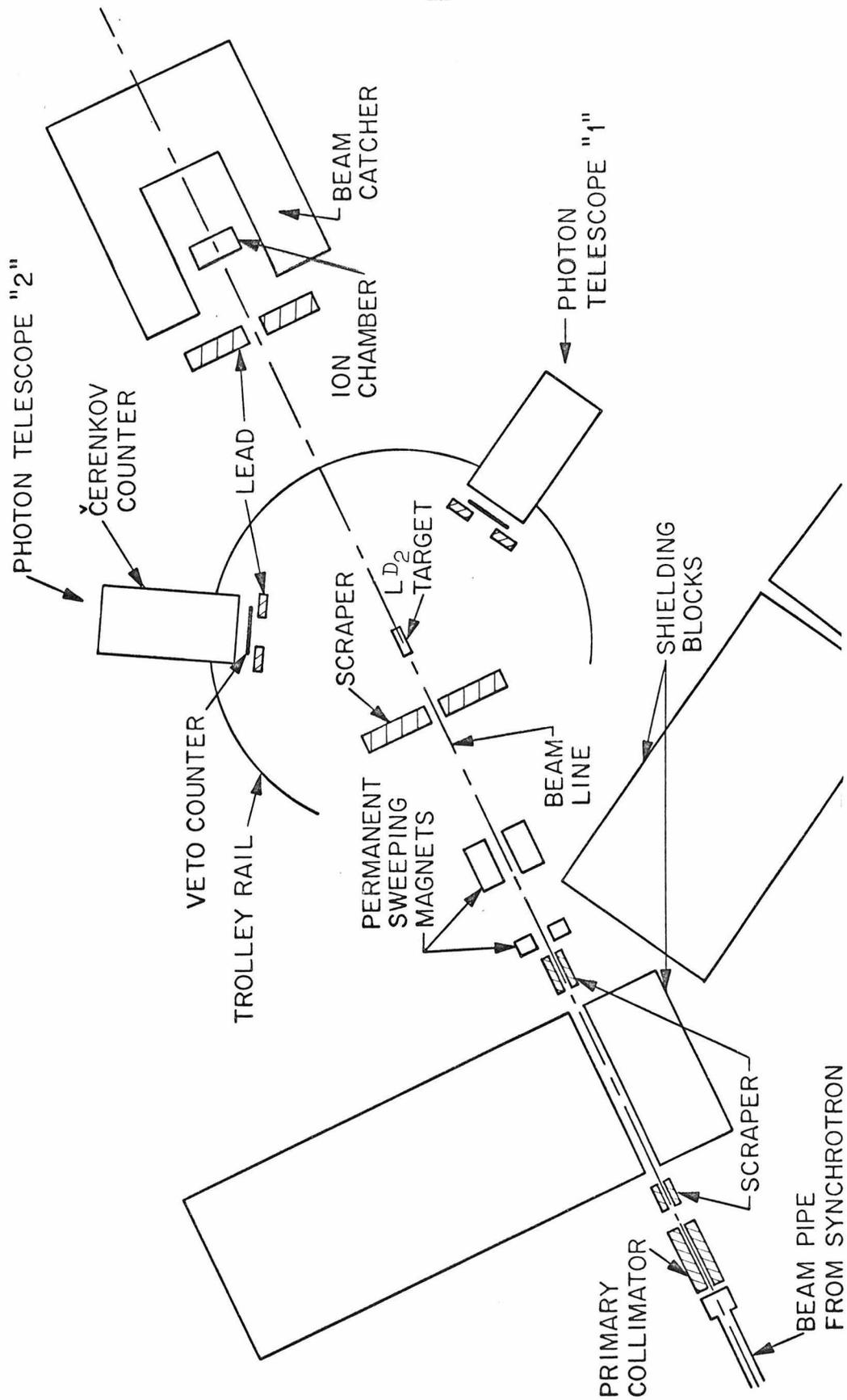


Figure 1. Experimental Layout.

TABLE II.1

Optimal Counter Half-Angles,  $\theta$ , for Detection  
of Etas Produced by Photons of Energy,  $k$ .

$k(\text{MeV})$	Forward	Backward
709	64	65
725	57	72
750	52	77
775	49	80
800	46	82
850	42	86
900	39	88
950	36	91
1000	34	93
1050	32	94
1100	30	96
1200	27	98
1300	25	100
1400	23	101
1500	22	103

## B. The Primary Beam

The incident photon beam was bremsstrahlung produced on a 1/32" tantalum radiator by electrons from the Caltech 1.5 GeV synchrotron.

The synchrotron was run at energies from below 710 MeV (threshold for eta production) up to about 1250 MeV, with energy monitored and regulated by measuring the synchrotron magnetic field during the dump. (2,3)

In each dump there were around  $10^{11}$  electrons hitting the radiator, causing about  $.5 \times 10^6$  photons per MeV to pass through the target. Often, unless the beam was unusually stable, these numbers would vary by typically a factor of two from one dump to the next. At energies below 925 MeV the dump was generally about 60 ms in length, while above 925 MeV it was stretched out to around 150 ms. The longer dump had the advantage of allowing fewer accidental coincidences in our logic, but could only be achieved by operating at one cycle per second, rather than two, resulting in a lower overall intensity. Furthermore, at some energies the magnetic field had a slope during the dump. This slope caused a slope in end point energy, and the resulting uncertainty in energy was proportional to the dump length. I should add, however, that for most runs the error in the synchrotron end point energy was mainly caused by the .5% uncertainty in the calibration of the beam energy meter. (3)

The size and shape of the beam spot at the target was determined by the primary collimator downstream from the tantalum

radiator. The diameter of the circular beam was measured (by placing photographic film in the beam) to be  $.76 \pm .02$ " at the position of the target. Two scrapers were used to clean off the spray from the primary collimator, and the beam was then swept clean of charged particles by a magnet upstream of a lead wall with an aperture. The hole in the wall was 1.4" in diameter, compared with a beam width at that point of about half that amount, so the wall served only to stop the particles swept out of the beam.

Since we wanted to measure a cross section, we had to accurately know the number and energy distribution of the photons passing through our target. The energy distribution was calculated by means of a program written here, BPAK I.<sup>(4)</sup>

To measure the total beam energy, an ion chamber was used in conjunction with an integrator of a type developed by R. Littauer.<sup>(5)</sup> In such an integrator, the charge from the ion chamber goes into a capacitor. The capacitor is placed across an input amplifier (which effectively amplifies the capacitance). When the capacitor reaches a certain voltage, an approximately equal charge of opposite sign is dropped on that capacitor from a precision capacitor which has been charged to a standard voltage. One full charge of the precision capacitor is defined to be a "bip" (beam integrator pulse). The amount of charge in a bip was measured at least once, and usually more than once, per day. This measurement was performed by dumping 200 pulses into the integrator input from another precision capacitor (of known capacitance) that was charged to a carefully measured

voltage. Typically, one bip would equal roughly  $2 \times 10^{-7}$  coulombs, with the exact number measureable for a given run to within better than .5%. Given the number of bips in a run, we were able to determine the total charge that had come from the ion chamber.

In order to find the integrated beam energy corresponding to that charge, we compared the ion chamber with a Wilson quantameter.<sup>(6,7)</sup> A Wilson quantameter contains nearly all the beam energy, numerically integrates the ionization produced in the showers, and compensates for shower loss through the sides. Arguments will soon be given for expecting totally absorbing Cherenkov counters to give light pulses proportional to shower energy. Similar arguments apply to the charge produced per shower in a Wilson quantameter. The charge per MeV can be theoretically calculated at least as well as it can be measured, and is independent of the energy of the photons producing the showers. For the dimensions, materials, and gas temperature and pressures in the quantameter we used, the calibration constant was  $5.78 \pm .18 \times 10^{18}$  MeV/coulomb.

The calibration of the ion chamber used to monitor our experiment was performed relative to the quantameter for several synchrotron endpoint energies. A thin ionization chamber was placed before the quantameter and the number of bips from the quantameter was measured for a certain amount of beam passing through the thin ion chamber (as measured by thin ion chamber bips). This procedure was repeated with the quantameter replaced by our ion chamber and with the same amount of beam through the thin ionization chamber.

The ratio between the charge from our chamber and that from the quantameter for the same integrated beam gives us the information we need to know to get our calibration constant. Unlike the Wilson quantameter, our ion chamber had a calibration constant that varied with synchrotron endpoint energy. This calibration constant is shown in Table II.2, where U is the number of MeV that must go into the chamber from a bremsstrahlung spectrum of endpoint energy E to produce one coulomb of charge.

TABLE II.2

## Ion Chamber Calibration Constant

$E_o$	$U \times 10^{18}$ (MeV/coulomb)
650	57.1
750	58.0
880	58.9
1000	60.0
1100	60.7

From what has been said above, it follows that one bip corresponded to about  $1.2 \times 10^{13}$  MeV. For use in analyzing our data, we standardize the bip to  $1.2132 \times 10^{13}$  MeV.

The calibration constant of our ion chamber was taken several times during our running, for the chamber had a slow leak and had to

be refilled every few months. In the several months between refillings the calibration constant declined 1 to 2%, as compared with a 3% uncertainty in the theoretical values used to determine the quantummeter calibration constant. In order to correct for the gradual change of the calibration "constant," for each run we corrected the calibration constant by assuming that it fell linearly with time between fillings.

### C. The Deuterium Target

Shown in Figure 2 is the condensing deuterium target used in this experiment. According to formulae given in Reference 16, hydrogen at a typical Pasadena pressure of 750 mm Hg boils at 20.2°K, while deuterium at the same pressure boils at 23.5°K, and at 20.2°K has a vapor pressure of only 243 mm Hg. By letting deuterium into our target at about atmospheric pressure, it could easily be made to condense by surrounding the target with liquid hydrogen. At the temperature of the liquid hydrogen, deuterium has a density of 0.171 g cm<sup>-3</sup>.

The hydrogen was allowed to boil off and escape into the atmosphere, but the deuterium, of course, was kept in a closed system. When the target was allowed (or forced) to heat up, the deuterium could escape into a storage tank.

In order to minimize the use of liquid hydrogen, the hydrogen dewar was surrounded by liquid nitrogen, and the entire system was kept well insulated from the rest of the laboratory by vacuum and

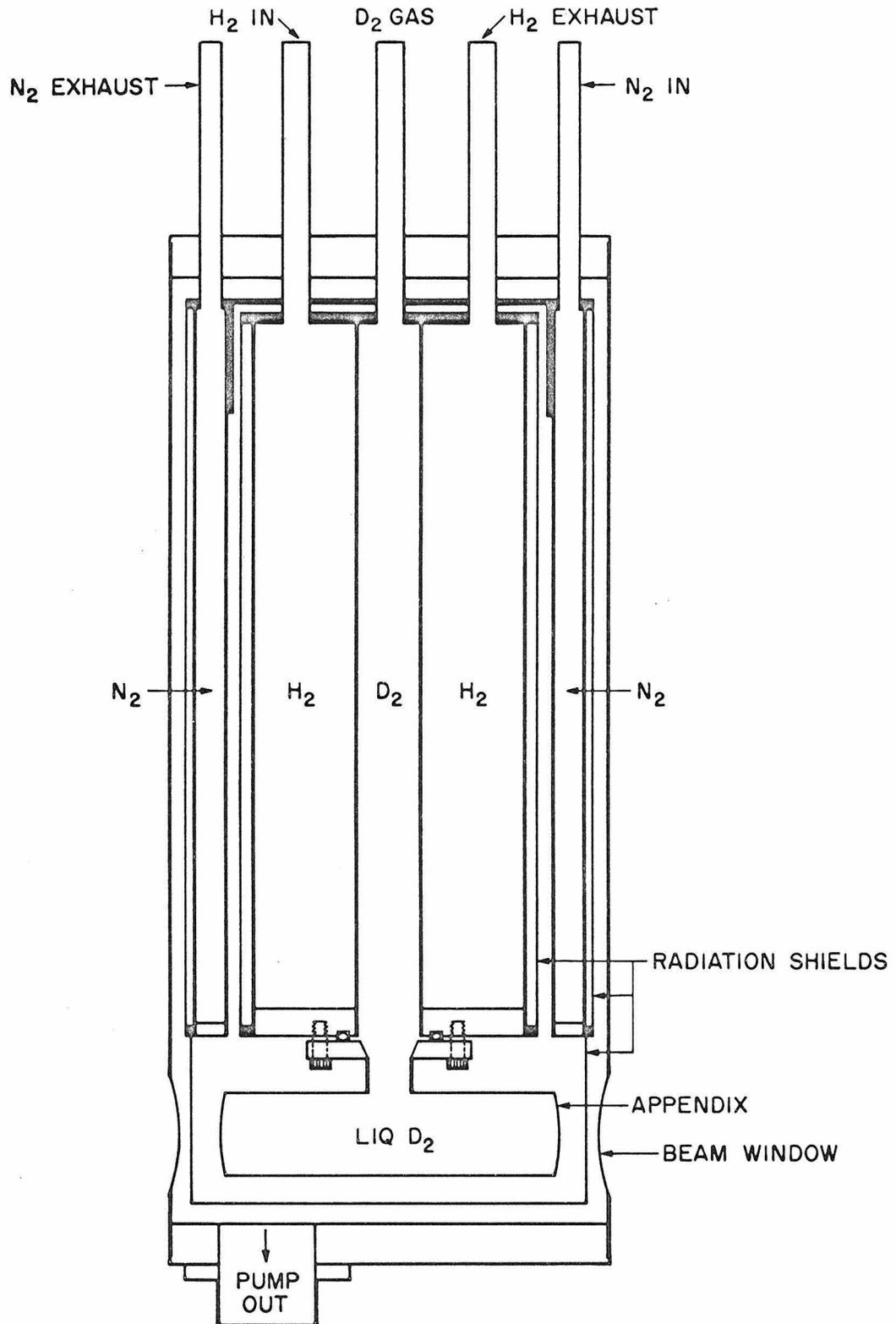


Figure 2. Condensing deuterium target.

highly reflecting aluminized mylar. For obvious reasons, the hydrogen and nitrogen baths did not extend down as far as the target itself, but only cooled the lower part of the tube leading from the deuterium storage tank to the target.

The composition of the "deuterium" used was measured before the experiment to be by weight 97.3% deuterium, .4% hydrogen, and 2.3% other (nitrogen and oxygen).

In the course of the experiment, three target lengths were used. We started with 6.639"; then, when a leak developed in the first target, we put in a new one of length 6.390"; and finally, for better kinematic resolution, we did some runs with a target of length 3.270".

In addition to deuterium, the incoming photons had to pass through a total of sixteen mils of mylar. The main obstruction in the target to the outgoing eta decay photons was the 1/16 inch aluminum outer case of the target (.018 radiation lengths).

#### D. The Cherenkov Counters

The photons from the eta decay were detected by two totally absorbing Cherenkov counters - one called PbG-1, the other called PbG-2. These counters are described in Appendix H. Here we describe the principles by which they work and some of the problems associated with their use.

Suppose we have an electron - or photon - induced shower in some material. Let  $P(E)$  be the probability per unit  $E$  that a given

small piece of the total track length of charged particles will be from a particle in the shower of energy  $E$ . Then according to calculations performed by Richards and Nordheim,<sup>(8)</sup>  $P(E)$  is independent of incident particle energy when that energy is sufficiently high (in the case of lead, the incident energy should be higher than the maximum of  $E$  and 70 MeV). It follows that the total path length of charged particles in the shower is proportional to the incident energy. This conclusion is not sensitive to small variations in  $P(E)$  with incident energy, for the ionization loss per track length is a slowly varying function of velocity for relativistic particles. Furthermore, the amount and spectrum of the Cherenkov radiation per path length is also a slowly varying function of  $\beta =$  velocity.

$$\frac{dN}{dv} = \frac{2\pi e^2 Z^2}{\hbar c^2} \left( 1 - \frac{1}{\beta^2 n_v^2} \right)$$

where  $A =$  atomic number and  $n_v$  is the index of refraction as a function of frequency,  $\nu$ .<sup>(9)</sup> We therefore can expect the spectrum of Cherenkov radiation to be independent of incident energy, and the total amount to be proportional to that energy. Because ionization losses and Cherenkov radiation are slowly varying functions of  $\beta$ , we can expect statistical variations in shower development to have only a small effect on the total Cherenkov light emitted.

Since there are no other particles near the eta mass that decay with any appreciable probability into two photons, there is, near the eta mass, only one peak in the invariant mass spectrum of the

photon pairs observed. It is the size of this peak that we require, rather than its precise position or shape. In order to measure the size accurately, the Cherenkov counters must have good resolution. Some effects which act to decrease this resolution are<sup>(10)</sup>

- 1) photo-electron statistics,
- 2) variation in the efficiency of light collection with variation in the position, direction and depth of the showers, and
- 3) failure of the counter to totally absorb the shower.

The first of the above mentioned effects appeared to be the major limitation on energy resolution. With seven phototubes producing a total of around a hundred photoelectrons in a typical shower, the gains had to be properly set in order to take full advantage of the amount of light available.

Suppose  $\sigma$  is the standard deviation of the signal from a single phototube in the counter,  $\mu$  is the mean of that signal, and  $G = \text{gain} = \frac{\sigma^2}{\mu}$ . Assume that for showers of a given energy, fluctuations in one tube are independent of those in another in the same counter, and that  $\frac{\sigma}{\mu}$  for a given tube is independent of its voltage. Then the best resolution is obtained when the gains for the different phototubes in the same counter are all equal. In Appendix A we discuss the assumptions, prove that the gains should indeed be matched, and discuss the effect of errors in the matching.

The gains were matched with the help of light-emitting diodes taped to the face of each counter. We would give the diodes

several thousand identical pulses, thereby simulating the pulses of Cherenkov light from identical showers. The signals from each of the phototubes would be pulseheight analyzed, and  $\sigma$  and  $\mu$  would be calculated, thereby giving the gain. By varying the voltages of each tube, the gains would be set equal.

Because gains could be matched so easily to within better than 10% (we had the help of an on-line computer), we were able to do the matching many times during the experiment. This experiment used the same counters as were used in a prior experiment, so by the time the data collection began, the phototubes had been "on" continually for nearly a year, and were relatively stable. Although gains were frequently matched, the voltage changes were always small; so even if we had matched gains only once, the drift in the course of the experiment would not have caused the gains to differ from a central value by more than 20%. As is shown in Appendix A, 20% matching is close enough to avoid seriously degrading the resolution.

The other effects tending to decrease the resolution are believed to be smaller than the effect of statistics at the photocathode. We now discuss these other limitations on the resolution.

Statistical fluctuations in shower development could indeed be expected to hurt resolution, for the phototubes did not all have the same conversion efficiencies, and even if they did, the absorption of short wavelengths by the glass would cause the pulse-height to depend on the depth of penetration of the shower. That is, statistical fluctuations in shower depth would broaden the resolution.

Cherenkov counters nearly identical to the ones we used had been made earlier at Caltech. In fact, we used the old lead glass blocks. Those who constructed the earlier counters made several different tests of their counters<sup>(10)</sup> in a monoenergetic positron beam. From these tests they concluded that the light from the showers was nearly isotropic (indicating either that there was much shower spread or that there was a large amount of ultraviolet light scattering). Furthermore, they found that there were many reflections from the faces of the counters. It could be anticipated, then, that the counters we built would be relatively insensitive to variation in the location and development of showers.

Our own tests in the same positron beam showed that varying the origin of the shower over the portion of the face used during the experiment varied the pulse height of the summed signal by less than 1%. Hence fluctuations in the lateral development of showers should have had little effect on the resolution. This result holds in spite of the fact that each counter consisted of two glass blocks imperfectly optically joined (the summed signals over the three phototubes in a given block would be twice as large when the shower occurred in the same block as when it occurred in the other).

So far as statistical fluctuations in depth are concerned, tests in the positron beam showed that the pulse height of the summed signal was nearly proportional to the positron energy (see Fig. 3). Since more energetic showers penetrate more deeply, this proportionality indicates that fluctuation in penetration depth

could not be a serious problem.

The third of the above-mentioned sources of decreased resolution, incomplete shower absorption, was not a serious problem. We worried that some photons could come in at a position and angle such that substantial leakage through the sides could occur, but tests in the positron beam indicated that fewer than 5% of the incoming photons would lose over 10% of their energy in this manner.

We consider, now, the possibility of showers escaping through the back, rather than the sides of the counter. Approximate formulae given by Crawford and Messel<sup>(11)</sup> lead to the conclusion that 500 MeV photon showers in lead have only 3.6% energy penetration to depth greater than fourteen radiation lengths. The calculations leading to the formulae can be trusted, for similar calculations on electron-induced showers in lead agree with experiment.<sup>(12)</sup> Our counters were fourteen radiation lengths deep and usually dealt with showers of less energy than 500 MeV. Since the Compton effect attenuates low-energy photons more per radiation length for low Z than high Z materials, we could expect less leakage through lead glass than calculations give through lead. Thus we can conclude that shower leakage through the backs of the counters was negligible.

A final problem associated with the use of these counters is that of drift in the gain of the phototubes and the pulse-height analyzing system. As has been already indicated, these drifts were not large enough to seriously upset the gain matching. Nonetheless, the overall drift of the summed pulse-height analyzed signal could

have been large enough to make it impractical to combine the data from several runs at the same setting. In fact, it turned out that the entire system was stable to within 2% over periods of days. This stability was monitored, and instabilities were compensated for, in the following manner:

For each counter we placed a radioactive source ( $\text{Bi}^{207}$ ) on a scintillator and attached the scintillator to the side of the counter's front face with the bonding agent RTV. At least once, and generally twice, a day we would pulse-height analyze the light flashes from Auger electrons in the scintillator, and we would record the location of the peak in the spectrum. Figure 4 shows an example of such a spectrum. The first "peak" is really just the effect of the cutoff imposed by a discriminator in the fast logic. By using positron-induced showers of known energy, it was found at one time that the source peak in PbG-1 had about the same location as the peak from showers of 170 MeV, while the source peak for PbG-2 corresponded to 184 MeV.

According to what has been said earlier about Cherenkov radiation, showers from initial electrons or positrons of a given energy should produce just as much Cherenkov radiation as showers from photons of the same energy. We were therefore able to use the information in Fig. 3 to find the relationship between shower energy and pulse height. We could fit our results to a straight line

$$h = A(E - E_0),$$

where  $h$  is the pulse height in volts,  $A$  is a constant which is irrelevant for our purposes,  $E$  is the shower energy, and  $E_0$  is an energy pedestal characteristic of the counter. Although the error in the measurement of  $E_0$  was about 15 MeV for each of the counters, the method described in part III C for extracting the yields was insensitive to such an error. For a pulse of  $h$  volts, the pulse height analyzer would give a channel number  $x = a + bh$ , where  $a$  was carefully monitored (see next section), and  $b$  is a proportionality constant. Then  $E(x) = E_0 + C(x-a)$  for some calibration constant,  $C$ . If  $E_s$  is the shower energy for showers producing pulse heights equal to those from the source peak, and if  $x_s$  is the channel number corresponding to the source peak, we can conclude that

$$C = \frac{E_s - E_0}{x_s - a} .$$

$E_s$  as obtained in the positron beam was not very reproducible. In practice, for each individual run we usually used the location of the observed etas to find  $C$  with reproducibility to within about two percent. Then we could use the above equation as an independent, more reliable determination of  $E_s$ . In those runs for which the quality of the data was too poor to accurately establish the energy calibration, we could determine  $C$  with the above equation using  $E_s$  as determined by other runs.

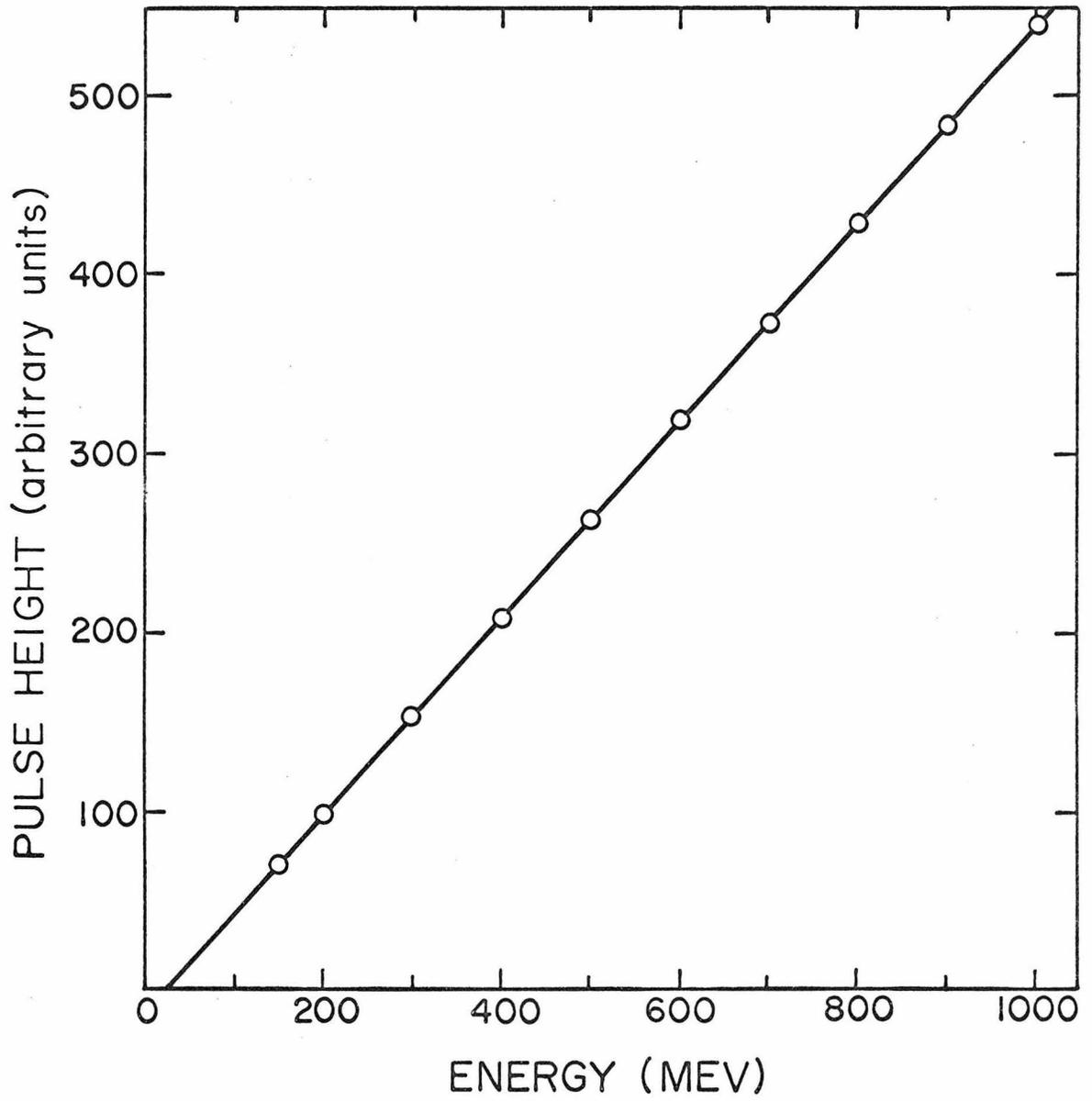


Figure 3. Cherenkov Counter Calibration

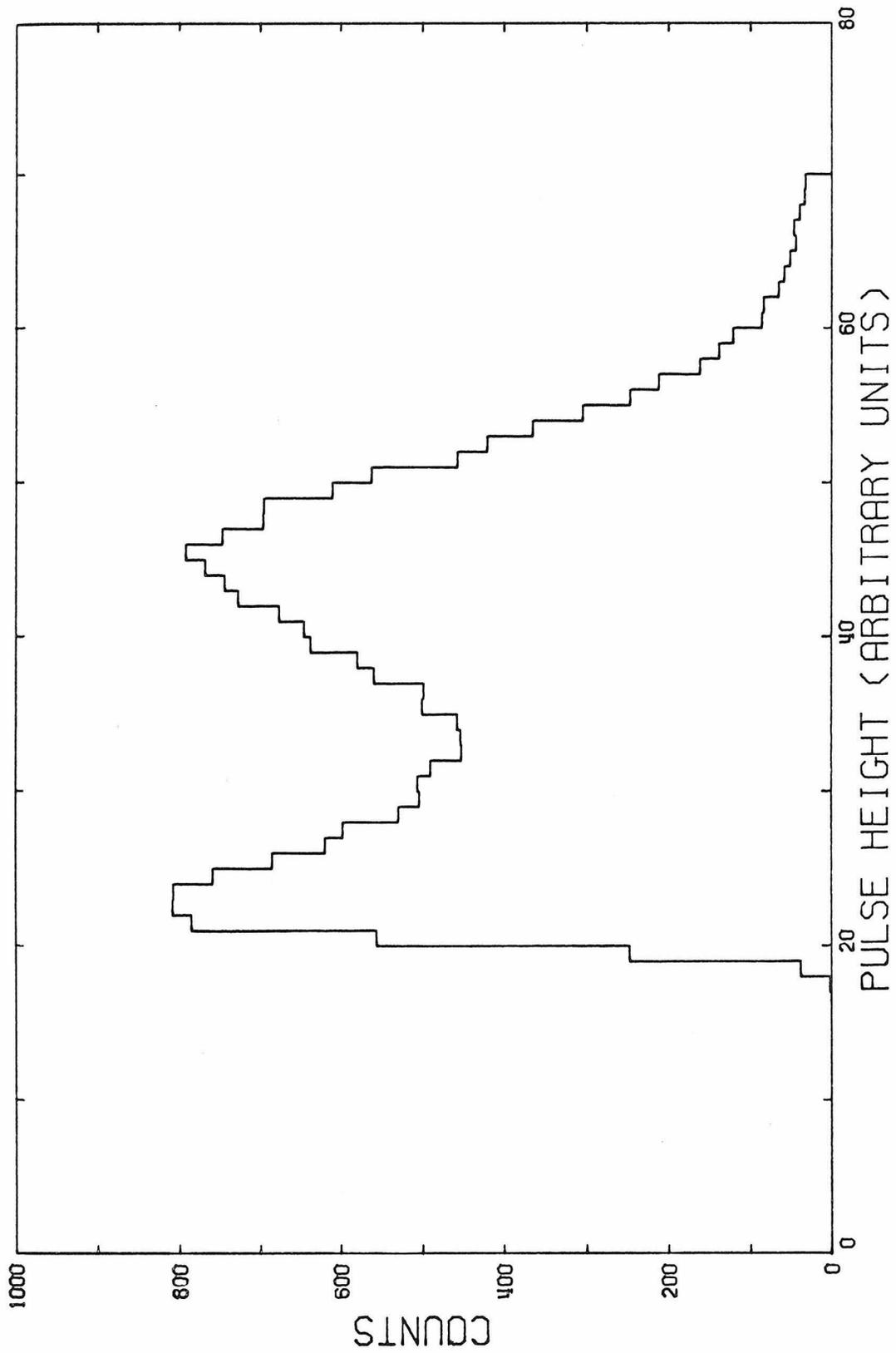


Figure 4. Spectrum of radioactive source.

## E. Electronics

Figure 5 shows the layout of the logic. Not shown are delay lines, certain "master" signals used to gate the pulse height analysis systems, inverting transformers used when the polarity of one module's output pulses would otherwise be incompatible with the input requirements of a succeeding module, and non-inverting transformers that were sometimes used to set D.C. levels to ground. Details of the pulse height analyzers and their interfaces with the computer are also omitted, as are certain other circuits considered irrelevant for understanding our method. Of the circuits denoted in Fig. 5, those in Table II.3 are described in Reference 14. The "name" is the name used in that reference.

Let us follow through the logic on the PbG-1 (lead glass counter number one) side of the system. The signals from the seven phototubes of PbG-1 entered a resistor network that served as a passive "adder." There was an additional input to the adder into which standard signals were sent for continuous calibration and monitoring during the run. Since the "adder" was really an averager, the output signal was rather small, and was amplified before being multiplexed. Three outputs came from the multiplexer. One was used in the fast logic, and the other two were used for pulse height analysis and timing purposes whenever the fast logic detected a possible eta. The "fast logic" consisted of the electronics to detect possible etas, while the "slow logic" consisted of the electronics to pulse height analyze the two PbG signals and determine

their relative time separation. The fast logic required of the order of a hundred nanoseconds to operate, while the slow logic required about a millisecond.

The multiplexer output for the fast logic went into a discriminator whose bias was set to exclude photons of energy less than about one hundred MeV. The eta meson has a mass of about 548.6 MeV. Because our counters were symmetrically placed, the decay photons tended to share the total eta energy about equally, and few etas were rejected by the requirement that each photon have energy greater than 100 MeV. As will be seen in section III C our method for determining the number of etas with both decay photons entering our counters automatically compensates for those etas with a photon rejected by a discriminator. The "x"'s in Fig. 5 represent discriminator outputs that were scaled. One of the discriminator outputs from the PbG-1 input was sent into coincidence with  $V_1$  to form the veto pulse,  $V_1 \times \text{PbG-1}$ , while another was sent into anticoincidence with that veto pulse to form  $\text{PbG-1} \times (\overline{V_1 \times \text{PbG-1}}) \equiv \gamma_1$ .  $V_1$  represents the signal from the veto counter before the PbG-1 Cherenkov counter.

The veto efficiency was measured both before and after the running of this experiment by detecting charged pions passing all the way through the Cherenkov counter and determining what percentage were vetoed by our system. At the highest counting rates used in this experiment, the efficiency was typically 95%. It was noticed, however, that during the running, one anticoincidence

circuit was somewhat less stable than expected, and the result could have been a lowered veto efficiency at times. This instability did not induce a systematic error, for it would merely increase the background which we subtract later. The method used in section III to subtract the background is not highly sensitive to the source of the background.

The  $\gamma_1$  pulse was sent to a discriminator, then if the  $\gamma_1$  pulse was of sufficient size, the discriminator gave a pulse that was placed in coincidence with a similar signal from the PbG-2 side of the logic. At this point, the beam gate also came in. The beam gate, a voltage level indicating that the beam is being dumped, was necessary to prevent the fast logic from being fooled by the noise and radiation produced during injection of electrons into the synchrotron. The resulting coincidence constituted the "master" trigger,  $\text{master} \equiv \gamma_1 \times \gamma_2 \times \text{beam gate}$ , and was used to gate the slow logic.

Whenever gated by a master pulse, the slow logic would digitize the PbG-1 and PbG-2 signals with a two dimensional Nuclear Data 160-F, 150-M analogue to digital converter (ADC) used in its 512 channel mode. The output of the Nuclear Data would be made available to the computer by a buffer in an interface, and the interface would signal the computer that a master pulse had occurred.

The time difference between the two signals was measured by

- 1) differentiating the signals with a capacitor,
- 2) sending the differentiated signals into zero crossing

discriminators<sup>(15)</sup> (ZCD's in Fig. 5 ),

- 3) whenever a trigger occurred, using one of the ZCD's to trigger a discriminator and the other to reset that discriminator, and finally
- 4) pulse height analyzing the integrated output of the triggered and reset discriminator.

The pulse height analysis of the discriminator output was not done with a Nuclear Data, but was done by sending the signal into a "super pulse-height analyzer"<sup>(13)</sup> (SPHA) whose output was a constant signal of length proportional to the charge of the input signal. The SPHA output was digitized in the usual manner (a clock and a scaler) by a Lecroy Digitizer. As in the case of the Nuclear Data interface, the digitizer interface held the digitizer output in a buffer and sent a pulse to the computer to indicate that data were about to become available.

The computer was a Programmed Data Processor-5 (PDP-5) manufactured by the Digital Equipment Corporation. As peripheral equipment it had an oscilloscope, a Teletype Model 33 ASR and a Hewlett-Packard digital magnetic tape unit. The PDP-5 had 4096 twelve-bit words, a cycle time of six microseconds, and had a single-level interrupt. External devices (including the interfaces mentioned above) could inform the computer of their status either at the computer's request (by way of a skip bus) or by driving the interrupt bus to ground.

Next we describe the PDP-5 machine language program used

with this experiment. This incomplete description is intended primarily as a framework in which to describe the data manipulation, monitoring, and calibration done with the help of the computer.

During a run, the computer program had two modes of operation -- a data input mode and a monitor mode. Normally the computer was in the monitor mode, but an interrupt caused by a pulse from an interface would send the program into the data input mode. In this mode it would wait for data to be available from the pulse-height analyzers, strobe those data in and store them in a buffer, then return to the place in the monitor program at which it was interrupted. When the buffer was full, the data input mode would continue storing the data in an overflow region, and would ring a bell on the Teletype either if the overflow region was unreasonably full or if the tape unit to which the data eventually would go was at the end of the tape. In the rare event (less than .01% of the time) that the data input mode was entered and no data were available within about a second, this fact would be noted in the memory, and the computer would return to the monitor mode without collecting the data.

The monitor mode had as its most essential function the storage of data on tape, but most of its time was spent either displaying or making computations required for display. By repeatedly examining switches set by the experimenter, the program would decide what to display on the oscilloscope, and would also determine the scale of the display. There were four possible displays -- a display of the buffer containing the PbG-1 pulse heights,

one of the buffer containing the PbG-2 pulse heights, one of the buffer containing the time differences, and a display of all binned information collected since the beginning of the run. We binned each of the input numbers, and also binned a fourth number calculated by the computer and intended to be nearly proportional to the invariant mass of the two-photon system detected. The computations and binning were interleaved with the display.

In order for the program to store its buffer on magnetic tape, the buffer had to be full and the beam gate had to be off (the computer had provision for testing the status of the beam gate). The overflow from the buffer was not immediately stored on magnetic tape. Instead, after the buffer had been stored on tape all numbers in the buffer were set to zero. Then any overflow was placed at the beginning of the buffer.

Some other functions of the monitor mode were counting the number of beam gates in a run, counting the number of bips in a run, and sending out calibration pulses.

The bips were counted with the help of a univibrator triggered by the beam integrator pulse. The computer tested the status of the univibrator frequently enough to count the number of bips to within 1% -- a negligible error resulting from electronic noise in the lab falsely triggering the univibrator. Each time a new bip was detected this information would be stored in the same buffer as the rest of the data. The bip count was also taken on a scaler not interfaced with the computer.

Calibration pulses could be used during the data collection, because during a period of several milliseconds before the beginning of each dump the beam gate was on. This time could be used for calibration without interfering with the data collection. Several times per bip the computer would, at the beginning of a beam gate, send out a pulse to fire a pulser that sent phototube-like signals into each of the two adders. These signals simulated eta events, and were treated as such by the electronics. They were distinguished from real events in the buffer by a single bit, which was set to one only when the incoming event was preceded by the output of a calibration pulse.

Figure 6 contains photographs of most of the possible oscilloscope displays as seen at the end of a typical run.

The bottom photograph shows an oscilloscope display of the part of the buffer containing the analyzed pulse heights from one of the Cherenkov counters. Each point represents one event. Events are displayed from left to right in order of their occurrence. The faint vertical dotted line on the left was produced by the computer's display subroutine, and could be shifted horizontally to any specified position on the display by appropriately setting the "switch register." The vertical position of each point in the photograph indicates the pulse height associated with the corresponding event, with zero pulse height events at the same height as the bottom of the dotted vertical line. The few events that apparently have zero pulse height will be explained shortly.

The top photograph shows the part of the buffer containing the time differences as received from the Lecroy Digitizer. As in the bottom photograph, the left-most events came in first, and the vertical position of each dot corresponds to the numerical value in the buffer. Above all the other points are ten equally high dots. These points represent the digitized time for calibration events. They are separated from the other events by virtue of the single bit which for calibration pulses is set to one. Below the calibration data, but above the rest of the timing data, are three equally high dots. These are the "bip markers", and each is stored in the buffer as if it were an event with three special values for timing and pulse height data. These same three "bip markers" can be seen as three dots at zero pulse height in the bottom photograph.

The buffer as displayed in the top and bottom photographs is full. The short line at lowest allowable height near the right of the two photographs represents part of the (empty) overflow region of the buffer.

The center photograph displays all the binned information. The left-most third of the photograph contains the invariant mass spectrum of all events in the run. The low-mass cutoff of the spectrum (i.e., the sharp rise of the left edge of the spectrum) is a consequence of discriminator cutoffs. Although most of the spectrum consists of background, on the tail of the mass spectrum is a small bump containing the etas. Moving to the right in the same photograph, the peak at the center of the photograph is part

of the time spectrum, with a full width at half maximum of about four nanoseconds. The two peaks in the right-most part of the photograph are the pulse height spectra from the two counters. They are each sharply cut off at low pulse heights by the discriminators. It takes a lot of imagination to see any evidence of etas in either of the single counter spectra.

Before and after each run, the computer helped with the collection and storage of other data. At the beginning of each run, the computer would request that such information as the synchrotron end-point energy, the counter angles, and the date and time be typed in. The experimenter and the computer would together find the channel number of the peak of the distribution from the radioactive sources in each Cherenkov counter (see the previous section for the use of these numbers). Also at the beginning of a run, the experimenter would send pulses into both adders from the same pulser used for calibration during the run. This would then be done again at twice the pulse height. With the help of the computer, the experimenter would find the pulse-height analyzer channels corresponding to each side of the logic and each voltage of the pulser. Recall from the previous section that if "h" represents the pulse height, the channel "x" followed the relationship  $x = a + bh$ . "a" is called the "pedestal," and was found easily with the use of the data from the pulser. Incidentally, it was found that the major instability in the above described relationship lay in the pedestal. The pulser was tested to be sufficiently stable that the shift of the pedestal

during the run could be determined by the calibration pulses taken during the run. When all the beginning-of-run data were available the computer would store the data on tape and the run would begin.

At the end of the run, the computer would type out the number of gates, bips, triggers, triggers without data, and the number of records put on tape. It would also type out requests for the contents of the scalers used to monitor the fast electronics. Finally, comments could be typed in, and the end-of-run data would all be stored on tape.

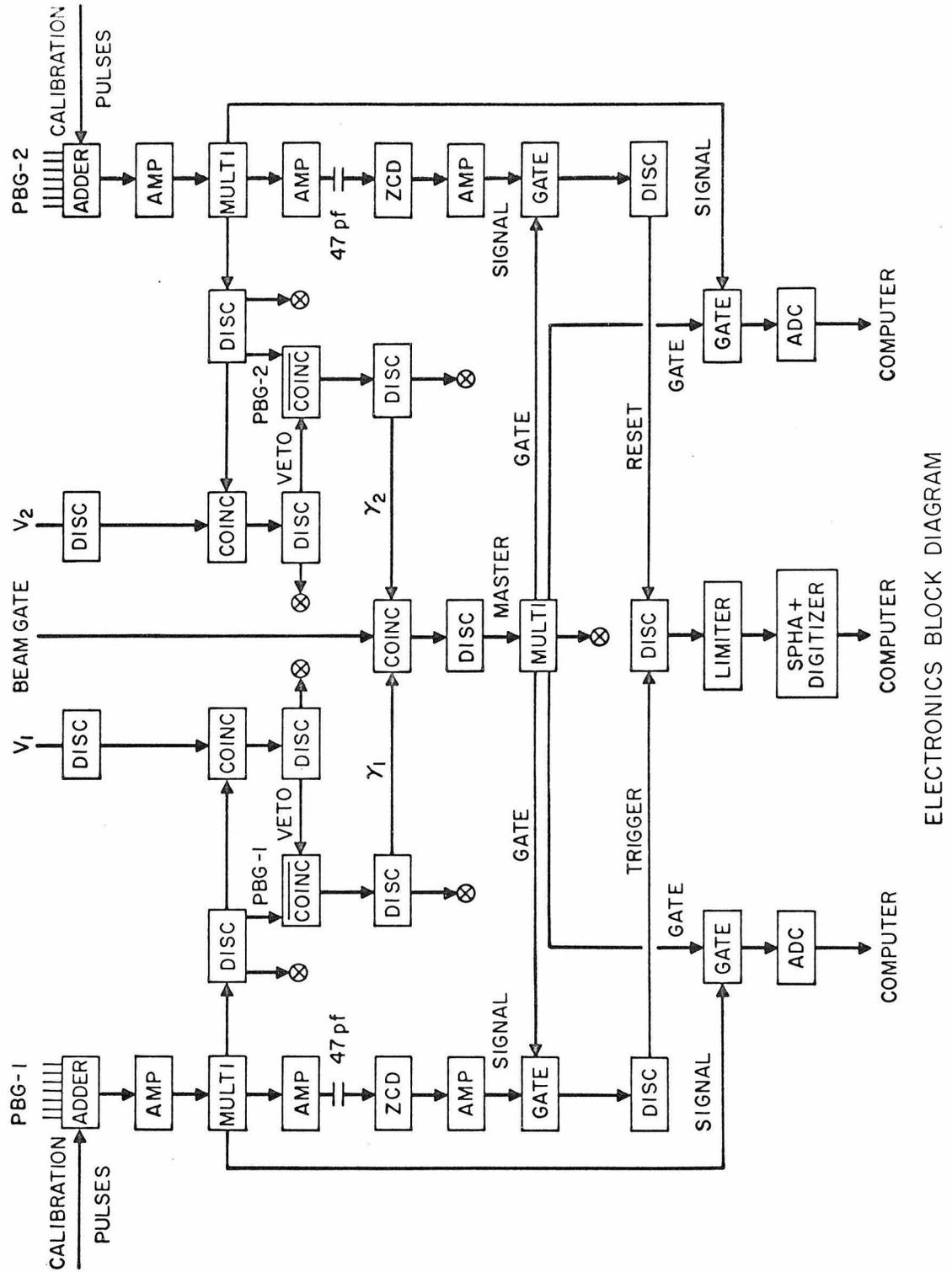


Figure 5. Electronics Block Diagram.

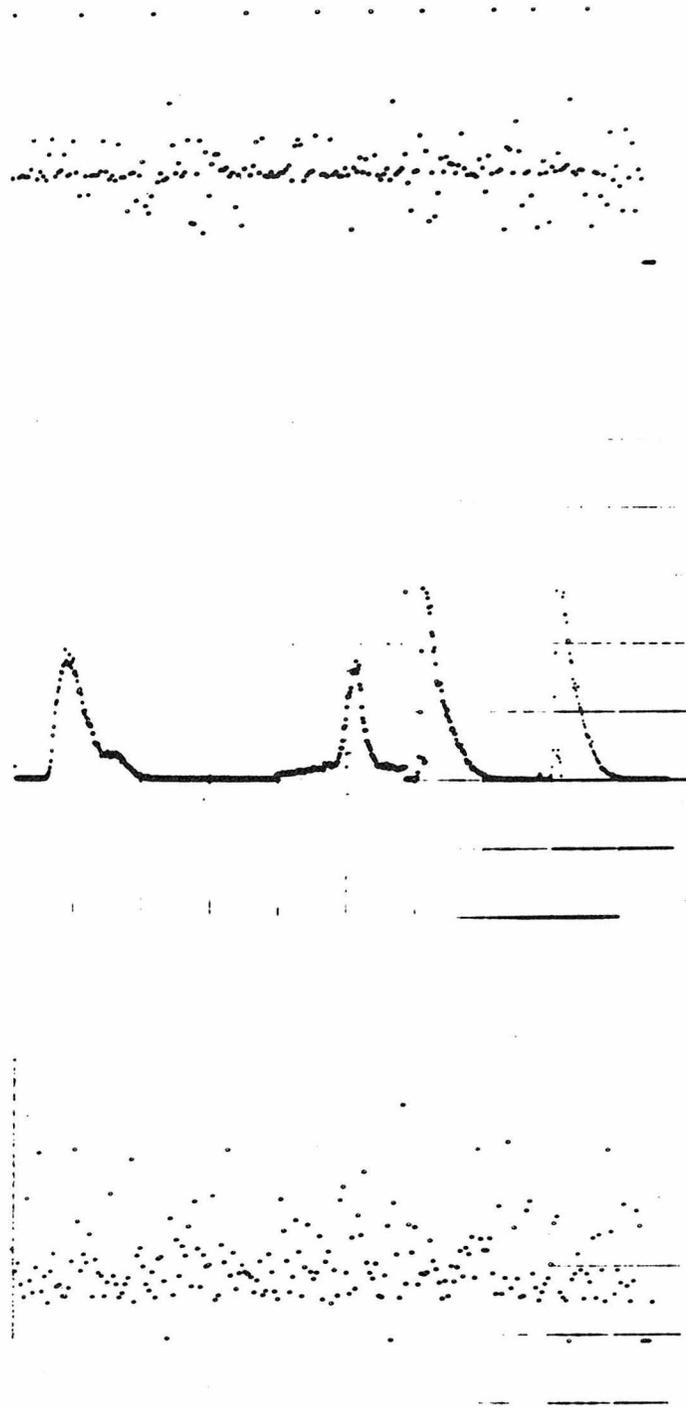


Figure 6. Oscilloscope Displays

TABLE II.3

## Circuits Used in This Experiment

CIRCUIT	NAME
Amp	FA-1
Multi	TM-4
	L-3*
Coinc	TC-6
$\overline{\text{Coinc}}$	TC-6
Disc	TVD-3
	TVD-4**
Gate	TG-3
Limiter	L-3

\* Limiters were used as multiplexers for the master signal.

\*\* TVD-4's are DC coupled, so were used instead of the AC coupled TVD-3's in parts of the logic with high counting rates.

## III. DATA ANALYSIS

## A. Raw Data

Data were taken at settings chosen according to the criteria given in Section II A. In Table III.1 we have a list of setting along with various information about the data collected at those settings. In the column labeled "target" is the length in inches of the appendix of the deuterium target used for the setting (see Section II C). The "energy" is the nominal end-point energy of the synchrotron before correcting for the error in calibration of the beam energy meter. This correction is discussed in Appendix F. The "angle" is the approximate angle between the center of each of the two symmetrically placed Cherenkov counters and the beam line. For purposes of data analysis, the geometry of the detection system was defined more precisely by surveying. The number of "bips" is a measure of the total photon beam energy passing through the target (see Section II B). "Events" represents the total number of triggers to which the computer responded at a given setting, excluding triggers from calibration pulses (see Section II E). The column labeled "etas" contains the estimated number of etas whose two decay photons entered the two Cherenkov counters. To obtain the number of etas, we had to separate the eta events from the large background contamination consisting of (1) accidentals in which two unrelated photons simultaneously entered the two counters, (2) true coincidences from processes other than  $\gamma + d \rightarrow \eta + n + p \rightarrow 2\gamma + n + p$ , and (3) both coincidences and accidentals in which at least one of the particles considered to be a photon was in reality a charged particle

that our veto counters failed to detect. Later (Section III C) we discuss the background subtraction along with the method used for estimating the number of etas. Essentially the method involved finding for the data of each run a good fit involving several parameters, one of which could be interpreted as the number of etas.

Table III.1 includes only data from runs which were ultimately used to obtain a cross section. Runs were rejected when the good fit needed to extract the number of etas could not be found (i.e., when the fit to the foreground or to the background made according to the method of Section III C was more than 2.5 standard deviations away from perfect). Runs were also rejected when some anomaly caught during or after the run made the data untrustworthy or when the amount of data collected was too small for us to trust the results of our method for extracting the number of etas (we included only runs with more than eighty bips).

TABLE III

Settings for Which Data Were Collected

Target	Angle	Energy	BIPS	Events	Etas
3.27	35	1150	2315	224572	470
3.27	37	1025	2314	139894	425
3.27	37	1100	1215	84723	349
6.39	40	925	201	28934	856
3.27	40	950	2529	101832	1266
6.639	40	875	266	30055	720
6.639	45	825	966	110231	4867
6.639	45	850	1031	121760	6673
6.639	50	775	1037	86376	4273
6.639	50	800	364	36461	2405
6.39	50	800	379	36527	2270
6.639	55	725	1025	66068	2395
6.639	55	750	1029	68467	5308
6.639	60	725	339	24458	1193
6.639	60	750	190	14077	1299
6.639	65	725	1020	50701	4271
6.639	65	750	516	48539	3181
6.639	70	725	116	5426	425
6.39	70	750	236	14630	1106
3.27	70	775	451	5448	563
6.639	75	750	311	15656	1117
6.39	75	750	236	12506	836

TABLE III, cont.

Target	Angle	Energy	BIPS	Events	Etas
6.639	75	775	512	23152	2362
6.39	75	775	237	10740	1084
6.639	80	800	871	32028	3208
6.639	80	850	936	30532	4550
6.39	85	825	482	21920	1177
6.639	85	875	1464	57771	4432
6.639	90	800	639	16070	515
6.639	90	925	1041	28908	1960
6.639	90	975	1050	24017	1766
6.639	93	1025	1058	21717	1474
6.639	93	1100	1271	28666	2001
6.639	95	1100	1068	22344	1185
3.27	95	1100	2629	13102	667
6.639	95	1175	541	12603	653
6.639	95	1195	541	12008	572
6.639	97	1175	363	7219	282
3.27	97	1175	2283	15368	409
6.639	97	1225	934	18540	787

## B. Accidental Background

In order to determine the number of etas detected at a given setting we must somehow estimate and remove the background. In this section, we discuss the estimate of the size and distribution of the accidental background.

Figure 7 displays the counting rate as a function of the time separation between the pulses from the two Cherenkov counters. There are two curves in this figure, both corresponding to the same run. The true coincidences are concentrated in the peaks, while the accidental coincidences are spread out over a wider range of time. The taller peak (drawn with a slightly thicker line) represents the result of correcting for an effect to be described shortly. Twenty-five units of "digitized time" correspond to about one nanosecond. Our resolution, then, is about three nanoseconds full width at half maximum. One simple way of excluding most of the accidentals is to make a timing cut (represented by vertical dotted lines in Figure and delete all events that are too far out of time.

To approximately account for the accidentals under the timing peak, we can estimate their number by interpolating the time distribution of the accidentals. We can obtain the distribution in pulse heights of those accidentals by assuming that the energy distribution of accidentals under the timing peak is the same as that of accidentals outside the peak. We return to these points later.

No matter how we deal with the accidentals, any improvement

in time resolution helps. We now discuss a small improvement in the time resolution that was made after the data had been collected.

Although we used zero-crossing discriminators in the experiment, there was still a small amount of slewing. That is, the time it took for a signal from one of the two Cherenkov counters to be detected by a discriminator depended to a small extent on the size of the incoming pulse. We estimated the amount of this slewing using the data collected during the normal course of the experiment. We binned the data with respect to  $E_1$ ,  $E_2$ , and  $T$  (the pulse-heights from the two counters and the relative time measured between the two pulses). For a fixed  $E_1$  and  $E_2$ , the time bin corresponding to coincidences can be estimated by, for example, seeing which time bin has the most counts.

Let  $T_{ij}$  be the time difference measured between coincidences of energy  $E_1 = E_i$  and  $E_2 = E_j$ . Then we expect to find the form

$$T_{ij} = (T_1)_i + (T_2)_j$$

if the time of detection of each pulse is a function of the incoming pulse-height. Knowledge of  $T_1$  and  $T_2$  allows us to correct for the slewing by finding  $i$  and  $j$  corresponding to the pulse-heights of the events and by then computing

$$T_{\text{corrected}} = T_{\text{observed}} - (T_1)_i - (T_2)_j .$$

Any determination of  $T_{ij}$  is subject to statistical errors. We use  $T_1$  and  $T_2$  because they can be determined with somewhat smaller

statistical errors. Consider

$$S = \sum_{ij} (T_{ij \text{ observed}} - (T_1)_i - (T_2)_j)^2$$

Then minimizing  $S$  with respect to the unknowns  $(T_1)_i$  and  $(T_2)_j$  is expected to give an estimate of those parameters. We get the equations

$$(T_1)_i = \frac{1}{N} \sum_j (T_{ij} - (T_2)_j)$$

III B.1

$$(T_2)_j = \frac{1}{N} \sum_i (T_{ij} - (T_1)_i),$$

where "N" is the number of values  $i$  (or  $j$ ) may take, and  $T_{ij}$  means  $T_{ij}$  observed.

Equations III B.1 do not have a unique solution. If  $(T_1)_i$  and  $(T_2)_j$  is a solution, then so is  $(T_1)_i + C$  and  $(T_2)_j - C$  for any value,  $C$ , that is constant over all  $i$  and  $j$ . Without loss of generality, we can choose  $C$  so that for some bin,  $k$ ,  $(T_1)_k = (T_2)_k$ . We obtain  $T_1$  and  $T_2$  by iteration:

- a) For all  $i$  and  $j$ , pick an initial guess for  $(T_1)_i$  and  $(T_2)_j$ .
- b) Evaluate a new set of  $T_1$  and  $T_2$  using equations III B.1.
- c) Adjust  $T_1$  and  $T_2$  with  $C$  chosen as described above.
- d) Return to step b.

This iterative procedure was found in practice to converge.

Some improvements were made in the method as described above. To better estimate  $T_{ij}$ , we interpolated the accidental background under the coincidence peak for each  $i, j$  pair, and after subtracting

the accidentals found the median of the coincidence peak. We also counted the number of events under the coincidence peak,  $N_{ij}$ , and (because we expected to be able to estimate  $T_{ij}$  better when  $N_{ij}$  was large) we weighted  $S$  according to

$$S = N_{ij} (T_{ij} - (T_1)_i - (T_2)_j)^2 .$$

The weighting factor,  $N_{ij}$ , changed the equations to be solved, but not the method of solution. The two curves of Figure 8 show the average over all runs of the slewing correction for each of the two zero-crossing discriminators. As can be seen, one of the two discriminators accounted for most of the slewing. Figure 7 shows the timing distribution for a particular run before and after correcting for the slewing, with the slightly narrower peak showing the slight improvement that was typically achieved. The "digitized time" is the number sent into the computer by the Lecroy Digitizer (see Section II E). At this point it should be noted that the numerical output of the digitizer was not linear in the time difference between the two signals. With the help of delay cables of known length, we were able to calibrate the digitizer output in terms of nanoseconds (see Figure 9). The calibration curve was not needed in the data analysis, but its non-linearity partly accounts for the non-uniformity of the accidental timing distribution of Figure 7.

Now let's consider the distribution in pulse-heights of the accidentals beneath the timing peak. Suppose we bin events for an entire run with respect to the two pulse-heights. Define  $A_{ij}$  to be the number of accidentals in one such bin for events outside the

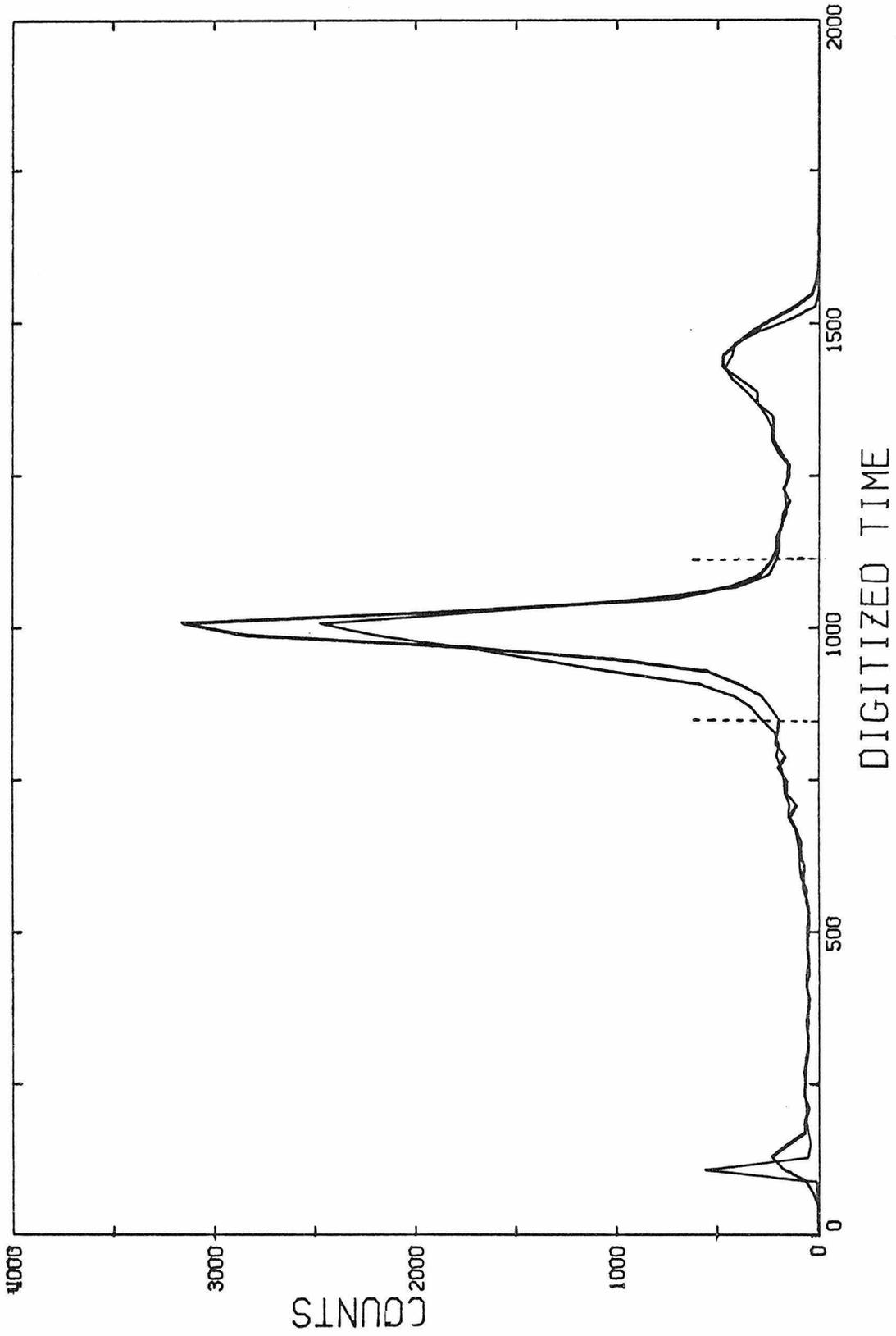


Figure 7. Timing Distribution

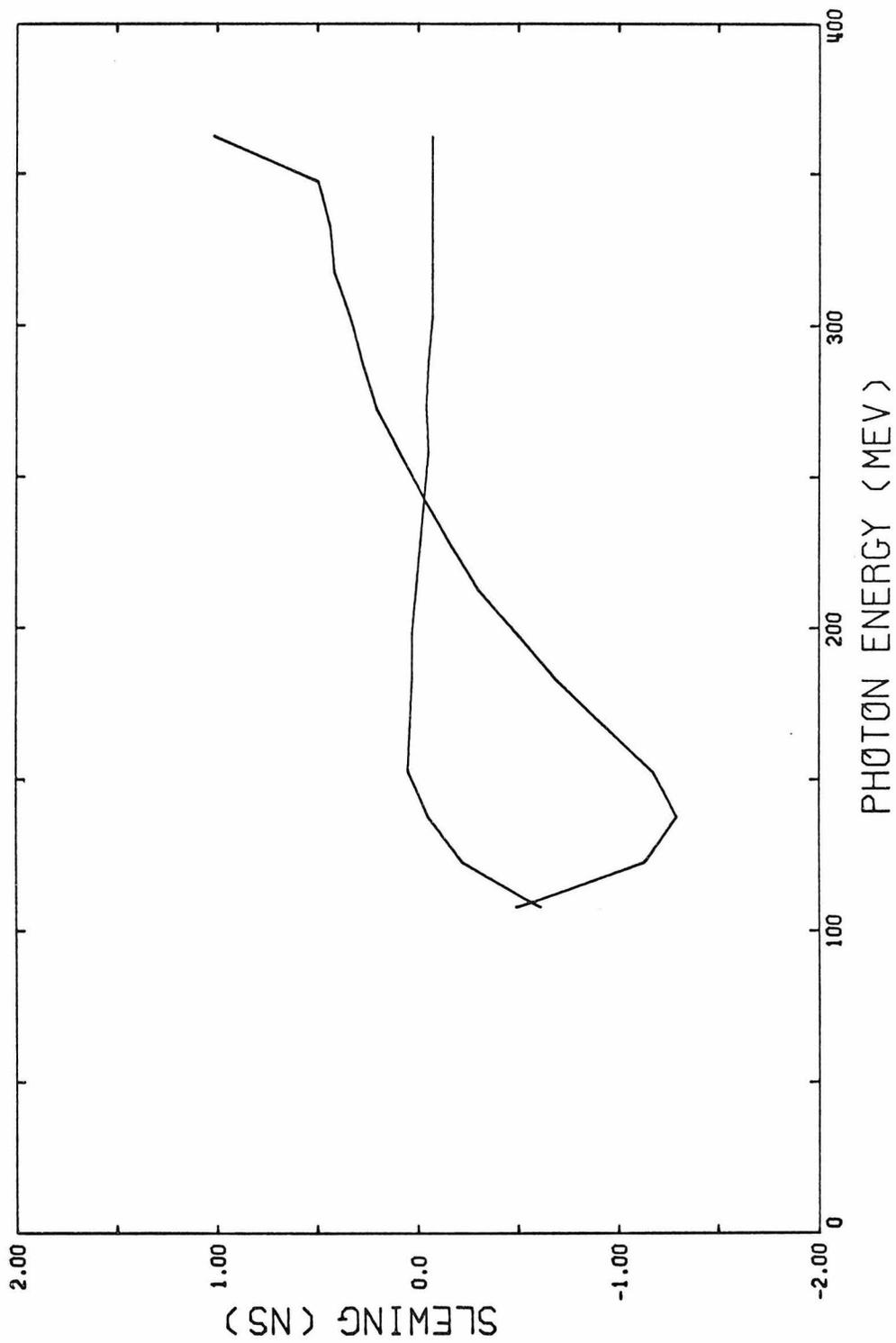


Figure 8. Average Slewing Correction.

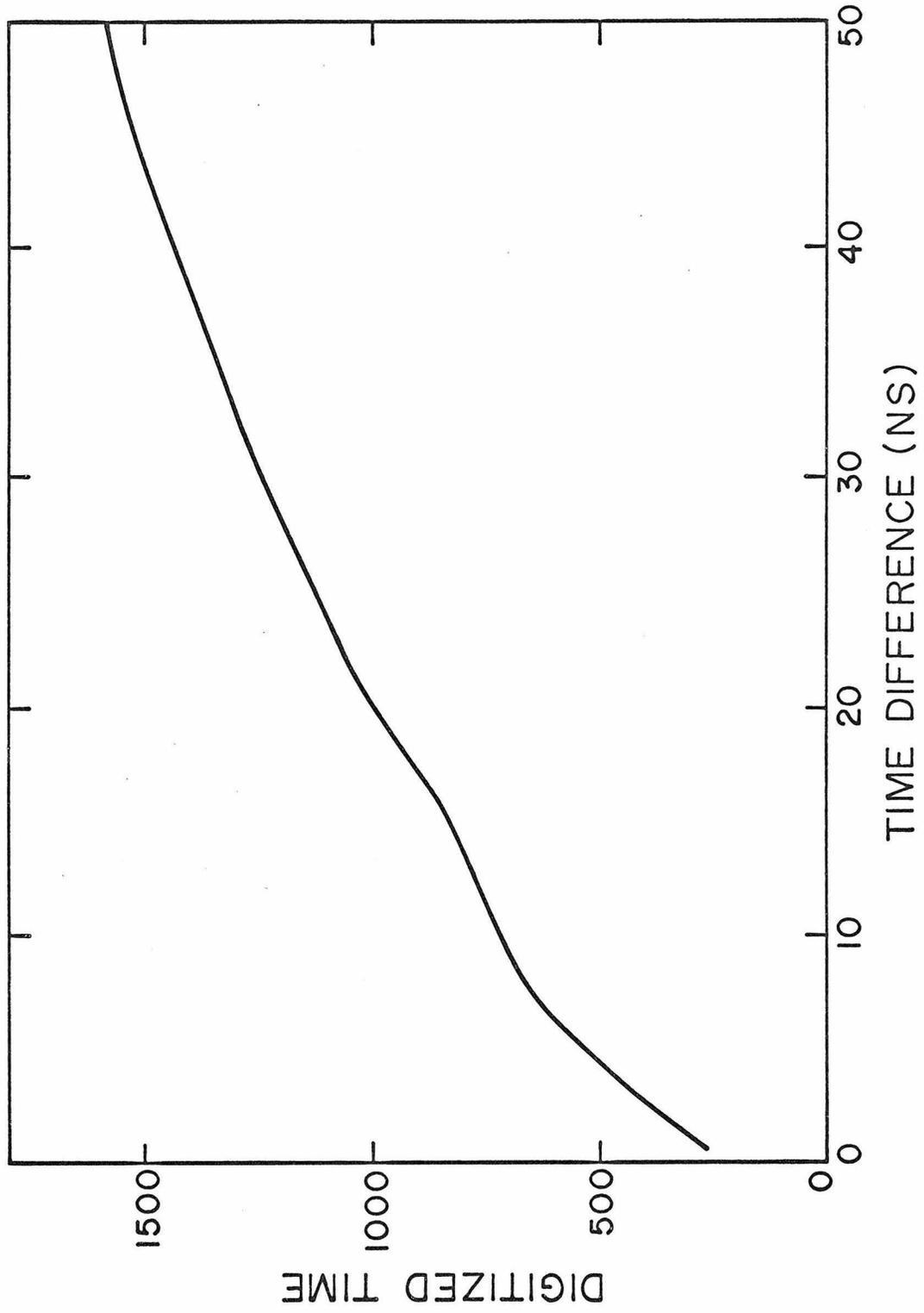


Figure 9. Calibration of Time Digitizer.

timing cut.  $A_{ij}'$  is directly measurable from the experiment. Define the non-measurable quantity,  $A_{ij}$ , to be the number of accidentals in the run beneath the timing coincidence peak and with pulse-heights in bin  $i, j$ . Even though we cannot measure  $A_{ij}$  directly, we can expect to have

$$\bar{A}_{ij} = f \bar{A}_{ij}' , \quad \text{III B.2}$$

where we use the notation "expectation value of  $x$ " =  $\langle x \rangle = \bar{x}$  and take  $f$  to be a constant for the run independent of  $i$  and  $j$ . One simple way of estimating  $f$  is to linearly interpolate the accidental rate outside the timing cut into the time zone of the coincidences. A more elaborate method that failed is described in the following paragraph.

As was discussed in Section II E, each time the meter monitoring the bremsstrahlung beam reset, a "bip" marker was placed in the incoming data stream. Between any two adjacent "bip" markers, consider the two quantities:

$N_{in}$  = the number of events during the "bip" that are within the timing cut, and

$N_{out}$  = the number of events during the "bip" that are outside the timing cut. The standard deviation of  $N_{out}$  tells us something about how the beam intensity varied during the run. For such a variation of beam intensity, the correlation between  $N_{in}$  and  $N_{out}$  can be used (in the absence of dead time) to compute the number of accidentals within the timing cut. This method would allow us to measure  $f$  if the statistics were good enough. Unfortunately, they

weren't, so no more will be said about this method. In a later section we will discuss the effect of errors in  $f$ . But now that we have equation III B.2, we are motivated to more closely examine  $\bar{A}_{ij}'$ .

The two photons of an accidental event cannot be correlated in energy. In other words, knowledge of the pulse-height from one of two photons from an accidental event gives no clue as to what the pulse-height from the other photon might be. Mathematically, this means that  $\bar{A}_{ij}'$  is of the form

$$\bar{A}_{ij}' = a_1(i) a_2(j) .$$

For a given  $\bar{A}_{ij}'$ , let us assume that during any infinitesimal time interval the probability of an event occurring in bin  $(i, j)$  is not dependent on when, whether, or in what bins other events occurred; so, for example, we are neglecting the effect of dead time. It follows from the above assumption that the probability distribution of the observed value of  $A_{ij}'$  will follow a Poisson distribution. Thus for a given set of  $a_1$  and  $a_2$ , the probability of the observed set of  $A_{ij}'$  is

$$L = \prod_{i,j} \frac{(\bar{A}_{ij}')^{A_{ij}'}}{(A_{ij}')!} e^{-(\bar{A}_{ij}')}$$

To find the best values of  $a_1(i)$  and  $a_2(j)$  for all  $i$  and  $j$ , we use the maximum likelihood method (Appendix E). That is, we choose  $a_1$  and  $a_2$  such that  $L$  is maximized.  $L$  is maximized when we maximize

$$W = \log[L \prod_{ij} (A_{ij}')!] = \sum_{ij} (A_{ij}' \log(\bar{A}_{ij}') - \bar{A}_{ij}')$$

with respect to  $a_1(i)$  and  $a_2(j)$ . Setting derivatives of  $W$  equal to zero gives the two equations

$$a_1(i) = \frac{\sum_j A_{ij}}{\sum_j a_2(j)} \quad \text{III B.3}$$

$$a_2(j) = \frac{\sum_i A_{ij}}{\sum_i a_1(i)} .$$

All solutions of equations III B.3 with the same product,  $a_1(i) a_2(j)$ , are equivalent. Without loss of generality, we can specify, for example, that

$$\sum_j a_2(j) = 1 .$$

Then equations III B.3 are the exact maximum likelihood solutions for  $a_1$  and  $a_2$ . With the help of equation III B.2, we will later use these solutions for  $\bar{A}_{ij}$  in the estimation of the total background contamination of the eta events. This estimation of the total contamination is the major problem to be dealt with in the next section.

### C. Background Subtraction

Figure 10 shows the distribution of events in the  $E_1$ - $E_2$  plane (where  $E_1$  and  $E_2$  are the observed pulse-heights in the two Cherenkov counters) after timing cuts have been made. For showers produced by photons of a given energy, the pulse-height is nearly proportional, on the average, to that energy (see section II.D). In Figure 10

the pulse-height is expressed in terms of the corresponding energy. The method of calibrating the energy vs pulse-height will be described shortly. The peak at high energies in Figure 10 comes from the etas. Etas are restricted to this region of the plane by virtue of the relationship

$$m = 2 \sin \theta \sqrt{E_1 E_2} ,$$

where  $\theta$  is half of the angle between the two photons,  $m$  is the invariant mass of the two photon system, and the  $E_i$  are the pulse-heights expressed in terms of energy. As the pulse-heights get lower in Figure 10, the density of etas drops to zero, but the background density rises until all events are cut off by biases in the electronic logic. Notice that even though we display in Figure 10 a setting with an unusually prominent eta peak, there is still a considerable background under the peak. The main problem involved in estimating the yields is the background subtraction.

The first step in making the background subtraction is to bin the events according to the pulse-heights. In our notation,  $i$  and  $j$  refer to the bins in which  $E_1$  and  $E_2$ , respectively, fall. For a particular run,  $N_{ij}$  is the number of events in the  $i, j$ 'th bin. Assume that we have already eliminated as many accidentals as we can by making timing cuts. Then  $A_{ij}$  is the number of accidentals remaining in the background, and its average is estimated by the method described in the preceding section. Let  $B_{ij}$  denote the non-accidental background and  $Y_{ij}$  denote the eta yield in bin  $i, j$ . In our notation,

$$N_{ij} = A_{ij} + B_{ij} + Y_{ij} \quad . \quad \text{III C.1}$$

Our intention is to determine the total yield of etas at a given setting essentially by estimating  $A_{ij}$  and  $B_{ij}$  under the eta peak and applying equation III C.1.

In those regions of the plane for which  $Y_{ij}$  is small, we can estimate  $B_{ij}$  by neglecting  $Y_{ij}$  in equation III C.1 and using the results of the preceding section for  $\bar{A}_{ij}$ . In order to determine  $B_{ij}$  beneath the eta peak, we use the approximation that as in the case of accidentals, the energies of the two photons from background events are uncorrelated. With this approximation, the average of  $B_{ij}$  is of the form

$$\bar{B}_{ij} = b_1(i) b_2(j) \quad \text{III C.2}$$

Insofar as the experiment is completely symmetric about the beam line of the incoming photons,  $b_1 = b_2$ . We do not, however, assume this symmetry.

In the region of the  $E_1 - E_2$  plane corresponding to  $\pi^0$  production, equation III C.2 is rendered invalid by the concentration of events with pulse-height pairs corresponding to the pion mass. If we are to have any hope of using equation III C.2, we must avoid that part of the background with low invariant mass. We exclude from consideration all events with measured invariant mass below about 240 MeV.

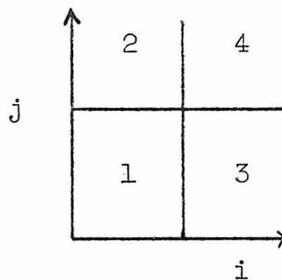
With the above proviso, we experimentally justify the use of

equation III C.2 in two ways:

1) When the target is hydrogen, rather than deuterium, eta photoproduction can be kinematically excluded for certain settings close to those for which etas are abundantly produced. Then  $Y_{ij}$  is everywhere zero, and we can test equation III C.2. We made such tests near settings for which etas moving forward in the center-of-mass are produced, as well as near settings for which etas moving backward in the center-of-mass are produced. For each of the settings, equation III C.2 was satisfied to within statistical errors.

2) For all runs, equation III C.2 can be checked in the region of the  $E_1 - E_2$  plane where the effect of  $Y_{ij}$  can be neglected. Such checks show almost no measurable deviation from zero correlation. Later we will discuss the (elementary) statistical methods used to check how well equation III C.2 is satisfied by the assumptions we make, including that of zero correlation. But first let us discuss how the assumption of no correlation allows us to determine  $B_{ij}$  under the eta peak.

Suppose we divide the  $i$ - $j$  plane into four regions as follows:



Suppose further that we know the background in regions 1, 2, and 3,

but not in region 4. If  $n_i$  is the total number of background counts in region  $i$ , then the assumption of no correlation gives within statistical errors

$$\frac{n_2}{n_1} = \frac{n_4}{n_3} ,$$

so that  $n_4$  can be computed in terms of measurable quantities.

The discussion of the previous paragraph was intended only to show that, in principle, the information on the background beneath the eta peak is contained in the measured background away from the peak. To obtain the value of  $B_{ij}$  for all bins requires a more elaborate treatment.

We divide the  $i$ - $j$  plane into two regions -- the "eta region," analogous to region 4 of the above discussion, and the "background region." In order to define the "eta region," we combine the results of a Monte Carlo program (which, as a byproduct, generates artificial eta events for each setting at various incoming photon energies) with some reasonably assumed eta photoproduction cross section and an assumed energy resolution of our counters (as independently measured in a positron beam -- see Appendix A). We obtain an expected distribution of etas in the  $E_1$ - $E_2$  plane. Then a contour about the eta peak can be drawn such that approximately 80% of the etas are expected to fall inside the contour.

Because this definition of the "eta region" is so arbitrary, we will have to show later that our method of estimating the yield is not very sensitive to the location of the boundary of the eta region.

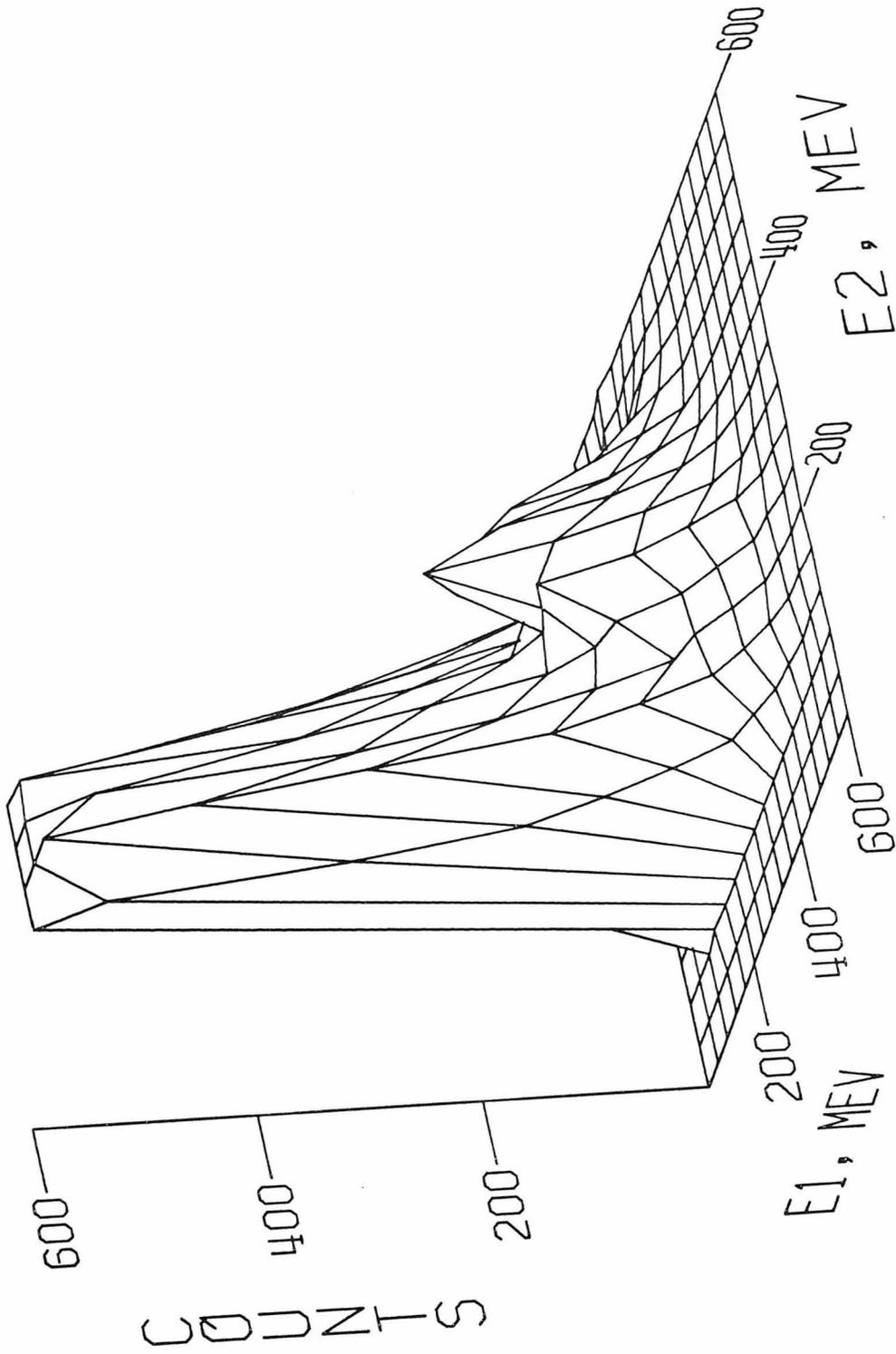


Figure 10. Distribution of Pulse-Heights for a Typical Run.

Figure 11 shows the above described division of the  $i$ - $j$  plane. The "low mass cut" shown in the figure has been already discussed. The "low  $E_1$  cutoff" and the "low  $E_2$  cutoff" are set slightly above the pulse heights at which the fast electronic logic begins to exclude events.

The same Monte Carlo results that were used to define the boundary of the eta region can be used to estimate the leakage of etas into the background region. Here the errors in some of the assumptions involved can hurt. We do not accurately know the energy calibration of the pulse heights, we don't know very well the resolution of the counters, and we don't know the cross section for eta photoproduction from deuterium. At this point, it should be pointed out that we are about to describe an iterative procedure for obtaining the yields. The initial assumptions may be inaccurate, but part of the output of each iteration will be improved estimates of the unknown energy calibration, resolution, and cross section.

Taking expectation values in equation III C.1 gives

$$\bar{N}_{ij} = f a_1(i) a_2(j) + b_1(i) b_2(j) + \bar{Y}_{ij} . \quad \text{III C.3}$$

We have used the results of the previous section, along with equation III C.2 of this section.

As in section III B, we use the maximum likelihood method to find equations for  $b_1$  and  $b_2$ . The sums over  $i, j$  are restricted to the background region and the equations are complicated by the presence of  $\bar{Y}_{ij}$  = the estimated eta leakage into the background region, and

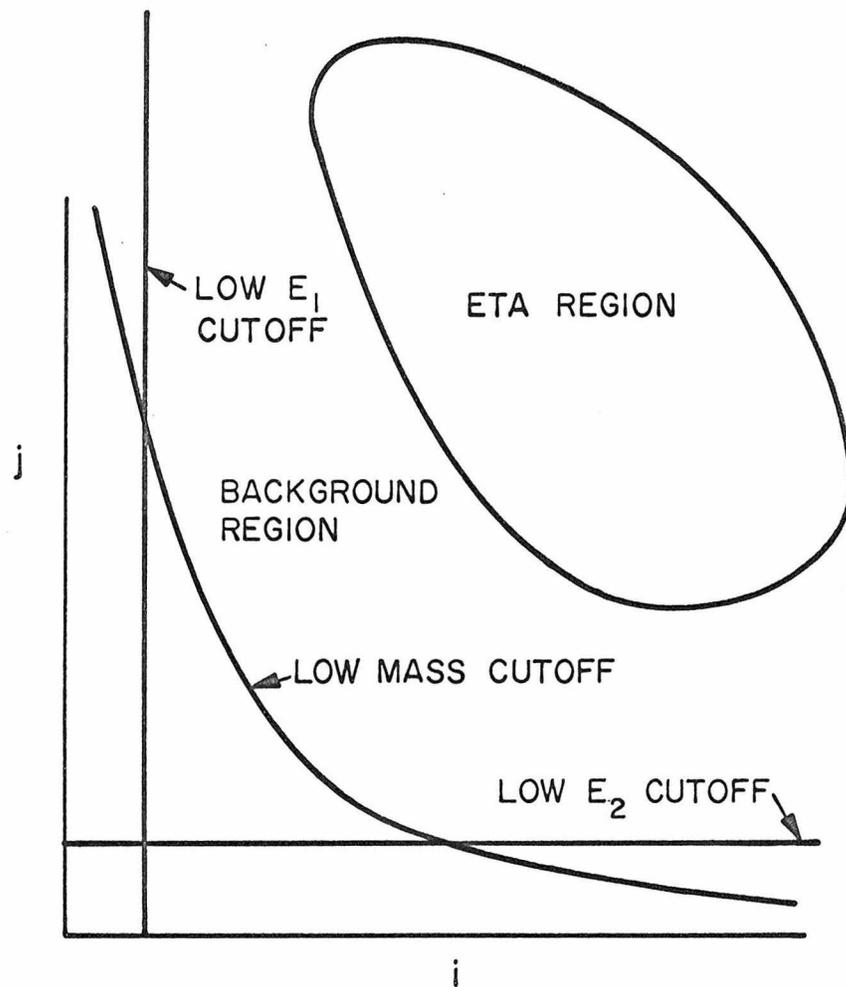


Figure 11. Division of the  $i$ - $j$  Plane.

by the existence of the accidentals. As a consequence of these complications, the equations resulting from the maximum likelihood condition are not trivially soluble. An iterative procedure for solving the equations is discussed in the thesis of W. A. McNeely.<sup>(27)</sup>

Although we solve for  $b_1$  and  $b_2$  using only information in the background region, the resulting solution applies to the eta region. The above-described method for obtaining  $B_{ij}$  really amounts to an extrapolation technique. The next step in the iterative procedure is to apply equation III C.3 to the eta region with  $Y_{ij}$  now considered to be a function of a few unknown parameters. These parameters are none other than the ones assumed as input to this iterative procedure. We have as unknowns:

- 1) the scale factors (the proportionality constants relating the pulse-heights to the expectation value of the shower energy),
- 2) the resolution, and
- 3) the size of the cross section.

The resolution is parameterized by  $\alpha$  in the equation

$$\left(\frac{\mu}{\sigma}\right)^2 = \alpha E$$

$E$  is the shower energy,  $\mu$  is the expectation of the pulse-height from showers of energy  $E$ , and  $\sigma$  is the standard deviation of the pulse-heights from such showers. This equation is an approximation to the results obtained in a monoenergetic positron beam (see Figure It is expected to hold exactly if the resolution is caused entirely by statistical fluctuations of the number of photoelectrons produced

in the cathodes and initial dynodes of the phototubes.

The shape of the cross section used for generating  $Y_{ij}$  is not critical; our geometry allows photoproduction to be observed only within a restricted kinematical range. It can furthermore be anticipated that when fitting the eta region the freedom of the scale factors and the resolution to vary would compensate for any reasonably small error in the shape of the cross section. We use a cross section with the shape of the cross section for photoproduction from protons and with an unknown proportionality constant to be determined for each run by the data of that run.

In terms of the above-described parameters, the final step in the iterative procedure involves maximizing the likelihood for the fit of  $\bar{N}_{ij}$  in equation III C.3 to the observed  $N_{ij}$ . Details of the fitting are given by McNeely. <sup>(27)</sup>

When fitting the eta peak, we increase the size of the "eta region." To see why, let us consider the effect on the background fit of errors in the estimated eta leakage. The maximum likelihood solution will try to adjust  $b_1$  and  $b_2$  so as to include this leakage, but background data far from the eta peak will carry enough weight to force a compromise. Some information about the shape of the eta peak will therefore be available in the background region after the background fit is made; so it makes sense to widen the eta region in order to better determine the parameters of the eta peak.

We have completed the description of one pass through the iterative procedure. The next step would be to go back to equation III C.3 as applied to the background region and again evaluate  $b_1$  and

$b_2$ . The end result of the procedure is a fit to both regions of the  $E_1$ - $E_2$  plane. For this fit we can compute the value of chi-squared per degree of freedom for each of the two regions of the plane (the number of degrees of freedom in a region is the number of  $i, j$  bins in that region minus the number of parameters which can be varied in the fitting to that region). Chi-squared per degree of freedom is expected to come out about one if our assumptions do not contradict the data. In practice, when we were deciding whether or not the fits obtained by the above described method were reasonable, we did not use chi-squared as calculated from all the bins in a region. Instead, we restricted the bins to ones with an expected number of counts greater than .5. As is explained in Appendix E, this restriction allows a stronger test of the goodness of fit than the unmodified chi-squared.

Figure 12 shows the computed background,  $\bar{B}_{ij} + \bar{A}_{ij}$ . Figure 13 displays the estimated foreground =  $N_{ij}$  - the computer background, and Figure 14 displays the calculated  $\bar{Y}_{ij}$ . Figure 15 shows the difference between the estimated foreground and the calculated  $\bar{Y}_{ij}$ . The spike at low  $E_1, E_2$  in Figures 13 and 15 occurs below the low mass cutoff of 240 MeV. All these figures are for the run whose pulse-height distribution,  $N_{ij}$ , is shown in Figure 10.

For a set of about fifty runs with especially prominent eta peaks we examined the scale factors and resolution (for this purpose

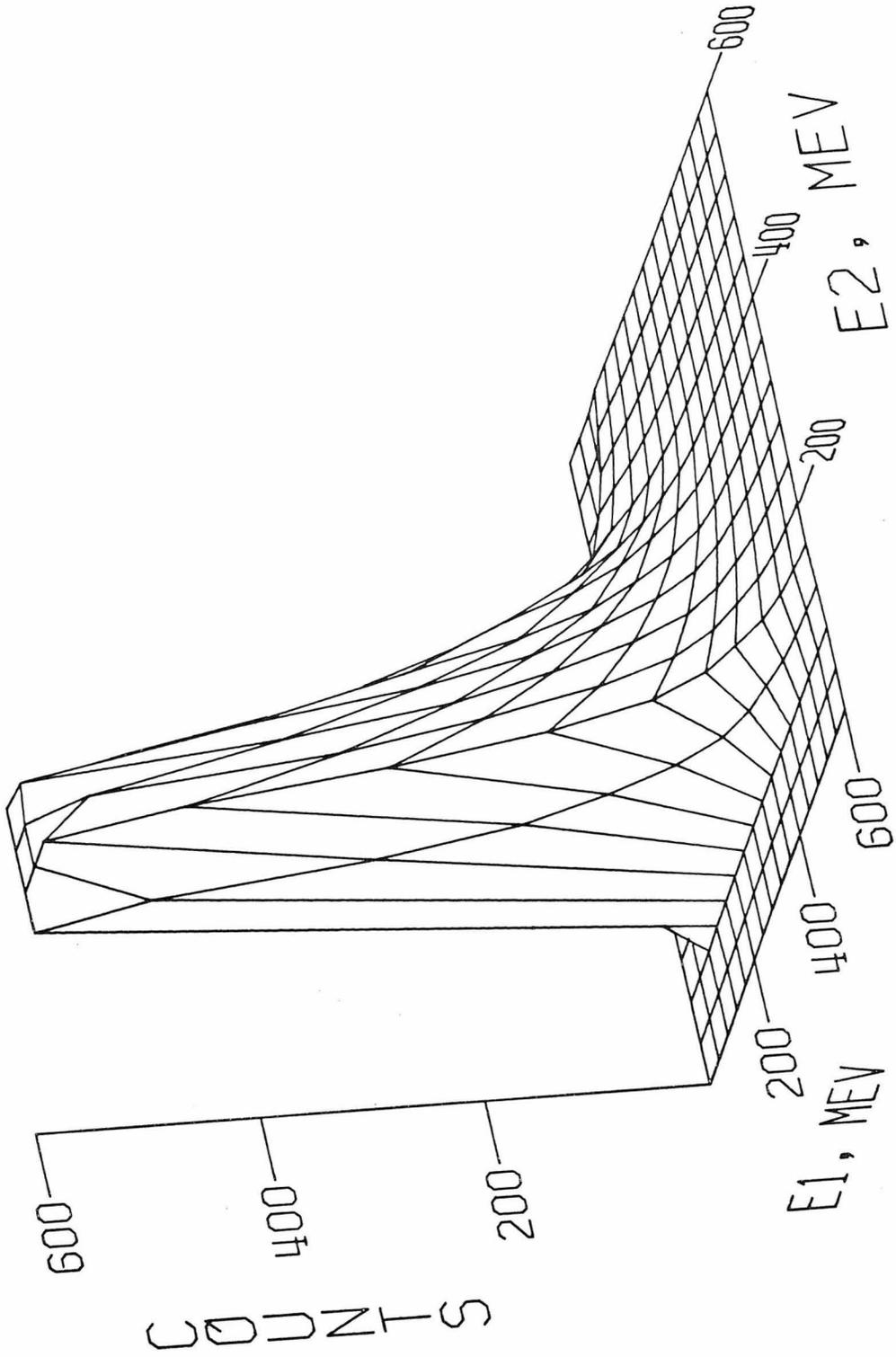


Figure 12. Computed Background for a Typical Run.

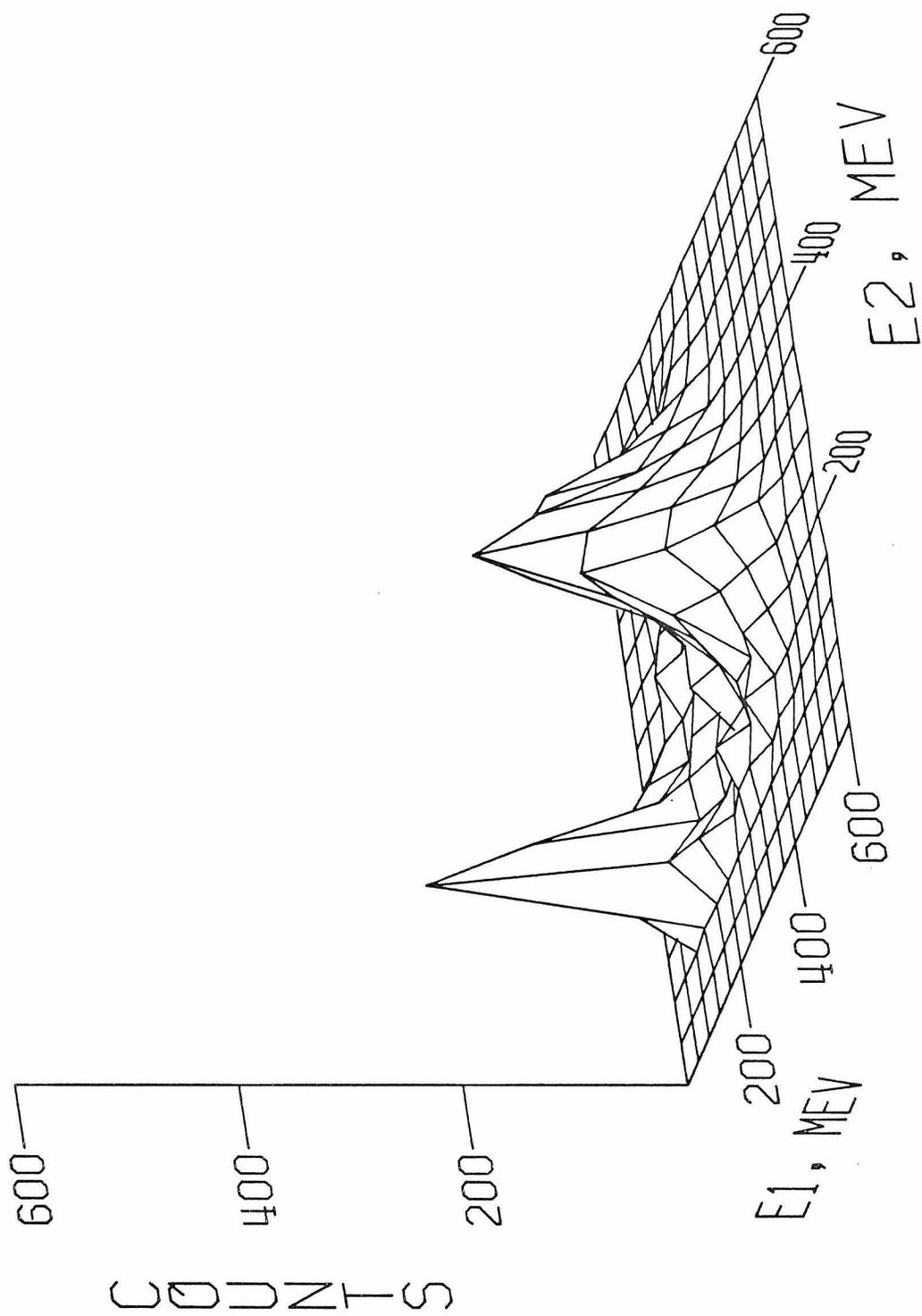


Figure 13. Pulse-Height Distribution Minus the Estimated Background.

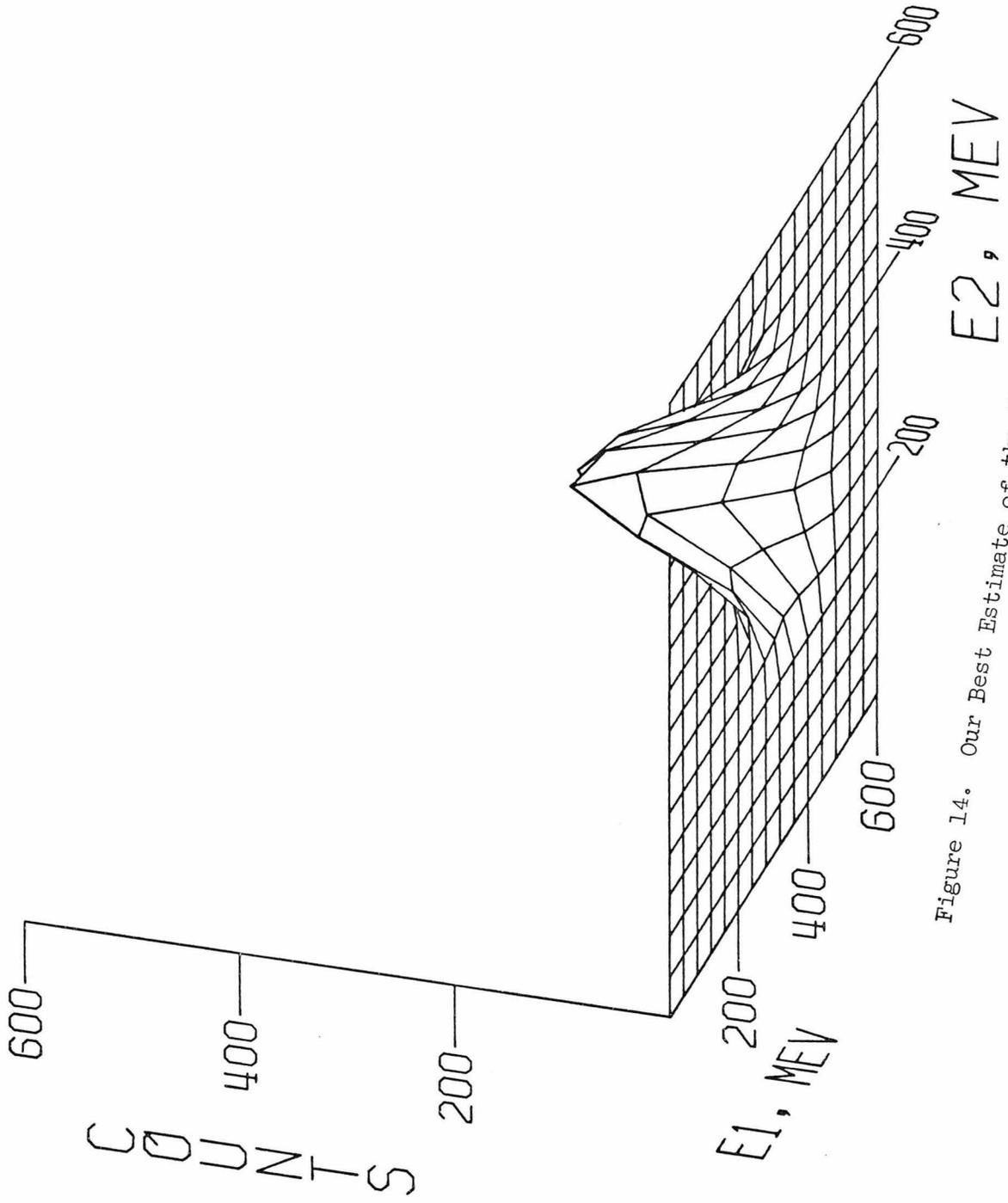


Figure 14. Our Best Estimate of the Foreground.

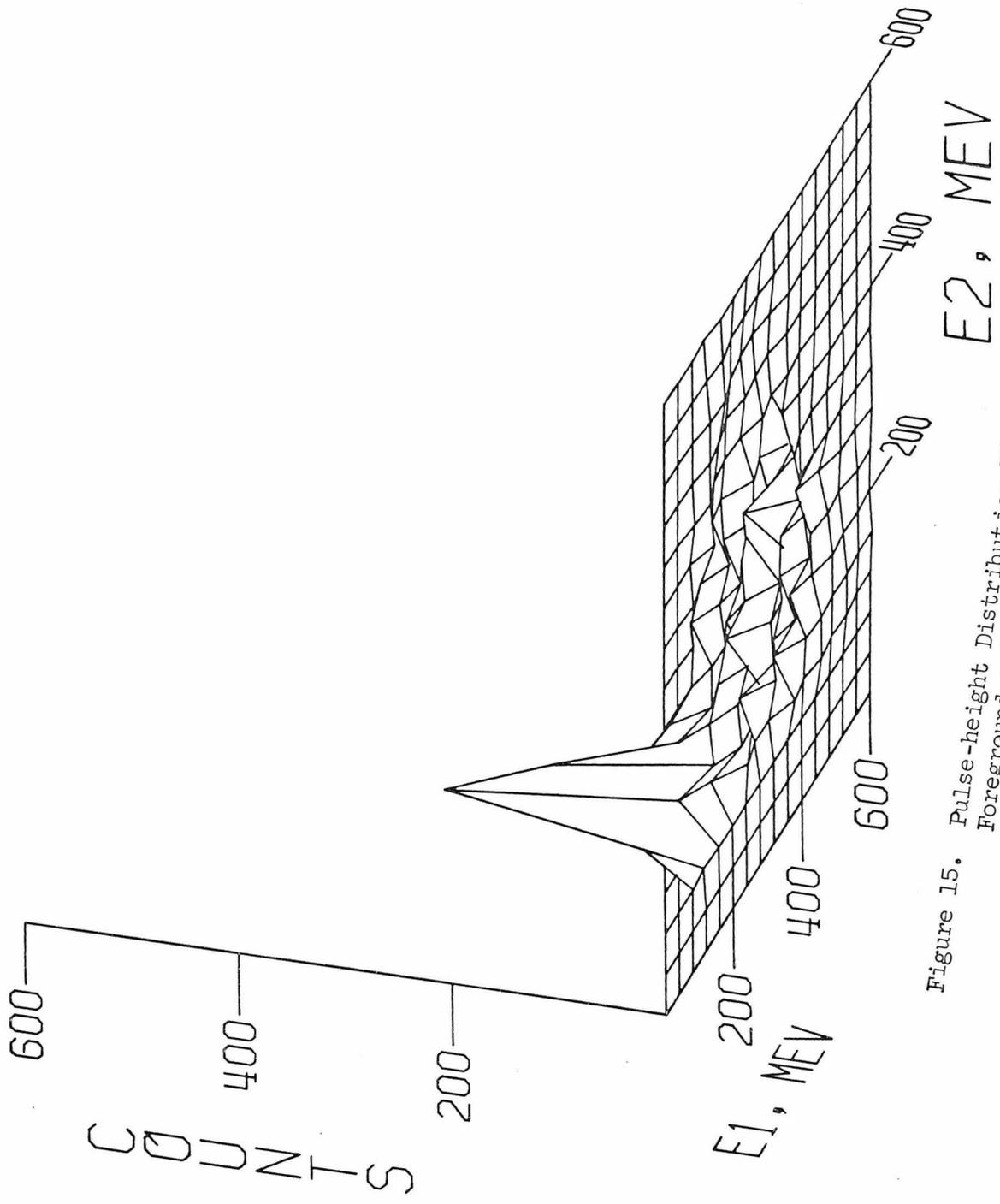


Figure 15. Pulse-height Distribution Minus the Combined Estimated Foreground and Background.

we used runs with hydrogen targets as well as ones with deuterium targets). We found that the scale factors could be usually predicted from the radioactive source peaks (see section II D) to within  $\pm 2\%$ , and the value of  $\alpha$  defined above could be generally predicted to within 15%. We observed a puzzling phenomenon:  $\alpha$  appeared to depend on the half angle of the Cherenkov counters. For angles greater than eighty degrees  $\alpha$  was typically .255, for angles less than fifty degrees it was typically .165, and for angles between fifty and eighty degrees it was typically .215 (with a standard deviation in  $\alpha$  of about .03). One cause of the broadening of the Cherenkov counter resolution was the possibility of low energy electrons, positrons, or photons entering the counter so close in time to the triggering event that the low energy events contribute to the measured pulse-height. This contribution to the broadening could explain the decrease in energy resolution with decreasing counter angle, for the background rate increases with decreasing angle. But so great is the uncertainty in the spectrum of low energy particles that we cannot tell whether or not this cause could produce as large an effect as was observed. In any case, when we say that  $\alpha$  was predictable to within 15%, we mean that given the angle of the counters, the error in the prediction of  $\alpha$  had a standard deviation of about 15% of the value of  $\alpha$ .

We conclude this section with one final comment about the usage of the above-described fitting method. Many runs had eta peaks so small that  $\alpha$ , and sometimes also the scale factors, could be far better determined by analogy with the runs with prominent eta peaks than by the data of the run itself. In such cases, we used the

maximum likelihood method with a priori probability distributions of  $\alpha$  and the scale factors. The a priori distributions were assumed to be gaussian with mean value equal to the value we would expect from the runs with prominent eta peaks. The standard deviation of the a priori distribution of  $\alpha$  was taken to be .08 and the standard deviation of the scale factors was chosen to be about 4%. Qualitatively, such a priori distributions amount to a weak clamping of the corresponding parameters. For runs with very obscure eta peaks, the maximum likelihood solution must settle near the maxima of the a priori distributions. For runs with prominent eta peaks, the solution must be almost unaffected by the presence of an a priori distribution (because in such cases the standard deviations of the a priori distributions are well above the standard deviations of the parameters as determined by the data alone). In Appendix E we explain how the maximum likelihood equations are modified by the assumption of a priori knowledge.

#### D. Corrections to the Yield

The yields as obtained in section III C were corrected for the effects of the empty target contribution, contaminants in the deuterium, dead time in the electronic circuitry, and the bias in our method of fitting the data. We now discuss those corrections.

In addition to passing through the deuterium, the beam had to pass through a total of sixteen mils of mylar. Making the approximation that the cross section for photoproduction of etas from mylar is about

the same per nucleon as that for photoproduction from deuterium (mylar =  $C_5H_4O_2$  resembles deuterium in that it contains approximately equal numbers of protons and neutrons) we can predict that typically three percent of the yield will come from the mylar when we use the relatively long targets (about  $6\frac{1}{2}$  inches) and about six percent from the mylar for the shorter 3.3" target. The size of this contribution was often non-negligible compared to the statistical error, and the effect was systematic. To correct for the empty target contribution, suppose we pick some reasonable cross section and use it with a Monte Carlo program to estimate the yield from the deuterium and the yield from the mylar, using the fact that the mylar was spacially concentrated somewhat differently from the deuterium. While the individual yields so calculated depend strongly on the size of the cross section used, the ratio between the expected yields depends only on the shape of the cross section. Because restrictions are imposed on the kinematics by the synchrotron end-point energy and by the counter geometry, the ratio of the yields cannot be sensitive to the shape of the cross section used (so long as the shape is realistic). We used a cross section approximately equal to that for photoproduction of etas from hydrogen, found the expected ratio between the contribution from the mylar and that from the deuterium, and corrected the yields by subtracting the appropriate fraction for each run.

Nitrogen and oxygen were the major impurities in the deuterium, comprising a total of 2.3% by weight of the contents of the target. Since nitrogen and oxygen each have about the same number of protons as neutrons, we can approximate their contribution to the yield in the

same way we approximated the effect of the mylar. The densities of liquid oxygen and nitrogen are so much larger than that of deuterium that we can approximately say that the amount of deuterium in the target is not appreciably decreased by the presence of those impurities. To correct for the effect of the impurities, then, we can simply decrease each yield by 2.3%. Instead, and equivalently, we decreased the cross section finally obtained by 2.3%.

Turning to the correction for dead time, in the logic which decided whether to accept an event, we consider two different sources. The veto counters, while intended to veto charged particles entering the counters, could accidentally veto photons. While we did not monitor the veto rates, we did monitor the rates of the Cherenkov counters, and tests showed that at settings for which those rates were high the veto rates were about thirty times higher. A simple estimate of the expected dead time showed that less than 3% of the photons were vetoed at all settings used (when the dead time from the veto counters was estimated at greater than 3%, the run was rejected). Our failure to monitor the veto rate constituted an error in our method, for the above mentioned estimate of the veto-induced dead time cannot be relied upon to within better than a factor of three. We did not correct for this source of dead time, but instead considered it to be a contribution to the error in our results.

Another source of dead time in our detection system was the time required for analogue-to-digital conversion of the signals whose pulse heights constituted the data of this experiment. Because a "master" scaler counted the total number of coincidences regardless of whether

or not there was time to digitize them, we could easily measure and correct for this source of dead time. This correction typically was between one and three percent. But for some empty target runs, runs with especially low counting rates, we noticed that the apparent dead time was far larger than the amount we would estimate from the time it took to pulse-height analyze signals. We believe that the "master" scaler was occasionally double counting or firing from electronic noise in the laboratory. For this reason, we looked with suspicion on runs with over five percent apparent dead time, and ultimately decided to reject one with 13% apparent dead time.

Lastly we mention the bias of our fitting procedure for the yields. When finding parameters through the maximum likelihood method, a systematic error is introduced. In Appendix E a formula for the approximate correction of this error is derived. This correction was applied to each of the yields, but because we excluded runs of less than eighty bips, the correction, which is inversely proportional to the statistics available, was small (generally less than two percent).

#### E. Extracting the Cross Section from the Yields

As is discussed in Appendix G, in the framework of the impulse approximation, <sup>(44,46,47)</sup> the yield can be considered to come partly from photoproduction from single nucleons in the deuteron and partly from an interference term. In the interference term, the eta must be considered to be photoproduced from both the nucleons at once. For now, we consider only the non-interference contribution to the yield.

The total cross section from non-interference terms is the sum of the cross sections from the individual nucleons.

The resolution in  $\underline{K}$  depends heavily on the setting, and is typically worse at higher energies. For example, for a setting with synchrotron nominal end-point of 725 MeV and counter half angle of 65 degrees, the range over which the cross section is measured is about 25 MeV. Such a setting corresponds to a near threshold measurement of photoproduction of eta mesons from a single nucleon. For a setting at 875 MeV, 40 degrees, the kinematic uncertainty is about 80 MeV; for 90 degrees, 975 MeV, the uncertainty is about 175 MeV. These uncertainties are typically five to ten percent worse than the corresponding ones for photoproduction from hydrogen because nucleons in the deuteron have Fermi motion.

The angular resolution is best at high energies. For the setting at 65 degrees, 725 MeV, the detection efficiency is almost independent of  $\theta^*$ . But for 40 degrees, 875 MeV, the production is detected almost exclusively for  $\cos \theta^*$  less than .1 from the value it takes for directly forward photoproduction (where  $\cos \theta^* = 1$ ). For 90 degrees, 975 MeV, eta photoproduction is detected almost exclusively with  $\cos \theta^*$  a distance less than .1 from its value for directly backward eta photoproduction, where  $\cos \theta^*$  is -1.

We were not, in this experiment, attempting to determine the cross section as a detailed function of  $\cos \theta^*$ . Our main problem involved unfolding the cross section as a function of  $\underline{K}$  while using yields measured with poor  $\underline{K}$  resolution. The situation facing us can

be described as follows (neglecting for now the interference term of the cross section): From Appendix G, and especially equations G.7 and G.8, the eta yield from deuterium for a particular kinematical setting can be expressed as

$$\text{Yield} = \int d\underline{K} \sum_{J=0}^{\infty} g_J(\underline{K}) f_J(\underline{K}) , \quad \text{III E.1}$$

$$\frac{d\sigma_p(\underline{K}, \cos \theta^*)}{d\Omega^*} + \frac{d\sigma_n(\underline{K}, \cos \theta^*)}{d\Omega^*} = \frac{q^*}{K^*} \sum_{J=0}^{\infty} f_J(\underline{K}) P_J(\cos \theta^*) .$$

$\frac{d\sigma_p(\underline{K}, \cos \theta^*)}{d\Omega^*}$  represents the differential cross section at center-of-mass angle  $\theta^*$  for photoproduction of eta mesons from protons at rest using photons of laboratory energy  $\underline{K}$ .  $\frac{d\sigma_n}{d\Omega^*}$  represents the corresponding differential cross section from neutrons.  $g_J$  is defined in G.9. It is a different function for each setting, and it includes such effects as the probability of detecting the eta given that it decays into two photons, the probability of the two-photon decay mode, the smearing of the kinematics caused by the nucleon Fermi motion, and the effect of folding in the synchrotron bremsstrahlung spectrum. From a set of yields known to within certain experimental errors, and from a corresponding set of known  $g_J(\underline{K})$  (computed by means of a Monte Carlo program), we wish to approximately unfold from III E.1 the unknown  $f_J$  for at least the lowest values of  $J$ .

We used two unconventional methods for extracting the cross sections from the yields. The first is discussed in Appendix C, and

is a generalization of the familiar technique in which a synchrotron end-point subtraction is performed -- i.e., in which an experiment is performed using bremsstrahlung from two slightly different end-point energies and the appropriately normalized results are subtracted. If the yields are appropriately normalized, then the difference between the two yields will receive contributions from the cross section primarily between the two different energies at which they were taken. Such a subtraction increases the statistical uncertainties in the value of the cross section determined, but decreases the energy range over which the cross section is being measured. Rather than simply making a single subtraction of yields, we found a method for choosing more general linear combinations of the yields such that the resolution is in some sense optimized. We considered only  $J = 0$ , and classified the settings into those corresponding to forward photoproduction and those corresponding to backward photoproduction. Then III E.1 becomes something of the form

$$\text{Yield}' = \int d\underline{K} g'(\underline{K}) f(\underline{K}) ,$$

where  $\text{Yield}'$  is a linear combination of the yields,  $g'(\underline{K})$  is the same linear combination of the  $g_0(\underline{K})$  for all forward (backward) settings, and  $\frac{q^*}{K^*} f = \frac{d\sigma_p}{d\Omega} + \frac{d\sigma_n}{d\Omega}$  in the forward (backward) direction. With carefully chosen linear combinations, the  $g'(\underline{K})$  can be chosen to be non-negligible only for  $\underline{K}$  within a narrow range.

All things considered, the above described methods lose information. Consider, for example, the method of end-point subtraction.

The information contained in the values of a pair of yields is also contained in the pair of numbers  $Y_+ =$  the sum of the two yields and  $Y_- =$  the difference of the two yields. But when making a synchrotron end-point subtraction,  $Y_+$  is thrown out because it corresponds to poor energy resolution. Thus information obtained during the experiment is lost during the analysis. Such a method does, nonetheless, allow one to describe the results of the experiment in terms of a set of cross section values with error bars, rather than as a set of yields with errors and with given resolution functions  $g(\underline{K}, \cos \theta^*)$ . We will say no more about this method because the details are in Appendix C, and the results from using it are in part IV.

The second method we used involved finding a function that (1) on a priori grounds is a plausible looking cross section, and (2) would lead to approximately the observed yields if it were in fact the true cross section. A well-known method of obtaining such a function,  $f(\underline{K}, \cos \theta^*)$ , is to pick a function of both  $\underline{K}$  and several unknown parameters such that for a wide range of the parameters the function satisfies condition (1); then vary the parameters until condition (2) is satisfied. A common criterion for best satisfying condition (2) is that of minimum  $\chi^2$ . For a given set of parameters, one computes the expected yields, then computes the value of  $\chi^2$  for the comparison between the observed and computed yields, and finally varies the parameters until  $\chi^2$  is minimized. An advantage of methods based on criteria (1) and (2) is that such methods, unlike synchrotron end-point subtraction and its generalization, do not give up experi-

mental information. It is possible, to be sure, that for a given experiment there are two or more completely different functions that satisfy criteria (1) and (2). But with care in choosing the experimental settings, such problems need not arise. One defect of such methods is that the resulting functions do not indicate clearly at what energies the function is determined by data and at what energies the function is only constrained by the a priori assumptions. For this reason, such methods should be used in conjunction with some method, such as end-point subtraction, for which the location of the data is clearly displayed.

A drawback to the use of minimum  $\chi^2$  can be seen by the following example. Suppose the a priori condition one had in mind was that the desired function should change slowly. In that case a parameterization would be chosen that would result in a slowly changing function when  $\chi^2$  is minimized. But then it could happen that after minimizing  $\chi^2$ , there would be for the same  $\chi^2$  a function that varies more slowly than the one found. If there is such a function, and if slowly varying functions are indeed more plausible on a priori grounds, then minimum  $\chi^2$  would have not found the function that best satisfies criteria (1) and (2). There are two reasons such a more plausible solution can, and almost certainly will, exist. In the first place, as soon as one begins minimizing  $\chi^2$ , the a priori preferences are no longer being taken into account. Functions with better a priori plausibility are not given special weight. In the second place, even if one has a particular set of a priori criteria in mind when choosing

the form of the function to be found, the form chosen may not be the best form for satisfying the a priori criteria. An a priori preference such as slow variation cannot be easily translated into a particular functional form. As soon as one has chosen a particular form for the function, the solution has been distorted in an unknown manner from the most plausible one.

The reader should understand that the above mentioned drawbacks are not fatal ones. If the data are good, then any fit that reasonably well satisfies criteria (1) and (2) will probably be pretty close to truth. But it is desirable to find some method that is able to avoid some of the disadvantages of minimum  $\chi^2$ . For this reason (and also because doing things in a different way is more fun) the objections to minimum  $\chi^2$  were met for a certain class of a priori criteria by a method described in Appendix D. In part IV, we try several a priori criteria in order to see how sensitive our results are to the initial assumptions. For example, one criterion we tried was that the function purporting to be the cross section must be smooth -- i.e., the derivative of the function must not change rapidly with  $\underline{K}$ . A reasonable measure of the smoothness of a function is

$$Y = \iint d\underline{K} d\cos(\theta^*) \left[ \frac{\partial^2 f(\underline{K}, \cos \theta^*)}{\partial \underline{K}^2} \right]^2 . \quad \text{III E.2}$$

When this quantity is small,  $f$  is smooth in  $\underline{K}$ ; otherwise it is not. Because we only took data near extreme forward and backward settings (except near threshold), it is reasonable to express this fact by taking  $f(\underline{K}, \cos \theta^*) \approx f_0(\underline{K}) + f_1(\underline{K}) \cos(\theta^*)$  and considering this form

to be valid only at  $\cos(\theta^*) = \pm 1$  (again except near threshold, where this form for  $f$  is valid at all  $\theta^*$ ). In order to avoid the possibility of there being a function which is smoother than the one we find, but which fits the data equally well, we set ourselves the mathematical problem of finding the smoothest function that fits the data to within a given value of  $\chi^2$ . This problem is solved in Appendix D. Which value of  $\chi^2$  is chosen is somewhat arbitrary; so it is wise to try several in order to see if any conclusions depend on how strongly one demands that the assumed function fit the data. In III E.2, the second derivative with respect to  $\underline{K}$  can be replaced by the  $n$ 'th derivative for any  $n \geq 1$ , thereby providing alternative measures of "smoothness."

Before we go on to display the results, we point out that  $f_J$  as defined in III E.1 is the  $J$ 'th Legendre moment of the cross section with the threshold factor,  $q^*/K^*$ , removed. We remove the threshold factor because our a priori assumption of smoothness is unreasonable if  $f_J$  has such a factor. It is easy, however, to write down an equation analogous to III E.1 but with the threshold factor remaining in the function assumed to be smooth. In this way we give ourselves a test of the effect of false assumptions on our method of unfolding the cross section.

Finally, we mention the effect of the interference term on the cross section (remember that until now we have been neglecting it). Equation G.11 gives a way of computing the contribution to the yield of the interference term given the amplitudes for photoproduction from

protons and neutrons individually. This contribution turns out to be generally less than ten percent for reasonable amplitudes, and it is not too terrible to simply neglect the effect. The cross section that results from neglecting the interference term can be used to estimate the photoproduction amplitude from neutrons. Then taking the amplitudes from say, the  $S'_{11}$  resonance as it appears in photoproduction from protons and from neutrons, one may estimate the interference contribution to the yield and see how this contribution can affect the results for our estimate of the non-interference term.

## IV. RESULTS

Figure 16 shows the cross section for eta photoproduction from deuterium. In this figure, there are two sets of three curves. The cross section is of the form

$$\frac{d\sigma_p(\underline{K}, \cos \theta^*)}{d\Omega^*} + \frac{d\sigma_n(\underline{K}, \cos \theta^*)}{d\Omega^*} \approx f_0(\underline{K}) + f_1(\underline{K}) \cos \theta^*$$

where subscripts "p" and "n" refer to the proton and neutron respectively, where  $\underline{K}$  is the photon energy in the rest frame of a nucleon, and where  $\theta^*$  is the center-of-mass angle of eta photoproduction. The set of three curves that start up from threshold and reach a peak of about two microbarns per steradian represent the  $f_0$  term. The center curve represents the value of  $f_0$ , while the two outer curves represent  $f_0 \pm$  estimated error in  $f_0$ . Similarly, the other three curves represent  $f_1$ , and  $f_1 \pm$  error in  $f_1$ . This set of six curves was generated from the data and detection efficiencies by Blackbox, a method and computer program described in III E and Appendix D. The criterion for "smoothness" used for generating Figure 16 was that of small

$$\int_{\text{Threshold}}^{1350 \text{ MeV}} d\underline{K} d \cos \theta^* \left[ \frac{\partial^2 (f_0 + f_1 \cos \theta^*)}{\partial \underline{K}^2} \right]^2$$

$f_0$  and  $f_1$  were not assumed to have the threshold factors  $q^*/K^*$  (from phase space) removed. Although forty-two settings were included in the data for Figure 16, only the twenty most important linear combina-

tions were used with a  $\chi^2$  of twenty (see Appendix D in order to understand how we separate out the "most important" linear combinations). Figure 16, then, represents the smoothest fit to the data with a  $\chi^2$  of twenty, with twenty settings, and with no threshold factor assumed to be present in the cross section. Notice that at energies above about 1125 MeV the errors increase sharply. Such large errors indicate a lack of data so severe that any cross sections displayed should not be taken seriously above 1125 MeV. The true cross section must go to zero at threshold.  $f_1$  is consistent with zero at threshold -- the observed small deviation from zero is consistent with the statistical errors we know to be present. But the deviation of  $f_0$  from zero at threshold appears to be inconsistent with the errors we display. At this point we must examine more closely what the errors given mean. The "error" represents the effect of statistical fluctuations in the data on the computed cross section when  $\chi^2$  is held fixed. The failure of  $f_0$  to go all the way to zero at threshold in Figure 16 is primarily a consequence of the fact that the assumption of a small second derivative of  $f_0$  is especially bad near threshold, where the second derivative with respect to  $\underline{K}$  of the true cross section is infinite. Statistical fluctuations in the data are not sufficient to overcome the a priori prejudice against an infinite second derivative of the cross section.

Another point that should be made about the errors is that they are optimistic. They include only the effects of statistical uncertainties in the number of etas observed during each run. Not

included are such difficult to estimate errors as those from veto-induced dead time (III D) and errors in the assumptions made during extraction of the yield (section III C and Appendix I). One way of detecting the fact that the errors are too small is by making various consistency checks. For example, frequently more than one run was taken at the same setting with the same length target, counter position, and synchrotron end-point. We could, therefore, check to see whether to within the statistical errors assumed for the runs, there was agreement between different runs. In fact, there almost always was such agreement. The few inconsistencies that were observed would be understandable if we were underestimating our random errors by a factor of about 1.5. If we increase all the errors by a factor of 1.5, then by finding the smoothest function for  $\chi^2 = 20/(1.5)^2 = 8.9$  we would obtain Figure 16 with 1.5 times as large an error. But in the limit of a large number of settings, a reasonable fit would be such that  $\chi^2 =$  the number of settings (if we knew the true cross section, our data should be consistent with it to within that value of  $\chi^2$ ). When unfolding the cross section with smaller fixed  $\chi^2$ ,  $f_0$  and  $f_1$  can be expected to exhibit greater fluctuations corresponding to statistical errors in the data.

In Figure 17 we have displayed the effect of making several changes in the method of extracting the cross section from the yield. Instead of having the cross section smooth, we have made the squared amplitude smooth. We have done this by taking

$$f_0 = \frac{q^*}{K^*} h_0 ,$$

$$f_1 = \frac{q^*}{K^*} h_1 ,$$

(where  $q^*$  = eta center-of-mass momentum,  $K^*$  = photon center-of-mass momentum) and requiring that  $h_0$  and  $h_1$  be smooth. The result of this change, alone, is not shown in Figure 17. Instead we also change the definition of smoothness by requiring that the first derivative, rather than the second, is small. Also displayed in Figure 17 is the cross section as calculated by the generalization of synchrotron end-point subtraction as discussed in III E and Appendix C. The backward cross section ( $\theta^* = 180$  degrees) is shown as points with both horizontal and vertical error bars, with the understanding that the horizontal errors (the energy range over which the cross section was being measured for a given point) are about the same for forward cross sections at a given energy as they are for backward ones at about the same energy. While the cross section represented as points with error bars ignores most of the information obtained by this experiment (see section III E) such a representation does serve to show where our data were taken, and it serves as a check on the results as displayed by the smooth curves.

With the exception of immediately above threshold,  $f_0$  and  $f_1$  of Figure 16 agree (to within the optimistic errors given) with  $f_0$  and  $f_1$  of Figure 17. This agreement does not mean that the errors given are correct; it means that the results are insensitive to the a priori assumptions made about the cross section.

Although Figure 16 and 17 agree to within statistical errors except at threshold, there are some disagreements that are not large compared to the statistical errors but deserve comment. At about 900 MeV, Figure 16 exhibits a bump in the error of  $f_0$  and  $f_1$ , but no significant structure in  $f_0$  itself. In Figure 17, there is not only a bump in the error, there is a shoulder in  $f_0$ . Other variations of our fitting procedure show that this shoulder is weakly indicated by the data. But in the case of Figure 17, the shoulder is especially pronounced. This enhancement of the shoulder is spurious, because when the expected statistical errors on the data are large, the requirement of a small first derivative of  $f_0$  can be expected to show up as a temporarily constant cross section in the region of large statistical errors. Another disagreement between Figures 16 and 17 lies in the large difference between the errors given in the region above which we claim to have no reliable results (above about 1125 MeV). But it is reasonable that a computed cross section constrained to have a small second derivative is more sensitive to statistical variations in the data than one constrained to have a small first derivative. The demand of a small second derivative allows an arbitrarily steep cross section in the region for which little data are available.

In Figure 18 we display the results of placing unreasonable stress on the demand for smoothness by setting  $\chi^2$  to 60 (using the optimistic statistical errors). There is an a priori threshold factor in the cross section and "smooth" means "small second derivative." As can be seen, the displayed errors are small, indicating that by placing such great stress on smoothness we become insensitive to

statistical fluctuations in the data. Most of the details formerly seen in Figures 16 and 17 are smoothed out by allowing such a high value of  $\chi^2$ .

Table IV.1 shows the differential cross section in microbarns per steradian. The "energy resolution" is an eyeball estimate of the full-width at half-maximum of the corresponding linear combination of efficiency functions. The "angular resolution" is  $1 - |\langle \cos \theta^* \rangle|$ , where  $\langle \cos \theta^* \rangle$  is estimated from the angular resolution of the detection efficiency for settings taken near the energy of the linear combination involved. The errors given in the cross section do not include the systematic errors from our method of determining the yields (errors estimated from the results of Appendix I to be of the order of the statistical errors -- about  $.1 \mu\text{b}/\text{ster}$ ). This table is not considered to be as informative as Figures 16, 17, and 18. It is placed here (1) because it presents the data in a conventional, hence easily used, form; (2) because it gives the reader an opportunity to decide for himself which method of unfolding the cross section to trust; and (3) because it helps one decide which features of Figures 16, 17, and 18 are most firmly established by the data.

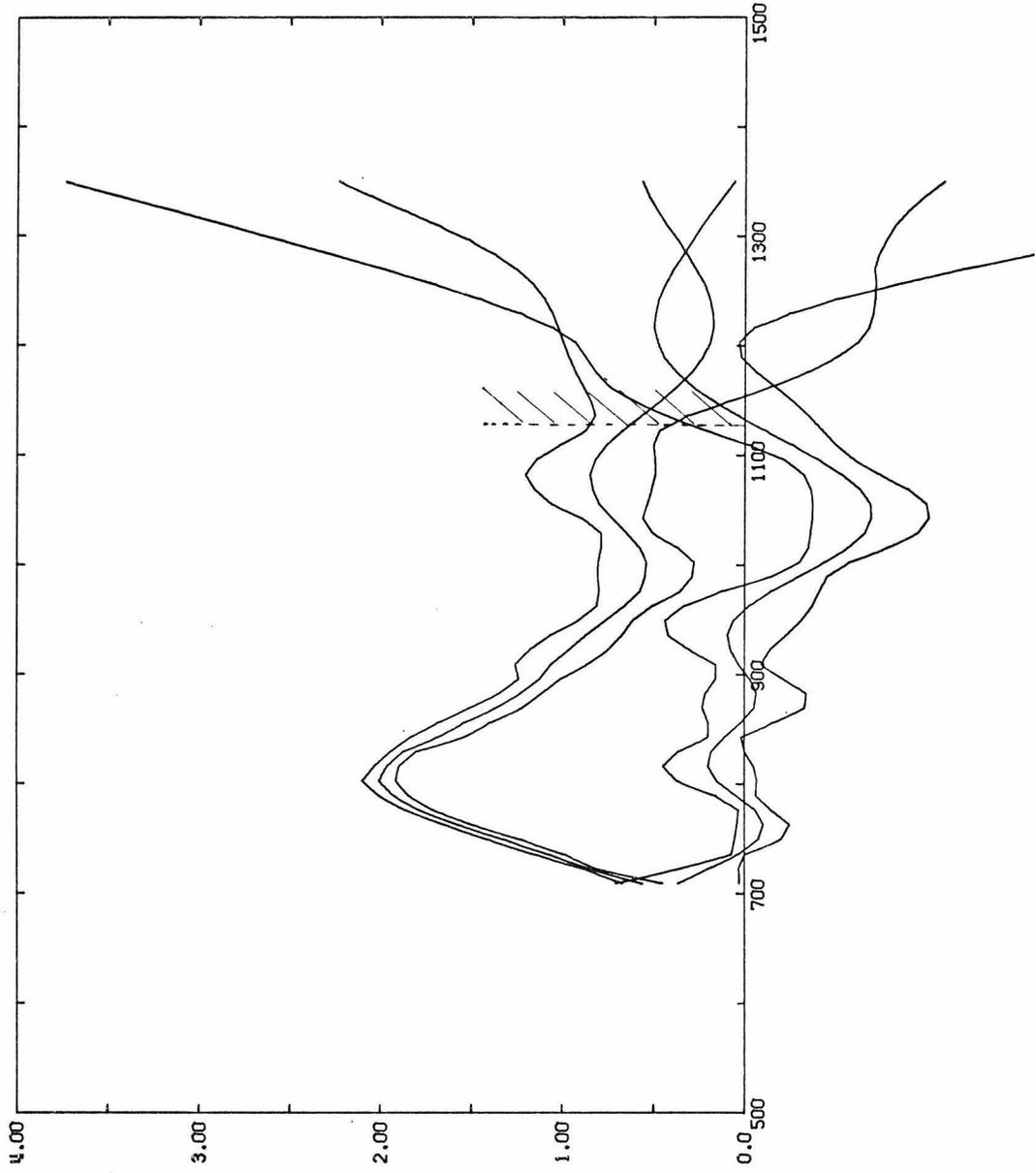


Figure 16. Differential Cross Section for Eta Photoproduction from Deuterium. No Threshold Factor,  $X^2 = 20$ .

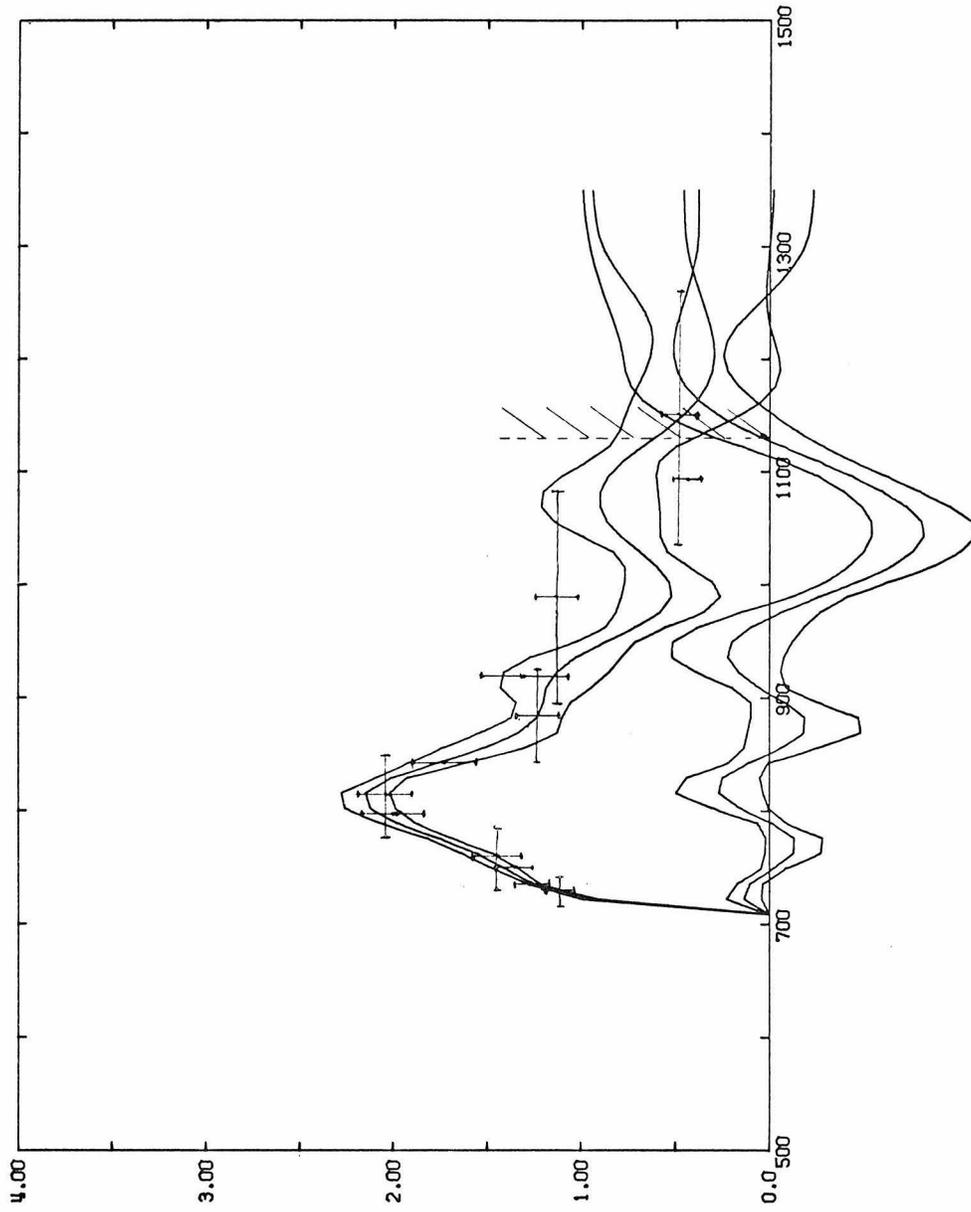


Figure 17. Differential Cross Section for Eta Photoproduction from Deuterium. Phase Space Threshold Assumed.  $\chi^2 = 20$ .

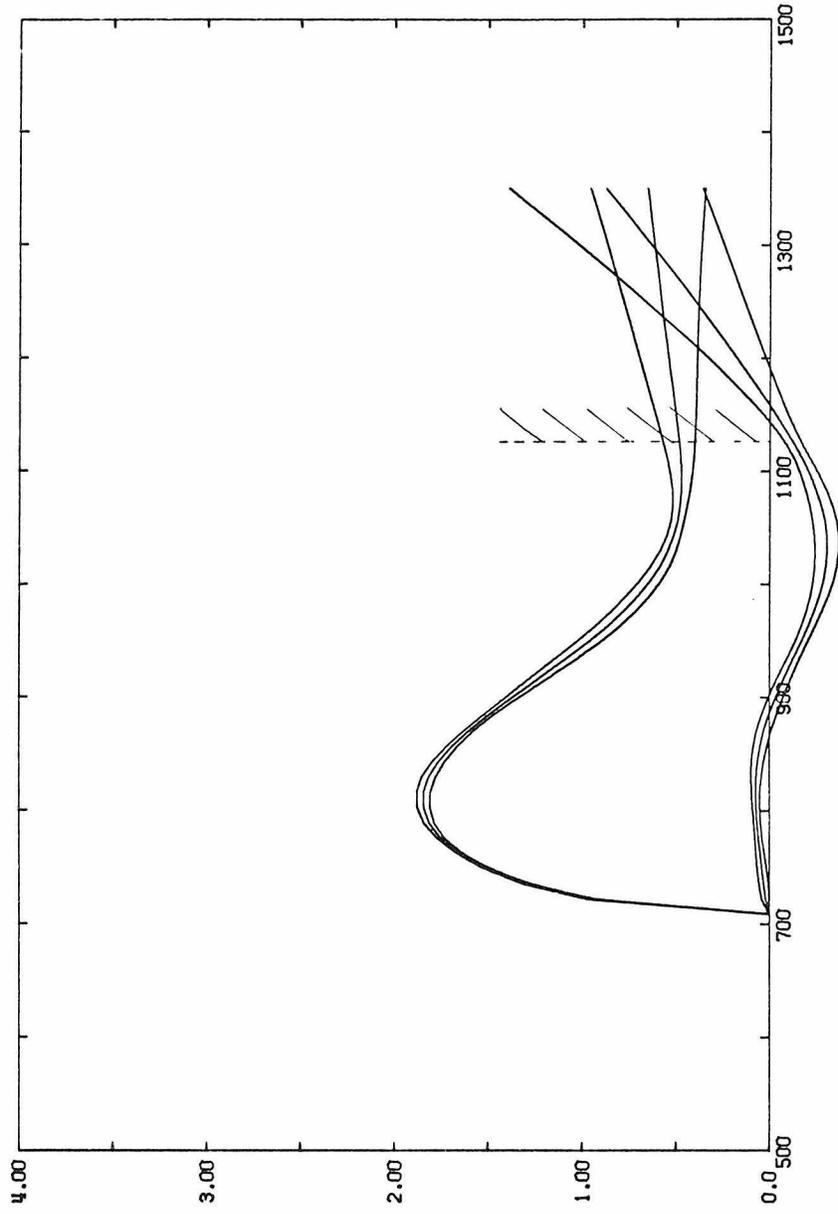


Figure 18. Differential Cross Section for Eta Photoproduction from Deuterium. Phase Space Threshold Assumed.  $X^2 = 60$ .

TABLE IV.1

Cross Section from Optimizing Resolution

Energy	Resolution		Cross Section		Error in Cross Section
	Energy (MeV)	Angle ( $\cos \theta^*$ )	Forward	Backward	
725	15	.7		1.10	.07
728	20	.7	1.25		.09
749	50	.4	1.33		.10
761	60	.35		1.40	.13
794	60	.25	1.95		.15
806	60	.3		2.01	.12
839	70	.17	1.70		.16
890	80	.12		1.21	.12
911	100	.10	1.28		.24
986	180	.09		1.13	.12
1095	140	.06	.42		.06
1154	220	.07		.46	.13

## V. CONCLUSIONS

In order of most firmly established conclusions first, we have:

1) The  $S'_{11}$  (1535) is about as prominent in photoproduction from neutrons as it is in photoproduction from protons. This conclusion is based on a comparison of Figures 16, 17, and 18 with corresponding results from a similar experiment performed at the same time but using a hydrogen target<sup>(27)</sup>. 805 MeV is within about 10 MeV of the maximum cross section for both hydrogen and for the cross sections as displayed in Figures 16, 17, and 18. At this energy, the isotropic term in the hydrogen differential cross section was found to be about  $.98 \pm .06$   $\mu\text{b}/\text{ster.}$  as compared with between 1.85 and 2.05  $\mu\text{b}/\text{ster.}$  for deuterium (depending on how the unfolding was done). We estimate that the ratio of the eta cross section from neutrons to that from protons is  $.99 \pm .10$  at 805 MeV. To within our experimental errors, the  $S'_{11}$  is photoproduced by either an almost pure isoscalar interaction or by an almost pure isovector interaction (we ignore the possibility of isotensor electromagnetic interactions).

2) Up to 900 MeV there is no detectable angular asymmetry in eta photoproduction from deuterium. If one somewhat discounts the probably statistical oscillations in the asymmetry term of Figures 16 and 17, then below 900 MeV we can be reasonably sure that the

$$\text{asymmetry} = \frac{\text{forward cross section} - \text{backward cross section}}{\text{forward cross section} + \text{backward cross section}}$$

is zero to within .1 for energies below 900 MeV, with an indication

of a slightly positive asymmetry over most of this range. Recall from the introduction that the absence of the Roper resonance ( $P'_{11}(1470)$ ) in photoproduction from protons could be explained if that resonance were in an  $SU(3)$   $\underline{10}^*$ . But if that explanation were valid, the Roper resonance could be visible in photoproduction from neutrons. By examining eta photoproduction, we isolate the  $I = 1/2$  from the  $I = 3/2$  intermediate states which also are produced in photon-nucleon interactions. The effect of the  $I = 1/2$  Roper resonance should then be seen as an angular asymmetry in the region between 750 MeV and 900 MeV, with strong energy dependence. Since such an asymmetry is as hard to see in photoproduction from neutrons as it is in photoproduction from protons, there is no longer any reason to suspect that the  $P'_{11}(1470)$  is in a  $\underline{10}^*$  irreducible representation of  $SU(3)$ .

3) From 975 MeV to 1100 MeV there is a negative asymmetry. This asymmetry also appears in photoproduction of eta mesons from protons, where it is consistent with an interference between the  $S'_{11}(1535)$  and the  $P''_{11}(1780)$ .

4) Somewhere between 1000 MeV and 1100 MeV, the isotropic part of the cross section seems to reach a minimum at about  $.5 \mu\text{b}/\text{ster.}$  and starts to turn up. Such a minimum has been observed in eta photoproduction from hydrogen<sup>(42,54)</sup> where the minimum is about  $.2 \mu\text{b}/\text{ster.}$  In the case of hydrogen, this minimum is followed by a broad peak, which again is attributed to the  $P''_{11}(1780)$ . In the case of deuterium, our experiment did not extend high enough to explore such a peak, but

it appears that the  $P''_{11}(1780)$  is at least as strong in photoproduction from neutrons as it is in photoproduction from protons.

5) In the range  $-.4 < \cos \theta^* < +.6$ , a Frascati group<sup>(23)</sup> found the deuterium cross section to be  $1.82 \pm .25$  (estimated from Figure 2 in their paper) at a photon energy of  $835 \pm 35$  MeV. Our value at this energy is about  $1.81 \pm .10$  for  $\cos \theta^*$  near  $\pm 1$ . But it is known that there is no significant  $\cos^2 \theta^*$  term in eta photoproduction at 835 MeV off protons.<sup>(21)</sup> Thus the experiments conducted here and at Frascati when compared show no  $\cos^2 \theta^*$  term in the differential cross section for eta photoproduction from neutrons. There seems to be much less of such a term in our reaction than there is at similar energy in  $\pi^- p \rightarrow \eta n$ .<sup>(20)</sup> We therefore confirm the prediction of Bietti<sup>(30)</sup> and the phase shift analysis of Walker<sup>(18)</sup> which have small  $|J_z| = 1/2$  components of the  $D'_{13}(1520)$  in photoproduction from neutrons (see part I B if you don't understand this sentence).

We wished to detect the sign of the neutron-proton interference term, but the effect turns out to be small compared with the errors of this experiment and compared with other sources of asymmetry.

To summarize our results, we see that it looks as if  $S'_{11}$  and  $P''_{11}$  can be invoked to explain  $\gamma + d \rightarrow \eta + p$  over the entire energy range and to within the errors of our experiment. The absence of any effect from the  $P'_{11}(1470)$  eliminates the Roper resonance as a plausible candidate for a member of an exotic multiplet. Its absence in photoproduction from protons is no longer a valid reason for suspecting it is a member of a  $10^*$ . Furthermore, from equation I A.3, we see that our experimental results require both the  $S'_{11}(1535)$  and the  $P''_{11}(1780)$

to be members of octet irreducible representations of  $SU(3)$ . They cannot be members of an  $SU(3)$   $\underline{10}^*$  because both are seen in photoproduction from protons. They cannot be members of an  $SU(3)$   $\underline{27}$  because their contribution from photoproduction off neutrons is comparable to that from photoproduction off protons. In the notation of Table I.1, the  $S'_{11}$  has  $\alpha_\gamma$  either very large in magnitude or approximately equal to  $3/4$ .

A more quantitative examination of the contribution of the states discussed here is in progress, and will hopefully soon give us a firmer handle in establishing the points touched on here; however we are confident that a more detailed discussion will have little effect on the prominent features of the conclusions presented here.

## VI. APPENDICES

## APPENDIX A Gain Matching

Certain assumptions must be made in order to conclude that phototube gains on a Cherenkov counter should be matched for best resolution.

Suppose we neglect statistical variations of the position and development of showers and consider only variations due to random fluctuations in the pulse height in each phototube for a given shower energy and development. Then it is plausible to assume that there is essentially no correlation between pulse height fluctuations in different phototubes. This assumption is the first one we shall use.

For the second of our assumptions, we say that if  $M_i$  and  $\sigma_i$  are the mean and standard deviation of the pulse height from tube  $i$ , then  $\frac{\sigma_i}{M_i}$  is independent of the phototube voltage. To see why this assumption is reasonable, suppose  $\sigma_i$  comes primarily from random fluctuations in the number of photons that convert. If the expected number of photoelectrons at the cathode of tube  $i$  is  $\bar{n}_i$ , then the probability of  $n_i$  photoelectrons is given by the Poisson distribution:

$$P(n_i) = \frac{(\bar{n}_i)^{n_i}}{n_i!} e^{-\bar{n}_i}$$

Because the phototube voltages were such that output pulse heights were at most a few tenths of a volt, and because we had capacitor banks on the last dynodes, saturation effects were eliminated. Hence the total pulse height can be taken proportional

to the number of photoelectrons, and we can speak of the gain of a phototube,  $G_i =$  pulse height per photoelectron. The Poisson distribution leads to

$$M_i = G_i \bar{n}_i$$

$$\sigma_i = G_i \sqrt{\bar{n}_i}$$

$$\frac{\sigma_i}{M_i} = \frac{1}{\sqrt{\bar{n}_i}}$$

When the phototube dynodes are operating in their proper voltage range, the number of photoelectrons should not change much with changing gain, for the voltage on a given dynode does not change much when the overall gain is changed substantially.  $\frac{\sigma_i}{M_i} = \frac{1}{\sqrt{\bar{n}_i}}$  should be independent of voltage over a large range of gain.

This interpretation of the cause of the fluctuations is confirmed by Figure 19, which shows that  $\frac{M^2}{\sigma^2}$  was approximately proportional to shower energy.

Even if the phototube pulse heights do not follow Poisson distributions, we can define  $G_i$  and  $\bar{n}_i$  by the above equations for  $M_i$  and  $\sigma_i$ , and similarly we can define overall  $G$  and  $\bar{n}$  in terms of the overall  $M$  and  $\sigma$ . When we speak of numbers of photoelectrons, we are referring to  $\bar{n}_i$  and  $\bar{n}$  even though these numbers may not really be physically the number of photoelectrons produced at the photocathode. For example, statistical fluctuations in phototube stages after the cathode can decrease these measured quantities and

our analysis will still be correct.

With pulse height nearly proportional to shower energy, maximum resolution means minimum  $\frac{\sigma}{M}$ . We use the above explained assumptions to show that  $\frac{\sigma}{M}$  is minimized when the voltages are such that the  $\frac{\sigma_i^2}{M_i}$  are all equal. This minimum value of  $\frac{\sigma}{M}$  satisfies

$$\left(\frac{\sigma}{M}\right)^{-2} = \sum_i \left(\frac{\sigma_i}{M_i}\right)^{-2}$$

In other words, we get best resolution when the  $G_i$  are all equal, in which case  $\bar{n}$  takes on its maximum value,  $N = \sum_i \bar{n}_i$ . A simple proof follows:

Independence of fluctuations in different phototubes implies

$$\sigma^2 = \sum_i \sigma_i^2 \quad \text{and} \quad M = \sum_i M_i .$$

Let the vector A be such that  $A_i = \frac{M_i}{\sigma_i}$  and the vector B be such that  $B_i = \sigma_i$ . By the Schwartz inequality,  $|A \cdot B| \leq |A| |B|$  with equality when A and B are parallel. Because  $A \cdot B = M$ ,  $A \cdot A = \sum_i \bar{n}_i$ , and  $B \cdot B = \sigma^2$ , the above stated conclusions follow trivially.

The gains of the phototubes need not be matched especially closely. To make this statement quantitative, define  $G'$  so that  $\sum_i \bar{n}_i (G' - G_i) = 0$  and define  $\epsilon_i$  so that  $G_i = G'(1 + \epsilon_i)$ . Then  $\sum_i \epsilon_i \bar{n}_i = 0$ .

We can interpret  $G'$  as a typical gain and the  $\epsilon_i$  as frac-

tional fluctuations about that gain. For nearly matched gains, the  $\epsilon_i$  are small. Note that  $G' = G\bar{n}/N$  is slightly below the overall gain as defined from the overall  $\sigma$  and  $M$ . If we define a measure of the effect of unmatched gains to be

$$f = \frac{\left(\frac{M}{\sigma}\right)_{\text{ideal}}^2 - \left(\frac{M}{\sigma}\right)_{\text{actual}}^2}{\left(\frac{M}{\sigma}\right)_{\text{actual}}^2} = \frac{N - \bar{n}}{\bar{n}}$$

then

$$f = \frac{1}{N} \sum_i \bar{n}_i \epsilon_i^2$$

can be easily shown.

Putting in numbers, if the gains are matched to within 20%, then  $f$  will be below .05. For the energies involved in this experiment,  $\frac{\sigma}{M}$  for ideal gain matching is around 10%. Then a .05 value of  $f$  will only increase  $\sigma/M$  from 10% to 10 $\frac{1}{4}$ %. As a check, using the light emitting diodes we calculated  $N$  to be 160 when  $\bar{n}$  was 154.

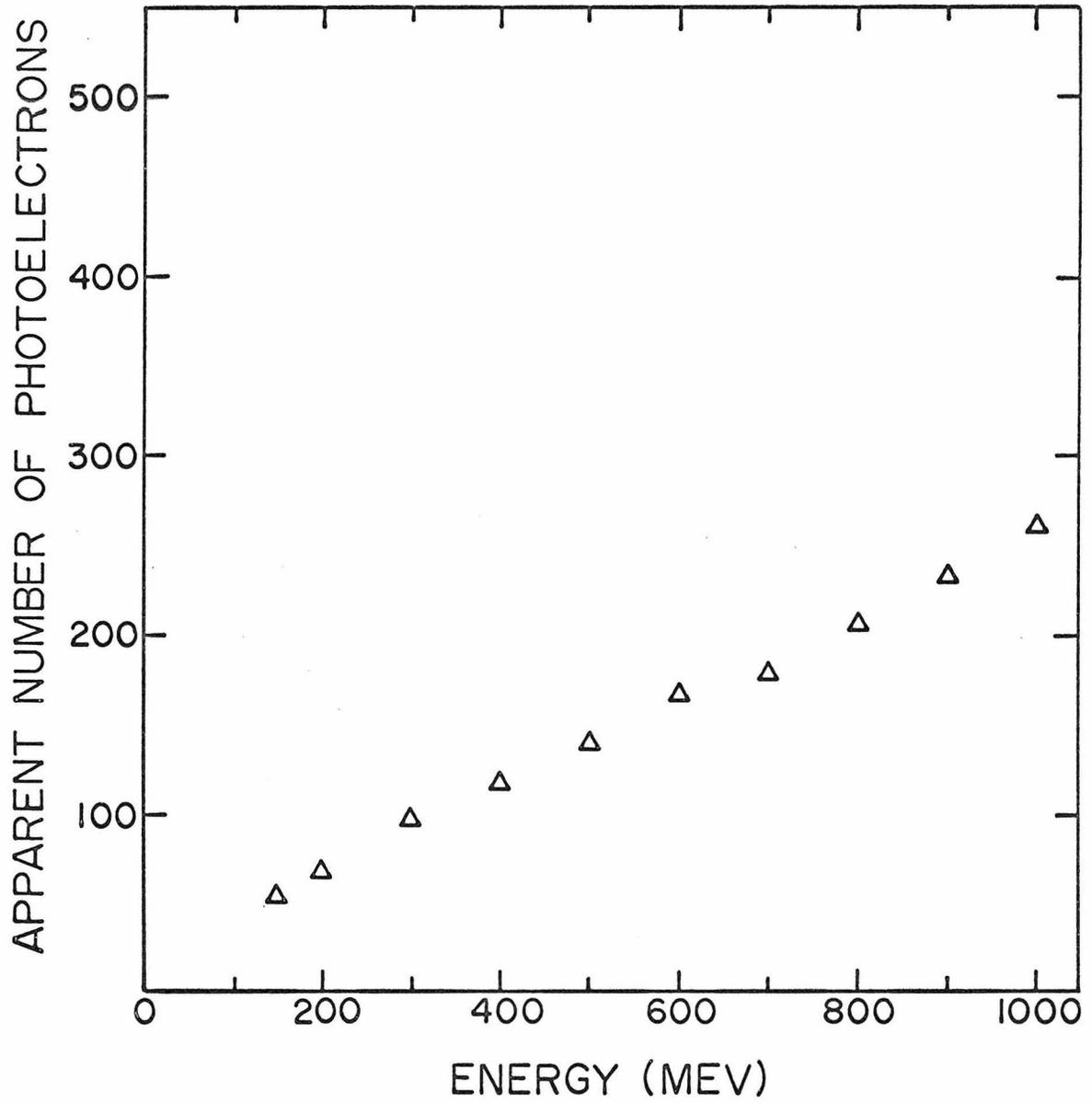


Figure 19. Apparent Number of Photoelectrons as a Function of Shower Energy.

## APPENDIX B Monte Carlo Method

The Monte Carlo method is so well known and so obvious that people can use it successfully without worrying about the mathematics behind it. To ease my mind, however, I built up a small personal store of Monte Carlo verities which I herewith present.

Suppose we can express an integral as

$$I = \int d^n x G(\vec{x}) f(\vec{x})$$

where  $f$ ,  $G$ , and the range of integration are known and where  $G(\vec{x})$  is everywhere positive. If we define

$$H = \int d^n x G(\vec{x}),$$

then  $G(\vec{x})/H$  is a probability distribution such that the expectation value of  $f$  for  $\vec{x}$  chosen according to the density function  $G/H$  is

$$\langle f \rangle = \int d^n x \frac{G(\vec{x})}{H} f(\vec{x}).$$

If  $\sigma^2(f) = \langle f^2 \rangle - \langle f \rangle^2$  is finite, then a consistent, unbiased estimate of  $\langle f \rangle$  is (5.2)

$$f_m = \frac{1}{m} \sum_{j=1}^m f(\vec{x}_j) \quad \text{B.1}$$

where the  $\vec{x}_j$  are chosen independently according to the density function  $G/H$ .

The variance of  $f_m$  is  $\frac{1}{m} \sigma^2(f)$ . If  $f^2$  has a finite variance,

then a consistent, unbiased estimate of  $\sigma^2(f)$  is<sup>(52)</sup>

$$\sigma_m^2(f) = \frac{m}{m-1} \left( \frac{1}{m} \sum_j f^2(x_j) - \left( \frac{1}{m} \sum_j f(x_j) \right)^2 \right) \quad \text{B.2}$$

From what has been said above, "I" and its variance, " $\sigma^2$ " can be estimated from

$$I \approx \frac{H}{m} \sum_{j=1}^m f(\vec{x}_j)$$

$$\sigma^2(I) \approx \frac{H^2}{m-1} \left( \frac{1}{m} \sum_{j=1}^m f^2(x_j) - \left( \frac{1}{m} \sum_{j=1}^m f(\vec{x}_j) \right)^2 \right) \quad \text{B.3}$$

In the simplest use of the Monte Carlo method,  $f$  is either one (a success) or zero (a failure) depending on the value of  $\vec{x}$ .

In such a case, B.3 becomes

$$\frac{I}{H} \approx \frac{\text{successes}}{\text{attempts}}$$

$$\frac{\sigma(I)}{H} \approx \frac{1}{\text{attempts}} \sqrt{\frac{\text{successes} \times \text{failures}}{\text{attempts}-1}} \quad \text{B.4}$$

It remains to find a way of picking  $\vec{x}_i$  according to a given density function  $G/H$ .

Define

$$G_0(x_1, \dots, x_n) = G(\vec{X})$$

$$G_1(x_2, \dots, x_n) = \int_{-\infty}^{\infty} dx_1 G_0$$

⋮

$$G_n = H$$

$$M_1(\vec{X}) = \int_{-\infty}^{x_1} \frac{G_0(t, x_2, \dots, x_n)}{G_1(x_2, \dots, x_n)} dt$$

B.5

$$M_2(\vec{X}) = \int_{-\infty}^{x_2} \frac{G_1(t, x_3, \dots, x_n)}{G_2(x_3, \dots, x_n)} dt$$

⋮

$$M_n(\vec{X}) = \int_{-\infty}^{x_n} \frac{G_{n-1}(t)}{G_n} dt$$

The  $M_i$  are all between zero and one, and the coordinate transformation  $\vec{X} \leftrightarrow \vec{M}$  has a Jacobian  $|\frac{\partial \vec{M}}{\partial \vec{X}}| = G/H$ . If we pick  $n$  random variables  $M_i$  independently and uniformly from the interval  $(0,1)$ , then the corresponding  $\vec{X}$  will have density function  $G/H$ .

## APPENDIX C Optimizing Resolution

We wish to determine a function,  $f(x)$ , using an experiment whose  $i$ 'th setting results in a yield,  $D_i$ . The yields have expectation values

$$\bar{D}_i = \int dx f(x) g_i(x) \quad \text{C.1}$$

with known efficiency functions  $g_i(x)$ .

Suppose  $g_i(x)$  is large over only a small range of  $x$  (centered about some value,  $\bar{x}_i$ ). Then we can approximately take  $f$  to be a constant over that small range and get

$$f(\bar{x}_i) \approx \frac{\bar{D}_i}{\int g_i(x) dx} \quad \text{C.2}$$

The smaller the range over which  $g_i$  is large, the better the "resolution" at the experimental setting, and the more accurate is Equation C.2. If the resolution is good enough, the approximation

$$f(\bar{x}_i) \approx \frac{D_i}{\int g_i(x) dx} \quad \text{C.3}$$

is almost correct to within the experimental error in the measurement of  $D_i$ . Even if the resolution is poor, we may be able to use approximation C.3 if we define better efficiencies by linear combinations of the original efficiencies.

Let

$$h(x) = \sum a_i g_i(x) \equiv \vec{a} \cdot \vec{g}(x), \quad \text{C.4}$$

where  $\vec{a}$  is a set of as yet undetermined numbers. Then

$$\bar{D}_h = \int dx f(x) h(x) = \sum a_i \bar{D}_i = \vec{a} \cdot \vec{D} \quad \text{C.5}$$

Define a function with the properties of a probability density:

$$\rho(x) = \frac{h^2(x)}{\int h^2(x)} \quad \text{C.6}$$

Then in terms of that probability density, the variance of  $x$  is

$$\sigma^2 = \int x^2 \rho(x) - (\int x \rho(x))^2 \quad \text{C.7}$$

We choose to use  $\sigma^2$  as a measure of the resolution. That is, the smaller the value of  $\sigma^2$ , the better the resolution. The reader should note that if  $h(x)$  were used in place of  $h^2(x)$  in Equation C.6,  $\rho(x)$  could not be considered a probability density when  $h(x)$  goes negative. If in place of  $h^2(x)$  we used the absolute value of  $h(x)$ , the author would not know how to optimize the resolution.

Using Equations C.6 and C.7, we will show how to minimize  $\sigma^2$  with respect to the  $a_i$ .

Let

$$P_{ij} = \int g_i(x) g_j(x) \quad \text{C.8a}$$

$$Q_{ij} = \int x g_i(x) g_j(x) \quad \text{C.8b}$$

$$R_{ij} = \int x^2 g_i(x) g_j(x) \quad \text{C.8c}$$

$$\lambda_1 = \frac{a^T R a}{a^T P a} \quad \text{C.8d}$$

$$\lambda_2 = \frac{a^T Q a}{a^T P a} \quad \text{C.8e}$$

where  $a^T$  is the transpose of  $a$  (we consider  $a$  to be a column vector and  $a^T$  to be a row vector). Then

$$\sigma^2 = \lambda_1 - \lambda_2^2$$

and minimization of  $\sigma^2$  with respect to  $\vec{a}$  gives the equation

$$(R + (2\lambda_2^2 - \lambda_1) P - 2\lambda_2 Q) a = 0. \quad \text{C.9}$$

To facilitate the solution of Equation C.9, we use a notation that is familiar from elementary quantum mechanics by defining

$$\begin{aligned} |\Psi\rangle &= \Psi(x) \\ \langle\theta|\Psi\rangle &= \int dx \theta^*(x) \Psi(x) \end{aligned}$$

where the complex conjugate of  $\theta(x)$ ,  $\theta^*(x)$ , is equal to  $\theta(x)$  for all functions to be considered here. In this notation

$$\bar{D}_h = \langle f|h \rangle$$

$$Q_{ij} = \langle i|x|j \rangle$$

$$\lambda_1 = \frac{\langle h|x^2|h \rangle}{\langle h|h \rangle}, \quad \text{etc.}$$

Suppose the functions  $g_i(x)$  formed a complete set (i.e., suppose all functions could be expressed as linear combinations of the  $g_i(x)$ ). Then we could find eigenfunctions of the operator,  $x$ . Those eigenfunctions would correspond to eigenvectors of matrix  $Q$ ,  $R$  would equal  $Q^2$ , and  $\sigma^2$  would be minimized to zero. Because the  $g_i$  do not form a complete set, all these statements are only approximations to the truth. For example, to see why we expect  $R$  to be approximately  $Q^2$  for an approximately complete set of functions, transform to an orthonormal basis (one for which  $P$  is the identity matrix) and use

$$\sum_i |i\rangle\langle i| \approx 1$$

(with equality only for a complete orthonormal set of  $|i\rangle$ ).

$$\begin{aligned} R_{ij} &= \langle i|x^2|j \rangle \approx \langle i|x(\sum_k |k\rangle\langle k|)x|j \rangle = \\ &= \sum_k Q_{ik}Q_{kj} = (Q^2)_{ij} . \end{aligned}$$

If a problem in quantum mechanics can be solved approximately, then perturbation theory makes it possible to get more exact solutions. We are therefore motivated to find the eigenfunctions of  $Q$ .

Let us begin the solution of Equation C.9 by transforming to an orthonormal basis of functions,  $|i\rangle$ . In other words, we define a new experiment in which the new data and  $g_i(x)$  are linear combinations of the original data and the original  $g_i$ , and such that matrix P is the identity. There are several ways of finding such linear combinations. Later we will make an especially judicious choice, but for now we assume the transformation has been made.

Computer programs are available for the diagonalization of any reasonably small real symmetric matrix Q. The eigenfunctions of Q are of the form

$$|\theta_i\rangle = \sum_j B_{ji} |j\rangle \quad \text{C.10}$$

where  $B^T Q B$  is diagonal and B is an orthogonal matrix. In this new basis of eigenfunctions,  $|\theta_i\rangle$ , P is still the identity. By the above arguments concerning approximate completeness, the eigenfunctions of Q are expected to be approximate solutions to Equation C.9. We wish to find the eigenfunctions,

$$|\Psi_i\rangle = |\theta_i\rangle + \text{small correction,}$$

of  $R - 2\lambda_2 Q$ .

In the basis of the states  $|\theta_i\rangle$ , define matrices H,  $H^0$ , and  $H'$  by

$$R - 2\lambda_2(i) Q = H = H^0 + H' \quad ,$$

where  $H^0$  is diagonal and  $H'$  has only off-diagonal elements. Then because

$$\lambda_2(i) = \frac{\langle \Psi_i | x | \Psi_i \rangle}{\langle \Psi_i | \Psi_i \rangle} = \frac{\langle \theta_i | x | \theta_i \rangle}{\langle \theta_i | \theta_i \rangle} + \text{second-order correction,}$$

we can use second-order perturbation theory: (48)

$$\begin{aligned} |\Psi_i\rangle &= |\theta_i\rangle + \sum_{j \neq i} \frac{H'_{ji}}{H_{ii} - H_{jj}} |\theta_j\rangle + \sum_{j, k \neq i} \frac{H'_{jk} H'_{ki}}{(H_{ii} - H_{jj})(H_{ii} - H_{kk})} |\theta_j\rangle = \\ &= \sum_j C_{ji} |\theta_j\rangle \end{aligned} \quad \text{C.11}$$

To second order,

$$|\Psi_i\rangle = \sum_j (BC)_{ji} |\theta_j\rangle .$$

If higher order corrections are desired, it is necessary to first correct  $\lambda_2(i)$ .

We have completed the description of our method of maximizing resolution. But the above described method is much more useful when we make a modification, which we motivate as follows:

Starting from our original functions,  $g_i(x)$ , we restrict ourselves to transformations,  $a$ , that are normalized to unity. If  $\sigma_i$  is the standard deviation of  $D_i$  and if the errors of the different  $D_i$  are independent, then  $\sigma_h^2 = (\text{variance of } D_h) = \sum a_1^2 \sigma_i^2 \geq (\text{minimum of the } \sigma_i^2) \equiv \sigma_m^2$ . But

$$\overline{D}_h^2 = (f h f)^2 \leq f^2 h^2 \quad .$$

So

$$\frac{\sigma_h^2}{\overline{D}_h^2} \geq \frac{\sigma_m^2}{f^2 h^2} \quad \text{C.12}$$

$f h^2$  may be so small that the above inequality forces  $\sigma_h^2 / \overline{D}_h^2$  to be of order one. In such a case, the corresponding linear combination of the data contains almost no experimental information, and that combination can be neglected with little loss to the accuracy of the determination of  $f$ .

Express the transformation,  $\vec{a}$ , as

$$\vec{a} = \sum_j b_j \vec{a}(j)$$

where  $\vec{a}(j)$  is the  $j$ 'th eigenvector of  $P$  and has eigenvalue  $\lambda_j$ . Then the requirement that  $\vec{a}$  be normalized to unity is equivalent to the requirement that  $\vec{b}$  be normalized to unity, and

$$f h^2 = \vec{a}^T P \vec{a} = \sum_j b_j^2 \lambda_j \quad .$$

It is clear, then that the smallest values of  $f h^2$  with  $\vec{a}$  normalized to unity will all be linear combinations of those eigenvectors of  $P$  with smallest eigenvalues. To make the method described in this appendix more useful, one should exclude from consideration those transformed settings corresponding to especially small eigenvalues of  $P$ . The smaller the size of the minimum eigenvalue allowed, the

better will be the resolution ultimately attained, but the worse will be the statistical errors on the  $f(\bar{x})$  obtained through Equation C.3. Too many eigenvalues may lead to statistical errors so large that the results are useless.

Once  $P$  is diagonalized and the less important linear combinations are excluded, the remaining submatrix of  $P$  can be transformed into the identity matrix by renormalizing the functions. These renormalized linear combinations are the functions  $|j\rangle$  to which we refer in Equation C.10.

As a final point, it should be noted that if  $s(x) = f(x)/r(x)$  for known  $r(x)$  is expected to behave more smoothly than  $f(x)$ , then Equation C.3 is more reliable for  $s(\bar{x})$  than  $f(\bar{x})$ . In such a case it is best to use the method of this appendix on  $s(x)$  and at the end evaluate  $f(\bar{x}) = r(\bar{x})s(\bar{x})$ .

## APPENDIX D Inside the Black Box

Suppose we wish to measure a function that can be expressed in the form

$$f(x,y) = \sum_{J=1}^{MU} r_J(y) f_J(x) \quad D.1$$

where the  $r_J(y)$  are an orthonormal set of functions, and  $MU$  is some finite integer. For example,  $y$  can be  $\cos \theta$  with  $r_J(y) = \sqrt{\frac{2J-1}{2}} P_J(\cos \theta)$ . Another special case is  $MU = 1$ , with  $r_1(y) = 1$ , in which case we really are talking about determining a function of  $x$  alone. In general,  $y$  can be a set of variables.

An experiment is performed at NSETS settings and data are obtained in the form of yields, with the yield at setting  $i$  equal to  $D_i$ . The efficiency function for setting  $i$ ,  $g_i(x,y)$ , is defined so that the expected value of the yield is

$$\bar{D}_i = \iint dx dy f(x,y) g_i(x,y) .$$

We assume the efficiency functions are zero outside the range  $x_0 < x < x_1$ . Define  $g_{iJ}(x)$  so that

$$g_i(x,y) = \sum_{J=1}^{\infty} r_J(y) g_{iJ}(x) .$$

Then

$$\bar{D}_i = \int_{X_0}^{X_1} dx \sum_{J=1}^{MU} g_{iJ}(x) f_J(x) .$$

Define the matrix of covariance to be  $\text{COV}_{ij}$  = expectation value of  $(D_i - \bar{D}_i)(D_j - \bar{D}_j)$ , so that chi squared =  $\chi^2 = (D - \bar{D})^T \text{COV}^{-1} (D - \bar{D})$  (superscript "T" means "transpose").

For some integer N we define a measure of smoothness of f:

$$Y = \int dx dy \left[ \frac{\partial^N f(x,y)}{\partial x^N} \right]^2 = \int_{X_0}^{X_1} dx \Sigma [f_J^{(N)}(x)]^2 \quad \text{D.2}$$

If f is now assumed to be the smoothest function such that  $\chi^2$  is fixed to be  $\chi_0^2$ , we can find f by minimizing the quantity  $\chi^2 + vY$  with respect to variations of f with fixed Lagrange multiplier, v. v can be chosen such that  $\chi^2 = \chi_0^2$  at the minimum. To carry out the minimization, first we find the minimum with  $\bar{D}$  fixed, then we vary  $\bar{D}$  so as to get an absolute minimum. But if  $\bar{D}$  is fixed, minimizing  $\chi^2 + vY$  is the same thing as minimizing Y.

From the calculus of variations with Lagrange multipliers  $\lambda_i$  (one for each constraint  $\bar{D}_i$ ) we get

$$\begin{aligned} \Sigma_J \left\{ [f_J^{(N)} \delta f_J^{(N-1)} - f_J^{(N+1)} \delta f_J^{(N-2)} \dots + (-1)^{N-1} f_J^{(2N-1)} \delta f_J] \right\}_{X_0}^{X_1} + \\ + (-1)^N \int_{X_0}^{X_1} dx [f_J^{(2N)} - \Sigma_{i=1}^{NSETS} \lambda_i g_{ij} \delta f_J] \Big\} = 0 \quad \text{D.3} \end{aligned}$$

where our notation is such that  $\delta f^{(n-1)}$  = arbitrary infinitesimal variation of the (N-1) derivative of f, etc.

Since the  $f_J$  are varied independently for each J, the individual terms of the sum over J are independently zero. Since

Equation D.3 holds for arbitrary variations, in particular it holds for variations that leave the boundary values fixed. Then we must have

$$f_J^{(2N)} = \sum_{i=1}^{NSETS} \lambda_i g_{iJ} = \lambda \cdot g_J \quad D.4$$

In order to satisfy D.3 with variations taken at the boundaries, at  $x_0$  and  $x_1$  we must have

$$0 = f_J^{(N)} = f_J^{(N+1)} = \dots = f_J^{(2N-1)} \quad D.5$$

Now define

$$G_i(0, J, x) = g_{iJ}(x)$$

$$G_i(L+1, J, x) = \begin{cases} \int_{x_0}^x G_i(L, J, t) dt & 0 \leq L < N \\ \int_{x_1}^x G_i(L, J, t) dt & N \leq L < 2N \end{cases}$$

In practice, these integrals can be obtained by first approximating the efficiency functions by piecewise quadratic functions, then performing all integrals exactly on the approximate functions. This method is an obvious generalization of Simpson's rule.

In order to use notation consistent with that of the Fortran program that was written to implement the mathematics described here, we define

$$\mathbb{T}\mathbb{T}_i(J, x) = G_i(N, J, x)$$

$$\mathbb{T}\mathbb{L}_i(J, x) = G_i(2N, J, x)$$

$$\mathbb{S}\mathbb{T}(i, J, k) = G_i(k, J, X_1) \quad .$$

Integrating Equation D.4  $N$  times and inserting boundary conditions, D.5, at  $X_0$  gives

$$f_J^{(N)}(x) = \lambda \cdot \mathbb{T}\mathbb{T}(J, x)$$

so that

$$Y = \lambda^T S \lambda, \text{ where } S_{ij} = \int_{X_0}^{X_1} dx \sum_{J=1}^{MU} \mathbb{T}\mathbb{T}_i(J, X) \mathbb{T}\mathbb{T}_j(j, X) \text{ can be}$$

computed by Simpson's rule.

Integrating  $N$  more times,

$$f_J(x) = \lambda \cdot \mathbb{T}\mathbb{L}(J, x) + \sum_{L=1}^N \beta(NU) (x - X_0)^{L-1} \quad \text{D.6}$$

where  $NU = J + (L-1) \cdot MU$  and where the  $\beta(NU)$  are constants chosen to satisfy the boundary conditions at  $X_1$ . Then by repeated application of integration by parts it can be shown that

$$\bar{D} = (-1)^N S \cdot \lambda + \sum_{NU=1}^{NUMAX} \beta(NU) P(NU) \quad \text{D.7}$$

where  $NUMAX = N \cdot MU$ , and we define

$$P_i^{(J+(L-1) \cdot MU)} = P_i^{(NU)} = \int_{X_0}^{X_1} (x-X_0)^{L-1} g_{iJ}(x) =$$

$$\sum_{k=1}^L (-1)^{k-1} (X_1 - X_0)^{L-K} \frac{(L-1)!}{(L-k)!} ST(i, J, k) .$$

Notice that "S" is a real, symmetric matrix, hence is diagonalizable. That is, there exists a matrix, A, such that  $A^T =$  transpose of A is also its inverse, and  $A^T \cdot S \cdot A$  is diagonal. Suppose we define  $D' = A^T \cdot D$ , and similarly redefine the efficiencies. Then  $COV' = A^T \cdot COV \cdot A$  and  $S' = A^T \cdot S \cdot A$ , so  $S'$  is diagonal. But the diagonal elements of  $S'$  are integrals of the form

$$S'_{ii} = \int \sum [TT'_{ii}]^2$$

so the eigenvalues are non-negative. The only way an eigenvalue can be zero (or almost zero) is if the corresponding  $TT'$  is zero (almost), i.e., if one efficiency function is a linear combination of others (almost). Assume for now that this doesn't happen. Then we can define

$$C_i = \sqrt{i^{\text{th}} \text{ eigenvalue}}$$

and we can define

$$D'_i = \frac{1}{C_i} \sum_{ij} A_{ji} D_j \quad D.8$$

(similarly for the efficiencies and covariance matrix). Hereafter, when we speak of  $D, \bar{D}, COV, P$ , etc., we assume this linear transformation has already been made. Then  $S$  is the identity matrix,  $I$ .

Solving D.7 for  $\lambda$  we get

$$\lambda = (-1)^N [\bar{D} - \sum_{NU} \beta(NU) P(NU)] \quad . \quad \text{D.9}$$

From the boundary conditions D.5, evaluated at  $X_{\perp}$ , we get  
for  $K = 1$  to  $N$

$$0 = \lambda \cdot ST(J,K) \quad .$$

Let

$$R_i(J + (L-1) \cdot MU) = R_i(NU) = ST(i, J, L); \quad B(NU, NU') = R(NU) \cdot P(NU') \quad .$$

Then

$$0 = \lambda \cdot ST$$

implies that

$$\beta(NU) = \sum_{NU'} B^{-1}(NU, NU') \bar{D} \cdot R(NU') = \bar{D} \cdot \text{TAU}(NU) \quad \text{D.10}$$

where

$$\text{TAU}_i(NU) = \sum_{NU'=1}^{\text{NUMAX}} B^{-1}(NU, NU') R_i(NU') \quad .$$

Define

$$\tilde{Q}_{ij} = \sum_{NU} \text{TAU}_i(NU) P_j(NU); \quad Q = I - \tilde{Q} \quad .$$

Then (using D.9 and D.10) so that D.6 becomes

$$\lambda = (-1)^N Q^T \bar{D} \quad ,$$

$$f_J(X) = \lambda \cdot T(J, X) + \sum_{L=1}^N \beta(NU)(X-X_0)^{L-1} = \bar{D} \cdot T(J, X);$$

$$T = (-1)^n Q \cdot T1 + \sum_{L=1}^N \text{TAU}(J+(L-1) \cdot \text{MU})(X-X_0)^{L-1} . \quad \text{D.11}$$

From the definition of TAU,

$$\sum_{NU'} B(NU, NU') \text{TAU}(NU') = R(NU) \Rightarrow \tilde{Q} \cdot R(NU) = R(NU) \Rightarrow \tilde{Q} \cdot P(NU) = P(NU) \Rightarrow$$

$$\Rightarrow \tilde{Q} \cdot \tilde{Q}^T = \tilde{Q}^T \Rightarrow \tilde{Q} \text{ and } Q \text{ are symmetric and } Q^2 = Q \text{ so that}$$

$$Y = \lambda \cdot \lambda = \bar{D}^T \cdot Q \cdot \bar{D} .$$

We have minimized  $X^2 + vY$  for fixed  $\bar{D}$ . It is easy, now, to perform the minimization with respect to  $\bar{D}$  of

$$X^2 + vY = (D-\bar{D})^T \cdot \text{COV}^{-1} \cdot (D-\bar{D}) + v \bar{D}^T Q \bar{D} .$$

The result is

$$\bar{D} = (1 + v \cdot \text{COV} \cdot Q)^{-1} D$$

$$X^2 = v^2 \bar{D}^T Q \cdot \text{COV} \cdot Q \bar{D} = v \bar{D}^T Q (D-\bar{D}) \quad \text{D.12}$$

$$\frac{dX^2}{dv} = 2 W \cdot \bar{D} ; \quad W = Q(1 + v \cdot \text{COV} \cdot Q)^{-1} (D-\bar{D})$$

To find the  $v$  appropriate to a given value of  $X_0^2$ , i.e., to find  $v(X_0^2)$ , we iterate using Newton's method:

$$v(X_0^2) \approx v(X^2) + \frac{X_0^2 - X^2}{dX^2/dv} .$$

In order to use this iterative procedure successfully, it is necessary to find a good initial choice of  $v$ . To do this we try to find a function of  $v$  that resembles  $\chi^2(v)$ , but can be solved for  $v(\chi^2)$ . From the easily seen fact that both  $Q$  and  $COV$  have all positive eigenvalues, it can be shown that for non-negative  $v$ , both  $\chi^2$  and  $\frac{d\chi^2}{dv}$  are non-negative. A little thought convinces one that  $\chi^2(\infty)$  corresponds to  $\chi^2$  for the best fit with a  $N-1$  degree polynomial, a fit that is easy to find. Finally,  $\chi^2(v) \propto v^2$  for small  $v$ . Then a function with the above described properties of  $\chi^2$  is

$$\chi^2(v) \approx \frac{\alpha v^2}{1 + \beta v + \gamma v^2}$$

where

$$\alpha = D^T Q \cdot COV \cdot Q D, \quad \alpha/\gamma = \chi^2(\infty),$$

and  $\beta$  is chosen to give the correct value of  $\frac{d^3\chi^2}{dv^3}$  at  $v = 0$ :

$$\beta = \frac{2}{\alpha} D^T \cdot Q \cdot COV \cdot Q \cdot COV \cdot Q \cdot D \quad .$$

With this approximate relationship, we get a reasonably good initial approximation to  $v(\chi^2)$ ; then we can apply Newton's method.

There are statistical fluctuations in our answer. If we keep  $\chi^2$  fixed and vary the input data,  $D$ , our answer varies. The amount by which the answer tends to vary is characterized by the square of the standard deviation of  $f_J(x)$ :

$$\langle (\delta f_J(x))^2 \rangle = \sum_{ij} \langle (\delta \bar{D}_i T_i) (\delta \bar{D}_j T_j) \rangle = T^T M T$$

where  $M_{ij} = \langle \delta \bar{D}_i \delta \bar{D}_j \rangle$ . If  $\delta \bar{D} = SS(\delta D)$ , then  $M = SS \text{ COV } SS^T$ .

We wish to find  $SS_{ij} \approx \frac{d\bar{D}_i}{d\bar{D}_j}$ , where  $v = v(D)$  is such that  $\chi^2$  is kept

fixed as  $D$  is varied.

Using

$$\frac{d\bar{D}}{dD} = \left. \frac{\partial \bar{D}}{\partial D} \right|_{v \text{ fixed}} + \left. \frac{\partial \bar{D}}{\partial v} \right|_{D \text{ fixed}} \frac{dv}{dD} \text{ and } 0 = \frac{d\chi^2}{dD} = \frac{\partial \chi^2}{\partial D} + \frac{\partial \chi^2}{\partial v} \frac{dv}{dD}$$

we get after some algebra

$$SS_{ij} = (1+v \text{ COV.Q})_{ij}^{-1} + \frac{U_i W_j}{\bar{D} \cdot W}; \quad U = (1+v \text{ COV.Q})^{-1} (\bar{D}-D).$$

We have completed the description of the method for finding the smoothest function for a given  $\chi^2$ . Now we return to a point glossed over earlier in this appendix.

The above described method does not depend crucially on making a linear transformation that sends  $S$  into the identity. It might seem that the transformation is made in order to simplify the algebra later on, but in fact there is a more important motivation, with the simplification coming as a bonus. The linear transformation is intended to deal with two difficulties of the method.

The first problem involves those linear combinations of efficiency functions that are essentially zero (i.e., combinations  $g$  such that  $\bar{D}_i = \sum f_J g_{iJ} \approx 0$  for all smooth  $f$ ). Because random

errors prevent  $D_i$  from being zero, there is no smooth function such that  $\bar{D} = D$  (such that  $\chi^2 = 0$ ). Since  $f_J(x) = \bar{D} \cdot T(J, x)$ , the non-existence of a smooth solution with  $\chi^2 = 0$  means that  $T$  must be a very jagged function, and is hard to calculate and manipulate well.

The second problem concerns the limitation imposed by this method on the number of experimental settings. We must, for example, invert the matrix  $(1 + v \text{ COV } Q)$ , and if there are NSETS settings, then the matrix is dimensioned NSETS X NSETS. We are limited to around 50 settings.

The first problem is partially cured by making the above described linear transformation, for then certain linear combinations of the input data can be very large in magnitude (notice the  $1/C_i$  in the definition D.8) so that  $D T$  can be jagged with  $T$  reasonably smooth. More important, diagonalization of  $S$  tells us which linear combinations of efficiency functions are especially small, so that the corresponding combinations of the yields will be theoretically about zero. Since we know that the experimental deviations from zero are, for these combinations, almost entirely caused by statistical errors, these combinations contain no information about  $f$ , and can therefore be ignored.

When there are too many settings for available computers to handle via Black Box, we can use the above method for deciding which linear combinations of the original, non-transformed data can be most safely ignored. Suppose  $h_J(x) = \vec{a} \cdot \vec{g}_J(x)$  for some vector,  $\vec{a}$ , and suppose  $D_h$  is the corresponding linear combination of the data. Call  $\sigma_h^2$  the variance of  $D_h$ . Then with  $\text{COV}_{ij} =$  the original, non-transformed covariance matrix, the relation

$$\frac{\sigma_h^2}{\bar{D}_h^2} \geq \frac{\mathbf{a}^T \text{COV} \mathbf{a}}{(\sum_J f_J^2)(\sum_J h_J^2)}$$

allows one to decide semi-quantitatively how much is lost by neglecting certain linear combinations (see equation C.12 and the associated reasoning).

## APPENDIX E Maximum Likelihood Method

We used the maximum likelihood method to find the number of etas in each run (see Section III C). Associated with this method is a systematic error (a "bias") for which we would like to correct. The purpose of this appendix is primarily to discuss this error and to derive a correction to the maximum likelihood method. On the way to the derivation, we will also obtain a formula for the random error of the method.

Suppose we perform an experiment = a set of  $N$  measurements, with the  $i$ 'th measurement giving result  $Z_i$ .  $Z_i$  may be a number, a set of numbers, or even a non-numerical result such as "true," or "false," or "she became pregnant." Assume the measurements are independent and assume that the probability (or probability density) that the  $i$ 'th measurement gives  $Z$  is  $P_i(Z; \vec{x})$ . The " $\vec{x}$ " consists of  $M$  real, continuous valued parameters which are to be varied to give a good fit to the experiment.

We have in mind two cases, though our results are more general. Case I: We assume the  $P_i$  are all the same function, and we let " $n$ " be the number of measurements.

Case II: The experiment is performed by taking a sequence of events and binning them according to some property. At the end of the experiment the set of measurements consists of the numbers of events in each bin. We assume that the probability of an event entering a given bin in a given instant during the experiment is independent of

the past history of the experiment. We also assume that there is some parameter,  $n$ , (characterizing the amount of statistics gathered) such that the expectation value of the number of events in a bin is of the form

$$\langle Z_i \rangle = nq_i(\vec{x}),$$

where  $q_i$  does not depend on " $n$ ". Then  $Z_i$  is Poisson distributed:

$$P_i(Z; \vec{x}) = \frac{(nq_i)^Z}{Z!} e^{-nq_i}; \quad Z = \text{integer} \geq 0.$$

Although " $n$ " is an integer for Case I, it need not be one for Case II.

For Case I, if the result of a measurement can only take on a finite number of values, then Case I is identical to Case II. If, on the other hand,  $Z$  can take on a continuum of values, then we can break up the range of  $Z$  into a large number of small regions and bin the measurements. Bin number " $i$ " is centered about  $Z_i$  and has volume  $dZ_i$ . We can define

$$q_i = dZ_i P(Z_i; \vec{x}) .$$

Instead of treating  $Z_i$  as the result of a measurement,  $Z_i$  is considered to be the property describing a bin, and the result of a measurement is the number of events in a bin. In the limit of infinitesimal bin size, Case I is then a specialization of Case II.

Suppose we define

$$\chi^2 = \sum_{k=1}^N \frac{(Z_k - nq_k)^2}{nq_k},$$

where  $N$  is the number of bins. Then it is straightforward to show that if  $x$  is fixed and known, then  $\chi^2$  will satisfy

$$\langle \chi^2/N \rangle = 1$$

$$\sigma^2\left(\frac{\chi^2}{N}\right) = \frac{2}{N} + \frac{1}{N^2} \sum_{k=1}^N \frac{1}{nq_k}$$

(Cramer<sup>(51)</sup> proves a similar result subject to the constraint that

$\sum_{k=1}^N Z_k = \text{constant}$ ). It is clear, then, that if  $nq_k$  is very small

for some  $k$ ,  $\chi^2/N$  will have a large variance and will therefore be a poor measure of the goodness of fit. One way of avoiding this problem is to combine bins (as recommended by Cramer). More simply, one may use the restricted sum

$$\chi_r^2/N_r = \frac{1}{N_r} \sum_{k=1}^{N_r} \frac{(Z_k - nq_k)^2}{nq_k},$$

where the sum runs over all  $k$  for which  $nq_k$  is larger than some minimum value. By seeing what happens to  $\sigma^2\left(\frac{\chi_r^2}{N_r}\right)$  when another bin is included in the sum, we find that  $\sigma^2$  will decrease if the new bin has  $nq \geq \frac{1}{2}$ . There is another complication.  $\vec{x}$  is not known a priori, but is fitted by the data. In that case, it is well known<sup>(51)</sup> that in the limit of very large  $n$ ,  $\chi^2/N_d$  can be expected to have mean

1 and standard deviation  $2/N_d$ , where  $N_d$  is the number of degrees of freedom. Although we haven't proved it, we expect that if there are parameters to be fitted, if there is finite  $n$ , and if the sums are restricted to  $nq_k \geq \frac{1}{2}$ , then

$$\langle \chi_r^2/N_d \rangle \approx 1,$$

$$\sigma^2(\chi_r^2/N_d) \approx 2/N_d + \frac{1}{N_d^2} \sum_{k=1}^{N_r} \frac{1}{nq_k}$$

with  $N_d = N_r - M$  ( $M = \text{dimension of } \vec{x}$ ). In our data analysis we use this relation to see how good our fits really are (see Section III C).

We now return to a more general discussion. For a particular set of parameters,  $\vec{x}$ , and set of experimental results,  $\vec{Z}$ , we define  $L(\vec{Z};\vec{x})$  to be the probability that the outcome of the set of experiments will be  $\vec{Z}$  given that the parameters were  $\vec{x}$ . Suppose that before the experiment is performed, one has an a priori opinion about what  $\vec{x}$  can be, and suppose the opinion is expressed as a probability distribution of  $\vec{x}$ ,  $Q(\vec{x})$ . Then the a priori probability that the parameters are  $\vec{x}$  and the experiment will give result  $\vec{Z}$  is  $L(\vec{Z};\vec{x})Q(\vec{x})$ . Summing this formula over all  $\vec{x}$  gives the a priori probability distribution of  $\vec{Z}$ ,  $R(\vec{Z})$ . Let  $W(\vec{x};\vec{Z})$  be the probability that the set of parameters is  $\vec{x}$ , given that we know the experimental result to be  $\vec{Z}$ :

$$W(\vec{x};\vec{Z}) = \frac{L(\vec{Z};\vec{x})Q(\vec{x})}{R(\vec{Z})} .$$

Given the result,  $\vec{Z}$ , there is then a most probable value for  $\vec{x}$ . This

value is found by maximizing  $W$ , which can be done by maximizing  $\text{Log}(W)$ , or equivalently by setting for each parameter

$$0 = \frac{1}{n} \frac{\partial \log(W)}{\partial x_j} = \frac{1}{n} \frac{\partial \text{Log}(LR)}{\partial x_j}$$

" $n$ " characterizes the amount of experimental data gathered, and the factor " $1/n$ " is included for convenience later on. For the previously discussed independent  $P_i(Z_i; \vec{x})$ ,

$$LQ = Q(\vec{x}) \prod_{i=1}^N P_i(Z_i; \vec{x}) .$$

In its conventional form, the maximum likelihood method assumes that  $Q(\vec{x})$  varies slowly with  $\vec{x}$ , so that

$$0 = \frac{1}{n} \sum_i \frac{\partial \text{Log}(P_i)}{\partial x_j} \tag{E.1}$$

If  $Q(\vec{x})$  is rapidly varying, Equation E.1 can still be used, provided  $Q$  is considered to be one of the  $P_i$ 's (say,  $Q(\vec{x}) = P_{N+1}(Z_{N+1}; \vec{x})$ ).

If  $\vec{x}_0$  is the true (though unknown) value of  $\vec{x}$ , then we can define

$$B_{j_1 j_2 \dots j_\ell}^{(\ell)} = \frac{1}{n} \sum_i \frac{\partial^\ell \text{Log}(P_i)}{\partial x_{j_1} \dots \partial x_{j_\ell}} \Big|_{\vec{x} = \vec{x}_0}$$

and expand Equation E.1 about  $x_0$ .

$$0 = B_j^{(1)} + \sum_k y_k B_{kj}^{(2)} + \frac{1}{2} \sum_k y_k y_l B_{k\ell j}^{(3)} + \dots, \quad \text{E.2}$$

where  $y_k = x_k - x_0$  and  $x_k$  is the maximum likelihood solution. This simple trick was used by Cramer<sup>(51)</sup> in a proof that the maximum likelihood estimator is asymptotically unbiased and efficient (which means roughly that when the statistics are good, there is no method for finding parameters that is more accurate).

Define

$$C_{j_1 \dots j_\ell}^{(\ell)} = - \langle B_{j_1 \dots j_\ell}^{(\ell)} \rangle.$$

One can easily show that for any differentiable function  $f(\vec{Z}; \vec{x}_0)$  we have

$$\frac{\partial}{\partial x_{j_0}} \langle f \rangle = \left\langle \frac{\partial f}{\partial x_{j_0}} \right\rangle + n \langle f B_j^{(1)} \rangle.$$

It follows (taking  $f=1$ ) that  $C_j^{(1)} = 0$ . Taking  $f = B_{j_1 \dots j_\ell}^{(\ell)}$  gives

$$C_{j, j_1, \dots, j_\ell}^{(\ell+1)} = \frac{\partial}{\partial x_{0j}} C_{j_1 \dots j_\ell}^{(\ell)} + n \text{cov}(B_j^{(1)}, B_{j_1 \dots j_\ell}^{(\ell)})$$

E.3

where we use the notation  $\text{cov}(r, s) = \text{"covariance of } r \text{ and } s\text{"} = \langle rs \rangle - \langle r \rangle \langle s \rangle$ .

For Case II,

$$B_{j_1 \dots j_\ell}^{(\ell)} = \sum_i \frac{Z_i - nq_i}{n} \left. \frac{\partial^\ell \text{Log}(q_i)}{\partial x_{j_1} \dots \partial x_{j_\ell}} \right|_{\vec{x}=\vec{x}_0} - C_{j_1 \dots j_\ell}^{(\ell)}$$

where  $C^{(\ell)}$  is independent of  $n$ . The mean of  $B^{(\ell)}$  is constant for all  $n$  and its standard deviation is a constant divided by the square root of  $n$ . Similar results on the mean and standard deviation are true for Case I. In what follows, we assume that  $B^{(\ell)}$  and  $\sqrt{n}$  (standard deviation of  $B^{(\ell)}$ ) both tend to constants in the limit of large  $n$ . We also assume that for large  $n$  the maximum likelihood solution for  $y_k$  is typically of order  $1/\sqrt{n}$ . For Case I, Cramer<sup>(51)</sup> gives a sufficient, reasonable condition for the validity of this last assumption. When we say " $\mathcal{O}(1/n^\alpha)$ " or "of order  $1/n^\alpha$ " we mean "goes to zero with increasing  $n$  at least as fast as  $1/n^\alpha$ ." Then if we take  $f = (B_{j_1 \dots j_\ell}^{(\ell)} - \langle B_{j_1 \dots j_\ell}^{(\ell)} \rangle) (B_{k_1 \dots k_m}^{(m)} - \langle B_{k_1 \dots k_m}^{(m)} \rangle)$  we find that

$$\langle B_j^{(1)} (B^{(\ell)} - \langle B^{(\ell)} \rangle) (B^{(m)} - \langle B^{(m)} \rangle) \rangle \cong \langle f B_j^{(1)} \rangle = \frac{1}{n} \left[ \frac{\partial}{\partial x_{0j}} \langle f \rangle - \left\langle \frac{\partial f}{\partial x_{0j}} \right\rangle \right]$$

is of order  $\frac{1}{n^2}$ . Furthermore, since Equation E.2 gives

$$y_k = \sum_j (C^{(2)})^{-1}_{kj} B_j^{(1)} + \mathcal{O}\left(\frac{1}{n}\right),$$

we have  $\langle y_j (B^{(\ell)} - \langle B^{(\ell)} \rangle) (B^{(m)} - \langle B^{(m)} \rangle) \rangle$ ,  $\langle y_j y_k (B^{(m)} - \langle B^{(m)} \rangle) \rangle$  and  $\langle y_j y_k y_\ell \rangle$  all of order  $\frac{1}{n^2}$ . The following results are valid only in the approximation that we can neglect terms of order  $1/n^2$ , and we use the above facts repeatedly in order to decide which terms are to be excluded. We have and will continue to make other unstated assumptions involving convergence of sums, exchanging order of differentiation, and probably several other things that only a mathematician would worry about. Since this is not a mathematical thesis, we will cross our fingers and plunge forward oblivious to such technicalities.

Take the covariance of Equation E.2 with  $B_{j_1 \dots j_\ell}^{(\ell)}$  :

$$0 = \text{cov} (B_j^{(1)}, B_{j_1 \dots j_\ell}^{(\ell)}) - \sum_k \left[ \text{cov}(y_k, B_{j_1 \dots j_\ell}^{(\ell)}) C_{kj}^{(2)} \right] + \theta\left(\frac{1}{n}\right)$$

Using E.3 we see that  $C_{jk}^{(2)} = n \text{cov} (B_j^{(1)}, B_k^{(1)})$  is non-singular unless one of the  $B_j^{(1)}$  is some linear combination of the others. In other words, the matrix  $C^{(2)}$  is singular only if there is some  $w$  (a linear combination of the  $x_j$ ) such that the likelihood is stationary for small variations of  $w$ . If this problem does not occur, we get

$$\text{cov}(y_k, B_{j_1 \dots j_\ell}^{(\ell)}) = \sum_j (C^{(2)})_{kj}^{-1} \text{cov}(B_j^{(1)}, B_{j_1 \dots j_\ell}^{(\ell)}) + \mathcal{O}\left(\frac{1}{n^2}\right)$$

E.4

If  $\ell = 1$ , Equation E.4 reduces (using Equation E.3) to

$$\text{cov} (y_k, B_{j_1}^{(1)}) = \frac{1}{n} \delta_{kj_1} + \mathcal{O}\left(\frac{1}{n^2}\right) .$$

If we take the covariance of Equation E.2 with  $y_\ell$  and use the above equation, we conclude that

$$\text{cov}(x_k, x_\ell) = \text{cov}(y_k, y_\ell) = \frac{1}{n} (C^{(2)})_{k\ell}^{-1} + \mathcal{O}\left(\frac{1}{n^2}\right) \quad \text{E.5}$$

This covariance is a measure of the random error of the maximum likelihood method.

Finally we take the expectation value of Equation E.2. With the help of Equations E.3, E.4, and E.5 we get

$$\langle y_k \rangle = \frac{1}{n} \sum_{j\ell i} (C^{(2)})_{kj}^{-1} (C^{(2)})_{li}^{-1} \left[ \frac{1}{2} C_{lij}^{(3)} - \frac{\partial}{\partial x_{i_0}} C_{j\ell}^{(2)} \right] + \mathcal{O}\left(\frac{1}{n^2}\right) \quad \text{E.6}$$

The above equation is an approximation to the systematic error of the maximum likelihood method. To correct for this error,  $\langle y_k \rangle$  should be subtracted from the maximum likelihood solution,  $x_k$ .  $\langle y_k \rangle$  is called the "bias" of the method. In Equations E.5 and E.6 it is necessary to know  $\vec{x}_0$  in order to compute the  $C^{(l)}$ . If one uses  $\vec{x}$  in place of the unknown  $\vec{x}_0$ , a random error of order  $n^{-3/2}$  is added to the bias and covariance on top of the systematic error that came from dropping terms of order  $\frac{1}{n^2}$ .

Specializing to Case II,

$$C_{ij}^{(2)} = \sum_m \frac{1}{q_m} \frac{\partial q_m}{\partial x_i} \frac{\partial q_m}{\partial x_j} \quad \text{E.7}$$

$$\langle y_k \rangle = -\frac{1}{2n} \sum_m \left\{ \frac{1}{q_m} \left[ \sum_{li} (C^{(2)})_{li}^{-1} \frac{\partial^2 q_m}{\partial x_{l_0} \partial x_{i_0}} \right] \left[ \sum_j (C^{(2)})_{kj}^{-1} \frac{\partial q_m}{\partial x_{j_0}} \right] \right\} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

For Case I, with  $Z_i$  a real number (or vector) taking on a continuum of values, we use the results for Case II with infinitely fine bins. Then sums over  $m$  become integrals and we get

$$C_{ij}^{(2)} = \int \frac{dZ}{P(Z; \vec{x}_0)} \frac{\partial P}{\partial x_{i_0}} \frac{\partial P}{\partial x_{j_0}}$$

E.8

$$\langle y_k \rangle = -\frac{1}{2n} \int \left\{ \frac{dZ}{P(Z, \vec{x}_0)} \left[ \sum_{\ell i} (C^{(2)})^{-1}_{\ell i} \frac{\partial^2 P}{\partial x_{\ell_0} \partial x_{i_0}} \right] \left[ \sum_j (C^{(2)})^{-1}_{kj} \frac{\partial P}{\partial x_{j_0}} \right] \right\}$$

Finally we obtain the bias of any function  $f(\vec{x})$ , and the covariance of any functions  $f(\vec{x})$  and  $g(\vec{x})$ . By expanding  $f$  about  $x_0$  it is easily seen that the bias of  $f$  is

$$\begin{aligned} \text{bias} = \langle f(\vec{x}) \rangle - f(\vec{x}_0) &= \sum_k \langle y_k \rangle \left. \frac{\partial f}{\partial x_k} \right|_{\vec{x}=\vec{x}_0} + \frac{1}{2} \sum_{kl} \text{cov}(x_k, x_l) \left. \frac{\partial^2 f}{\partial x_k \partial x_l} \right|_{\vec{x}=\vec{x}_0} + \\ &+ \mathcal{O}\left(\frac{1}{n^2}\right) . \end{aligned}$$

In the above equation we can replace  $f$  everywhere by  $g$  or by  $fg$ .

Then it is a matter of algebra to show that the covariance of  $f$  and  $g$  is

$$\text{cov}(f, g) = \langle fg \rangle - \langle f \rangle \langle g \rangle = \sum_{kl} \text{cov}(x_k, x_l) \left. \frac{\partial f}{\partial x_k} \frac{\partial g}{\partial x_l} \right|_{\vec{x}=\vec{x}_0} + \mathcal{O}\left(\frac{1}{n^2}\right) .$$

Example:

For an experiment that measures the polarization of a particle by scattering it in a substance, let  $x$  be the polarization,  $\theta$  be the angle of the scatter of a given event, and  $Z$  be parameters needed to

specify the analyzing power (such as the energy and inelasticity of of the scatter). Take  $Q(Z)$  = probability distribution of  $Z$ , and  $A(Z)$  = analyzing power. Then

$$P(\theta, Z; x) = \frac{1}{2\pi} (1 + xA(Z) \cos(\theta))Q(Z) .$$

Since  $P$  is linear in  $x$ , from Equation E.8 we see immediately that to lowest order the maximum likelihood method is unbiased for this example.

Also from E.8 we find that the standard deviation of a polarization measurement with  $n$  events is

$$\sigma_x \approx \frac{x}{\sqrt{n} \left\langle \frac{1}{\sqrt{1-x^2 A^2(Z)}} \right\rangle} \xrightarrow{x^2 A^2(Z) \text{ small}} \frac{1}{\sqrt{n} \langle A^2(Z) \rangle}$$

(compared with  $\frac{1}{\langle A(Z) \rangle} \sqrt{\frac{2}{n}}$  used in Reference 22).

## APPENDIX F Synchrotron End Point Correction

The synchrotron's beam energy meter measured the magnetic field in which the electrons circulated. For a given radius of circulating electrons, the end point energy was proportional to the magnetic field at the time the beam struck the radiator. The proportionality constant used by the operators for setting the synchrotron end point was inaccurate, with the calibration by Thiessen<sup>(3)</sup> giving a correction factor of  $1.021 \pm .003$ . This factor was determined by accurately measuring the integrated voltage output from a coil (of known size) as the synchrotron's magnetic field increased from zero to its end point value. The uncertainties in the result arise mainly from the uncertainty in the area of the coil and possibly from non-circularity in the shape of the orbit and drifts in the electronics.

An experiment using the same apparatus and methods as ours gave an independent measurement of the correction factor. The cross section for photoproduction of etas from protons rises very rapidly at threshold. Using the known location of the threshold, we were able to take advantage of this behavior to measure the true synchrotron end point at the eta threshold, and thereby get the correction factor. For this purpose we considered only those settings with nominal end point of 750 MeV or less, and we used Blackbox (see Appendix D) to fit the data with various assumed correction factors. We used only the six most informative linear combinations of the data, where our criterion for "informative" is given in Appendix D.

One of the results given by Blackbox was the value of  $\chi^2$  for the best straight line fit to the magnitude of the squared amplitude for photoproduction (for conventions relating amplitudes to cross sections, see Appendix G). The lowest value of  $\chi^2$  for the straight line fit occurred at a correction factor of  $1.0210 \pm .0005$ . This factor is the one used in our data analysis, although " $\pm .0005$ " represents the uncertainty in the location of the best straight line fit to the squared amplitude; it is far smaller than the uncertainty of our result for the factor.

We also used Blackbox to give the smoothest curves of amplitude squared that fitted the six linear combinations with  $\chi^2 = 6$ . These smoothest functions are shown in Figure 20 for correction factors of 1.021 and  $1.021 \pm .003$  (the range of error in Thiessen's measurement). If the correction factor is especially low we can qualitatively reason that the cross section must have been high near threshold to give the observed data for  $\gamma + p \rightarrow \eta + p$ . Quantitatively, this feature is shown as the squared amplitude curves upward on approaching threshold for factor = 1.018. Conversely, if the correction factor is high the squared amplitude must curve down as threshold is approached (as for factor = 1.024). Assuming an S-wave dominates near threshold, we expect the amplitude to move smoothly to a non-zero value at threshold. From Figure 20, we see that a value of the correction factor outside the error bars given by Thiessen would give unreasonable threshold behavior. We can turn this argument around to see whether the eta appears to be photoproduced in an S-wave or in a P-wave (either of

which by itself would give the spherically symmetric cross section observed near threshold<sup>(21)</sup>). Assuming the validity of Thiessen's measurement and his expected error, we see from Figure 20 that 1.018 and 1.021 are both incompatible with P-wave, but that a P-wave dominance is possible if the true correction factor is at the upper end of the range of error of Thiessen's measurement.

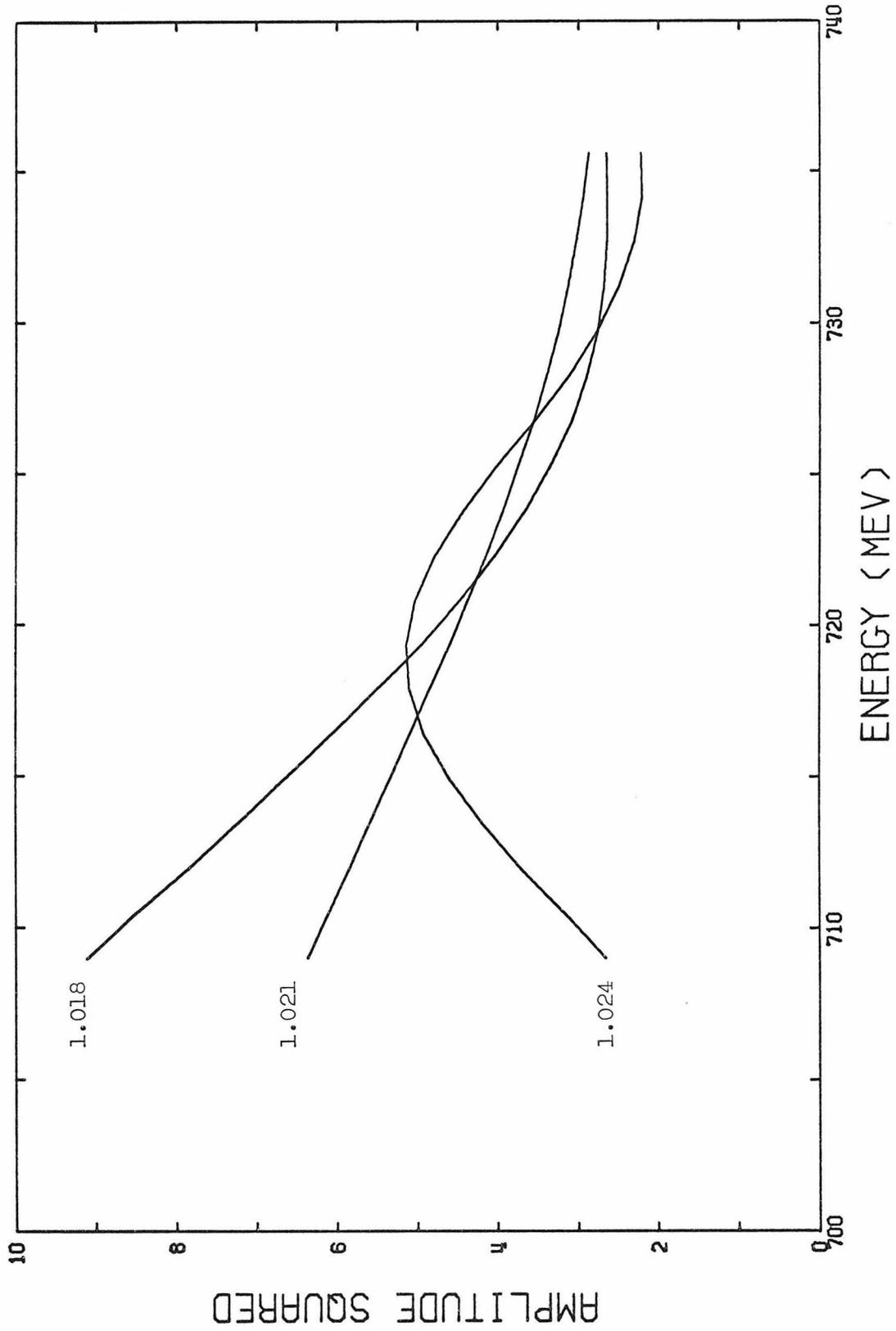


Figure 20. Effect of Varying Synchrotron End-point.

## APPENDIX G Kinematics

Let  $|\vec{P}\rangle$  be a state of a system with three-momentum  $\vec{P}$ . We normalize such states so that  $\langle\vec{P}'|\vec{P}\rangle = \theta E_p \delta^3(\vec{P}'-\vec{P})$  where  $\theta$  is some Lorentz invariant and  $E_p$  is the energy of the system. This form of the normalization is chosen so that Lorentz invariant operators will have Lorentz invariant matrix elements. Gasiorowicz<sup>(37)</sup> takes  $\theta = 2$  for bosons and  $\theta = 1/M$  for fermions. Barut<sup>(49)</sup> takes  $\theta = (2\pi)^3$  for all particles.

From the relationship

$$|\vec{X}\rangle = \int \frac{d^3\vec{P}}{(2\pi)^3} e^{-i\vec{P}\cdot\vec{X}} |\vec{P}\rangle \quad (\hbar = c = 1)$$

we get

$$\langle\vec{X}|\vec{P}\rangle = \frac{\theta E_p}{(2\pi)^3} e^{i\vec{P}\cdot\vec{X}}$$

so that our normalization corresponds to  $\frac{\theta E_p}{(2\pi)^3}$  particles per unit volume. We use the words "particle" and "system" interchangeably.

We use S and T matrices related according to

$$S = 1 + (2\pi)^4 i \delta^4(P_f - P_i) T,$$

where  $P_i$  is the total four-momentum of the initial state and  $P_f$  is that of the final state. Then the differential cross section for two initial systems colliding to form a set of final systems is

$$d\sigma_{fi} = \frac{(2\pi)^{10} \delta^4(P_f - P_i)}{K^* W \prod_{j \in f, i} \theta_j} |T_{fi}|^2 \prod_f \frac{d^3 \vec{K}_f}{E_f} \quad G.1$$

$K^*$  is the magnitude of the momentum of either of the initial systems in the center-of-mass,  $W$  is the total energy in the center-of-mass,  $\prod_{j \in f, i} \theta_j$  means the product of the  $\theta$ 's for all particles, and the  $\vec{K}_f$  are the momenta of the final particles.

In the special case of two particles in the final state,

$$\frac{d\sigma_{fi}}{d\Omega} = \frac{(2\pi)^{10}}{\prod_{j \in f, i} \theta_j} |T_{fi}|^2 \frac{q^*}{K^* W^2}$$

where  $d\Omega$  is the differential solid angle in the center-of-mass, and  $q^*$  is the magnitude of the center-of-mass momentum of one of the final particles.

Consider the case of photoproduction of scalar mesons from nucleons and specialize to  $\theta = (2\pi)^3$  ( $E$  particles per unit volume). Let  $\hat{K}$  be the unit vector along the photon direction,  $\hat{e}$  be the polarization of the photon, and  $\hat{q}$  be the vector along the meson direction. Then

$$\frac{d\sigma_{fi}}{d\Omega} = \frac{q^*}{K^*} |A_{fi}|^2 \quad G.2$$

where

$$A_{fi} = \frac{T_{fi}}{2\pi W} = L + i \vec{\sigma} \cdot \vec{K} \quad G.3$$

is the set of Pauli spin matrices (acting on the initial and final nucleons). In terms of the CGLN amplitudes(50)

$$L = \hat{q} \cdot (\hat{K} \times \hat{\epsilon}) F_2 = \hat{\epsilon} \cdot (\hat{q} \times \hat{K}) F_2$$

G.4

$$\vec{K} = \hat{\epsilon} F_1 + \hat{q} \times (\hat{K} \times \hat{\epsilon}) F_2 + \hat{K} (\hat{q} \cdot \hat{\epsilon}) F_3 + \hat{q} (\hat{q} \cdot \hat{\epsilon}) F_4 \quad .$$

For relations between the above defined amplitudes and helicity amplitudes, helicity coefficients, helicity elements, and CGLN multipole coefficients, see reference 18. A resonance contributes to the electric dipole,  $E_{0+}$  if and only if it is  $S_{11}$ . A resonance contributes to the magnetic dipole,  $M_{1-}$ , if and only if it is a  $P_{11}$ .  $E_{0+}$  contribute only to  $F_1$  and  $M_{1-}$  contributes only to  $F_2$ .

Note that in equation G.3,  $A_{fi}$  is a 2X2 matrix, while in equation G.2  $A_{fi}$  is considered to be one of the matrix elements of that 2X2 matrix. In what follows, when I ignore nucleon spins it is because I am considering quantities such as A to be matrices which are to be placed between spinors only when it is time to sum over final and average over initial spins.

In our experiment, instead of a photon-nucleon initial state, we have a photon-deuteron initial state. We can approximately describe the deuteron as a linear combination of proton-neutron pairs:

$$|\vec{0}\rangle_d = \int \frac{d^3P}{(2\pi)^3} \phi(\vec{P}) |\vec{P}\rangle_p |-\vec{P}\rangle_n \quad G.5$$

where  $|\vec{0}\rangle_d$  is the state vector for a deuteron at rest,  $|\vec{P}\rangle_p$  is the state

vector for a proton of momentum  $\vec{P}$ , and  $|\vec{-P}\rangle_n$  is the state vector for a neutron of momentum  $-\vec{P}$ . Our normalization convention requires that

$${}_d\langle\vec{P}_d | \vec{0}\rangle_d = (2\pi)^3 E_{P_d} \delta^3(\vec{P}_d),$$

where for small  $P_d$  we can obtain  $|\vec{P}_d\rangle_d$  by making a small Lorentz boost on equation G.5. It follows easily that

$$\int \frac{d^3P}{(2\pi)^3} \frac{E_p}{2} |\phi(\vec{P})|^2 = 1 .$$

Following Chew and Lewis,<sup>(44)</sup> we assume the validity of the impulse approximation and say that

$$T = T_p + T_n$$

where  $T_p$  is the amplitude for photoproduction off protons, while  $T_n$  is that for photoproduction off neutrons. Our reaction begins with a photon of momentum  $\vec{K}$  colliding with a linear combination of proton-neutron pairs with proton momentum  $\vec{P}_p = \vec{P}$  and neutron momentum  $\vec{P}_n = -\vec{P}$ . As a consequence of the collision we form an eta of momentum  $\vec{q}$ , a proton of momentum  $\vec{P}_p'$ , and a neutron of momentum  $\vec{P}_n'$ . Then  $T_p$  is more precisely described by the equation

$$\langle\vec{P}_p', \vec{P}_n', \vec{q} | T_p | \vec{P}_p, \vec{P}_n, \vec{K}\rangle = \langle\vec{P}_p', \vec{q} | T_p | \vec{P}_p, \vec{K}\rangle \langle P_n' | P_n \rangle$$

and similarly for  $T_n$ . Figure 21 shows the impulse approximation in terms of diagrams.  $M_d$  is the mass of the deuteron.

From this diagram, it is intuitively clear that the impulse approximation is valid insofar as the proton and neutron of the deuteron

are so loosely bound that they have little effect on each other during the short time in which photoproduction takes place. We neglect multiple scattering and the effect of one nucleon shading the other. References 44 and 46 give a more complete justification of the method.

From what has been said above, one finds that

$$T_{fi} = E_n' \phi(-\vec{P}_n') T_1 + E_p' \phi(-\vec{P}_p') T_2$$

where  $T_1$  is the matrix element for photon + proton of momentum  $-P_n'$  forming  $P_p' + q$ ,  $T_2$  is the  $T_n$  matrix element for photon + neutron of momentum  $-P_p'$  forming  $P_n' + q$ , and  $T_{fi}$  is the matrix element for photon + deuteron  $\rightarrow P_p' + P_n' + q$ .

For convenience, define  $u(\vec{P}) = \phi(\vec{P}) \sqrt{\frac{E_p}{2}}$ . Then  $u(\vec{P})$  is a conventionally normalized wave function.

Recall that  $A = T/2\pi W$  for photoproduction from either protons alone or neutrons alone. Let  $A_1$ ,  $W_1$ , and other quantities with subscript "1" refer to photoproduction from protons, while quantities with subscript "2" refer to photoproduction from neutrons. The absence of a numerical subscript indicates corresponding quantities for photoproduction from deuterons. Equation G.1 becomes

$$\begin{aligned} d\sigma_{fi} = & \frac{2 \delta^4(P_f - P_i)}{(2\pi)^3 K^* W} \frac{d^3 P_p'}{E_p'} \frac{d^3 P_n'}{E_n'} \frac{d^3 q}{\omega} \left[ |u(-P_n')|^2 E_n' W_1^2 |A_1|^2 + \right. \\ & + |u(-\vec{P}_p')|^2 E_p' W_2^2 |A_2|^2 + \\ & \left. + 2 \sqrt{E_n' E_p'} W_1 W_2 R_e (A_1^* u^*(-P_n') A_2 u(-\vec{P}_p')) \right] \end{aligned} \quad G.6$$

$\omega$  is obviously the eta energy.

Before discussing G.6, we consider the question of spins. With  $A_j$  ( $j = 1$  or  $2$ ) in the form  $A_j = L_j + i \vec{\sigma}_j \cdot \vec{K}_j$  it is easy to sum over final, and average over initial, spins. We use the fact that the two nucleons of the deuteron have a combined spin of one to find that we must make the replacement

$$|A_j|^2 \rightarrow |L_j|^2 + |\vec{K}_j|^2$$

$$A_1^* A_2 \rightarrow L_1^* L_2 + \frac{1}{3} \vec{K}_1^* \cdot K_2 \quad .$$

Notice the factor of three suppression of the  $S_{11}$  interference term (from G.4 and the sentences following shortly after G.4, it is seen that the  $S_{11}$  resonance affects  $\vec{K}$  but not  $L$ ).

Now we consider the term of G.6 proportional to  $|A_1|^2$ . Because  $\delta^4(P_f - P_i)$  is an invariant, we can integrate out the delta function in the center-of-mass frame of the outgoing proton and eta.

$$\int \delta^4(\vec{P}_f - \vec{P}_i) \frac{d^3 P_p}{E_p} \frac{d^3 q}{\omega} = \frac{q_1^* d\Omega_1}{W_1}$$

where  $q_1^*$  is the momentum of the eta in the final proton-eta center of mass, and  $\Omega_1$  is the eta direction relative to the incoming photon in the same frame.

The contribution from the  $|A_1|^2$  term is

$$(\mathrm{d}\sigma_{\mathrm{fi}})_1 = \frac{2W_1}{K^*W} \frac{\mathrm{d}^3P}{(2\pi)^3} u^2(P) |A_1|^2 q_1^* \mathrm{d}\Omega_1$$

where  $\vec{P}$  = momentum of initial proton =  $-\vec{P}'_n$ .

$K^*W = (\vec{K} \cdot \vec{P})_{\mathrm{d.c.m.}} = (\vec{K} \cdot \vec{P})_{\mathrm{lab}} = KM_d$  (K evaluated in the lab;  $K^*$  evaluated in the center-of-mass).

If we define  $K_1$  to be the photon energy in the rest frame of the initial proton, then  $K_1^*W_1 = K_1M$  with  $M$  = mass of proton = mass of neutron, so that

$$(\mathrm{d}\sigma_{\mathrm{fi}})_1 = \frac{2M}{M_d} \left(\frac{K_1}{K}\right) \frac{\mathrm{d}^3P}{(2\pi)^3} |u(P)|^2 \frac{q_1^*}{K_1^*} |A_1|^2 \mathrm{d}\Omega_1 .$$

At this point we modify our equations to approximately include the kinematic effect of the final state interaction. We approximate the situation by saying that the eta is unaffected by the spectator nucleon, but that the two nucleons attract each other sufficiently to slow them down enough to conserve overall energy. For the term in G.6 involving  $|A_1|^2$ , then, we act as if there were no final state interaction. Then we find the factor by which the phase space decreases when the final nucleons are slowed down and correct our results by that factor.

We define the yield to be the number of etas detected per unit beam energy at a given synchrotron endpoint and at a given counter setting. The unit of beam energy is the "bip" as standardized for this experiment in section II.B (1 bip =  $1.2132 \times 10^{13}$  MeV).

We would like to express the proton's contribution to the yield as

$$\text{Yield}_1 = \int d\mathbf{K}_1 \frac{dY_1}{d\mathbf{K}_1}$$

so that we can use Blackbox to unfold the cross section (see Appendix D). Take the z axis along the direction on the photon beam. Then

$$\frac{K_1}{K} = \sqrt{\frac{P^2 + M^2 - P_z}{M}}$$

Keeping  $\vec{P} = -\vec{P}_n$  constant, we get

$$\frac{dY_1}{d\mathbf{K}_1} = \frac{dY_1}{d\mathbf{K}} \frac{K}{K_1} .$$

For a given counter setting, let the probability of detecting the eta be called  $\mathcal{P}(\vec{q})$ .  $\vec{q}$ , the laboratory momentum of the eta, is to be considered a function of  $\vec{K}_1$ ,  $\vec{P}$ , and  $\Omega_1$ . Suppose there are  $N_d$  deuterons per unit area of the incoming photon beam and let "D" be the number of MeV per bip.  $E_0$  is the synchrotron endpoint energy and  $B(K/E_0; E_0)$  is the bremsstrahlung function normalized so that the number of photons per MeV is

$$\frac{B(K/E_0; E_0)}{KE_0} .$$

Then

$$\frac{dY_1}{d\mathbf{K}_1} = \frac{DN_d}{E_0 K_1^*} \int d\Omega_1 |A_1|^2 \int \frac{d^3 P}{(2\pi)^3} |u(\vec{P})|^2 q_1^* \frac{B(K/E_0; E_0)}{K} r \mathcal{P}(\vec{q}) ,$$

where  $r$  is the factor by which phase space is corrected for the final state interaction.

If we define functions  $f_J^{(p)}$  by

$$|A_{\perp}|^2 = \sum_{J=0}^{\infty} P_J(\cos \theta_{\perp}) f_J^{(p)}(K_{\perp}) \quad G.7$$

then

$$\frac{dY_{\perp}}{dK_{\perp}} = \sum_{J=0}^{\infty} g_J(K_{\perp}) f_J^{(p)}(K_{\perp}) \quad G.8$$

where

$$g_J(K_{\perp}) = \frac{q_{\perp}^{*D} N_d}{K_{\perp}^{*} E_o} \int \frac{d^3 P}{(2\pi)^3} |u(\vec{P})|^2 d\Omega_{\perp} P_J(\cos \theta_{\perp}) \frac{B(K/E_o; E_o)}{K} r \mathcal{P}(\vec{q}) \quad G.9$$

The effect of the contribution from the  $|A_2|^2$  term is to replace  $f_J^{(p)}$  with  $f_J^{(p)} + f_J^{(n)}$ .

Applying the operator TP (time reversal times parity) to equation G.5 leads to the conclusion that if the deuteron Hamiltonian is TP invariant then  $u(\vec{P})$  can be taken real. To be sure, G.5 ignores spin, but including spin explicitly doesn't change this conclusion.

A common approximation to  $u(\vec{P})$  is the Hulthén wave function:<sup>(44)</sup>

$$u(\vec{P}) = C \left( \frac{1}{P^2 + \alpha^2} - \frac{1}{P^2 + \beta^2} \right)$$

$$C = \frac{1}{\beta - \alpha} \sqrt{8\pi \alpha \beta (\alpha + \beta)}$$

with  $\alpha$  and  $\beta$  chosen empirically to make the Hulthén wave function a good approximation to the actual wave function.  $\alpha$  is about 45.4 MeV, and  $\beta$  is about 276 MeV.

The deuteron form factor is defined by

$$F(Q) = \int d^3r \underline{u}^2(\vec{r}) e^{i\vec{Q}\cdot\vec{r}},$$

where  $\underline{u}(\vec{r})$  is the Hulthén wave function in terms of positional coordinates. Equivalently,

$$F(Q) = \int \frac{d^3P}{(2\pi)^3} u(\vec{P}) u(\vec{Q}-\vec{P}) = \frac{C^2}{4\pi Q} \left[ \tan^{-1} \frac{Q}{2\alpha} + \tan^{-1} \frac{Q}{2\beta} - \tan^{-1} \frac{Q}{\alpha+\beta} \right].$$

The relevance of this form factor to our experiment will be seen shortly.  $F(Q)$  is graphed in Figure 22 and  $u(P)$  is displayed in Figure 23.

We used the Hulthen wave function to evaluate  $g_J$  by a Monte Carlo method (see Appendix B). We chose random  $\Omega_{\perp}$  with uniform density over a  $4\pi$  solid angle and chose  $\vec{P}$  at random with density  $\frac{u^2(\vec{P})}{(2\pi)^3}$ . The computation of  $\mathcal{P}(\vec{q})$  can also be done by the Monte Carlo method, but to do so consumes several hours of IBM 360-75 computer time. In one of two programs for computing  $g_J$ ,  $\mathcal{P}(\vec{q})$  was taken from the results of a program for computing the detection efficiency for photoproduction from hydrogen. In the other program, the Monte Carlo program for computing the quantities corresponding to  $g_J$  for hydrogen was modified to include the effect of the deuteron fermi motion. While much slower, the second program was helpful as a check on the first.

For simplicity, we will evaluate the interference term of G.6 with more drastic approximations than those hitherto used. To exhibit more clearly the nature of the other terms in  $d\sigma_{fi}$ , we will temporarily apply those same approximations to the terms proportional to  $|A_1|^2$  and  $|A_2|^2$ , although when analyzing this experiment we really handled those terms with equations G.7 through G.9.

From the graph of  $u(P)$ , it can be seen that the contribution from  $P$  greater than 200 MeV is strongly suppressed. In fact, the probability of a nucleon in a deuteron having more than 200 MeV Fermi momentum is only about 6%. To within a few percent we can approximate

$$E_P = \sqrt{P^2 + M^2} \approx M.$$

$W_1^2 = M^2 + 2K(E_p - P_z)$  (where the  $z$  axis is along the direction of photon motion) varies more rapidly with  $\vec{P}$  than does  $E_p$ , but even so is typically within a few percent of its value for  $\vec{P} = 0$  when  $u(\vec{P})$  is large. We would like to be able to say that  $A_i$  is nearly constant as a function of  $W_i$  over the range in which the Hulthen wave function is large. But the strong enhancement of the cross section near threshold (attributed to the  $S'_{11}(1535)$ ) indicates an amplitude that at least for photoproduction of etas from protons varies rapidly with  $W_1$ . Nonetheless, we neglect the variation of the amplitudes with  $P_p'$  and  $P_n'$ . We also neglect final state interactions.

With a given counter setting, the probability of detecting a photoproduced eta depends only on the laboratory momentum of the eta. Since we don't detect the final nucleons, it is appropriate to

integrate equation G.6 over  $\vec{P}_p'$  and  $\vec{P}_n'$ , obtaining with the above described approximations

$$d\sigma_{fi} = \frac{2d^3q \delta(E_f - E_i)}{\omega K^*W} \left[ W_0^2 \frac{|A_1|^2}{E_p'} + \frac{2W_1W_2 F(Q)}{\sqrt{E_p' E_n'}} \operatorname{Re}(A_1^* A_2) + W_0^2 \frac{|A_2|^2}{E_n'} \right]$$

where  $\frac{|A_1|^2}{E_p'}$  is evaluated at  $\vec{P}_n' = 0$ ,  $\frac{|A_2|^2}{E_n'}$  is evaluated at  $\vec{P}_p' = 0$ ,

and  $W_0$  is the same as  $W_1$  (or  $W_2$ ) evaluated at  $\vec{P}_n'$  (or  $\vec{P}_p'$ ) = 0. With

$\vec{P}_n' = 0$ ,  $K_1^*W_0 = KM$ , while with  $\vec{P}_p' = 0$ ,  $K_2^*W_0 = KM$ .  $\vec{Q} = \vec{P}_p' + \vec{P}_n'$ .

$\vec{Q}$  equals  $\vec{P}_p'$  for the  $|A_1|^2$  term and  $\vec{P}_n'$  for the  $|A_2|^2$  term. For

the term proportional to  $\operatorname{Re}(A_1^* A_2)$ , we notice that  $u(-\vec{P}_n') u(-\vec{P}_p')$

tends to be large when either  $\vec{P}_n'$  or  $\vec{P}_p'$  is zero. Following an

argument of Hadjioannou<sup>(45)</sup> we consider the interference term (including

its delta function over energy) to be averaged over the situation

in which  $\vec{P}_n'$  is zero and the situation in which  $\vec{P}_p' = \vec{Q} - \vec{P}_n' = 0$ . Then

$$d\sigma_{fi} = \frac{2M}{M_d} W_0 \frac{d^3q}{\omega K^*E_Q} \left[ |A_1|^2 + |A_2|^2 + 2 \frac{E_Q}{M} F(Q) \operatorname{Re}(A_1^* A_2) \right] \delta(E_f - E_i),$$

where  $K^*$  is evaluated in the center-of-mass of the photon with a

nucleon at rest in the lab,  $E_Q$  is the energy of a nucleon of momentum

$\vec{Q}$ ,  $E_f =$  final energy =  $\omega + M + E_Q$ , and  $E_i =$  initial energy =  $K + M_d$ .

We neglect the difference between  $M$  and  $\frac{1}{2}M_d$  and the difference between

1 and  $\sqrt{\frac{E_Q}{M}}$ . Then it can be shown that

$$\begin{aligned} \frac{d\sigma_{fi}}{d\Omega} &= \frac{q^*}{K^*} \left[ |A_1|^2 + |A_2|^2 + 2F(Q) \operatorname{Re} A_1^* A_2 \right] = \\ &= \frac{d\sigma_p}{d\Omega} + \frac{d\sigma_n}{d\Omega} + \frac{2q^*}{K^*} F(Q) \operatorname{Re} (A_1^* A_2) \end{aligned} \quad \text{G.10}$$

where  $q^*$  is in the same frame as  $K^*$ .

The contributions to the yield of the first two terms of G.10 have already been discussed to a better approximation than G.10. The third term, the interference term, contributes to the yield

$$\frac{dY_{int}}{dK} = 2 DN_d \frac{q^*}{K^*} \operatorname{Re} (A_1^* A_2) \frac{B(K/E_o; E_o)}{K E_o} \int d\Omega \mathcal{P}(q) F(Q), \quad \text{G.11}$$

where except for  $A_1^* A_2$  all quantities are evaluated with kinematics corresponding to photoproduction from a single nucleon at rest.

$A_1^* A_2$  is evaluated half with the proton at rest and half with the neutron at rest, and we neglect angular variation of  $A_1$  and  $A_2$  because  $S_{11}$  is expected to dominate near threshold. Strictly speaking,  $\operatorname{Re}(A_1^* A_2)$  depends on  $\Omega$  through the variation with  $\Omega$  of the energy at which the  $A_i$  corresponding to the moving nucleon is evaluated. But for backward photoproduction, the small  $F(Q)$  prevents the interference term from contributing significantly; so we don't have to worry about errors from evaluating one of the  $A_i$  at an incorrect energy. For forward photoproduction, both  $\mathcal{P}(q)$  and  $F(Q)$  are largest when the direction of  $\Omega$  is along the beam line. So the major contribution of  $\operatorname{Re}(A_1^* A_2)$  occurs when  $\Omega$  is along the beam line, and we evaluate  $\operatorname{Re}(A_1^* A_2)$  for such  $\Omega$ .

The detectors of the photons from eta decay were placed symmetrically about the beam line at half angle  $\theta$ . Such a setting is optimal for detecting etas moving along the beam line with velocity such that  $\beta = \cos(\theta)$ . An eta of this velocity can be photoproduced by a photon striking a nucleon at rest if the photon has energy  $K(\beta)$ . A typical momentum transfer for a given detector setting is then

$$Q(\theta) = K(\beta) - q(\beta) = \frac{MT(2 - T)}{2(1 - T)}$$

$$T = \frac{M_\eta}{M} \tan \theta/2 \quad (M_\eta = \text{mass of the eta}) .$$

$F(Q(\theta))$  is then a measure of the importance of the interference term at a given setting, and is displayed in Figure 24. For backward photoproduction ( $\theta$  greater than 65%) the interference term is negligible.

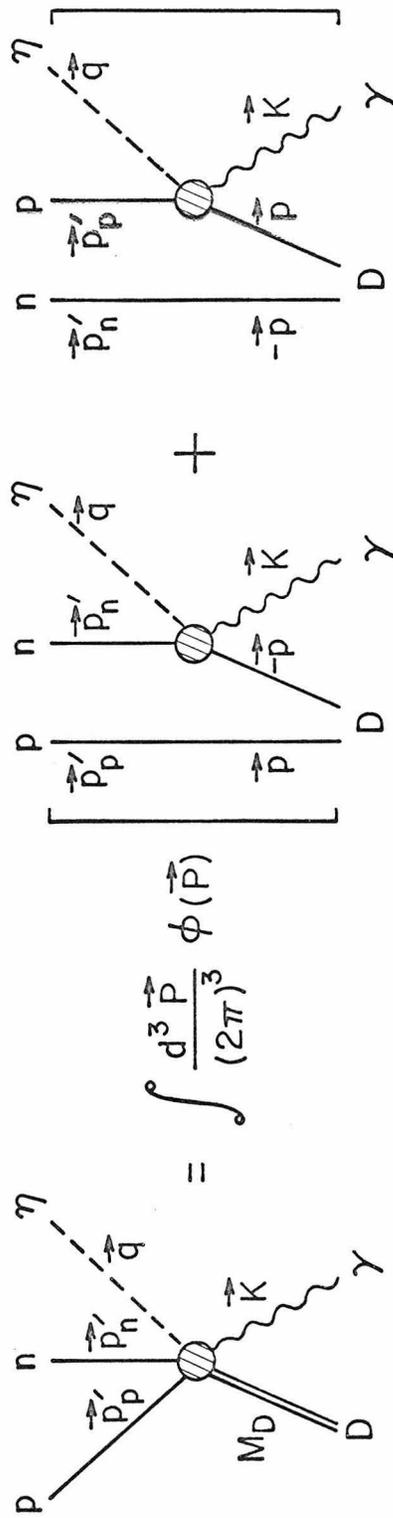


Figure 21. Impulse Approximation.

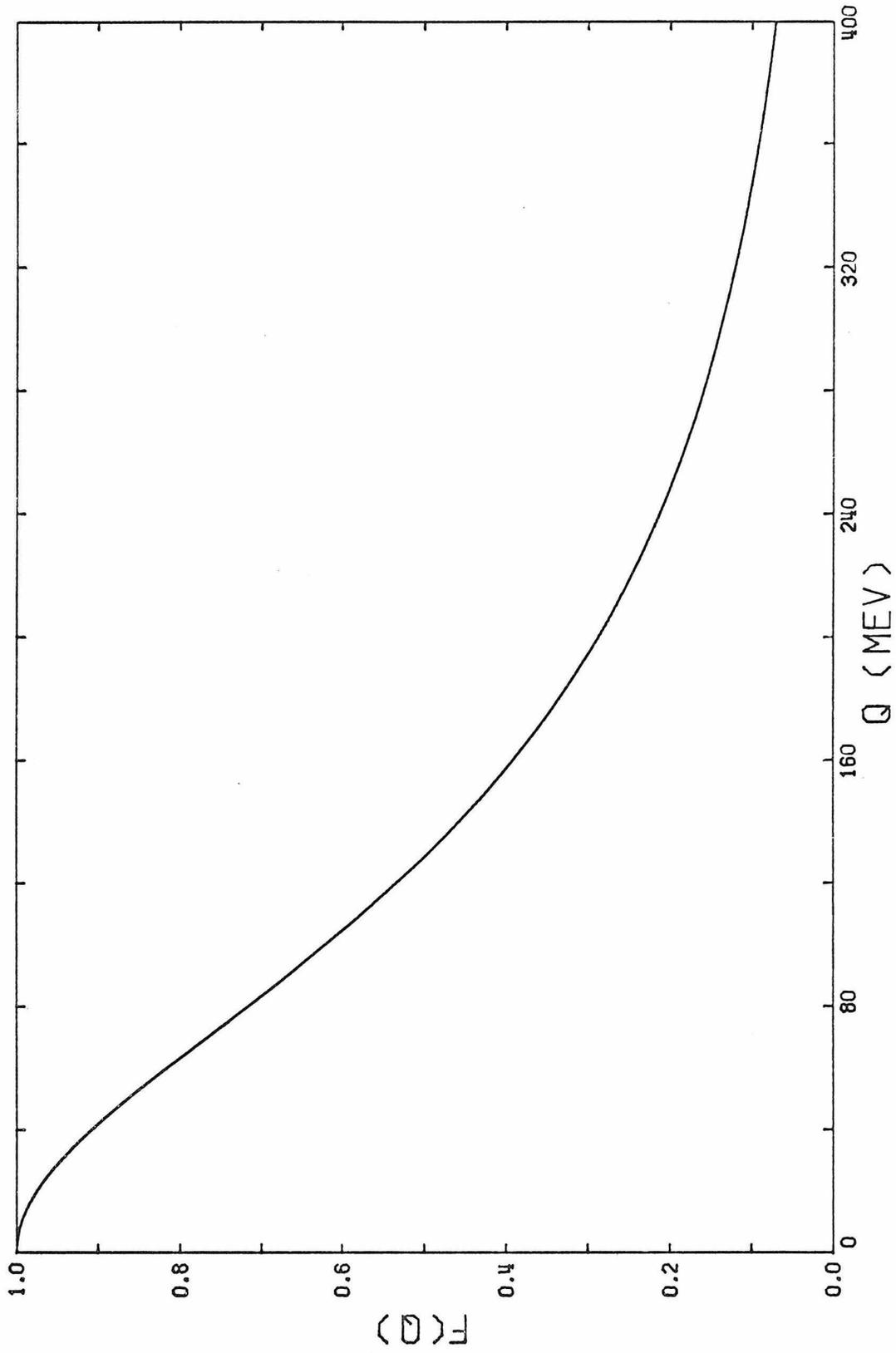


Figure 22. Form Factor for Hulten Function.

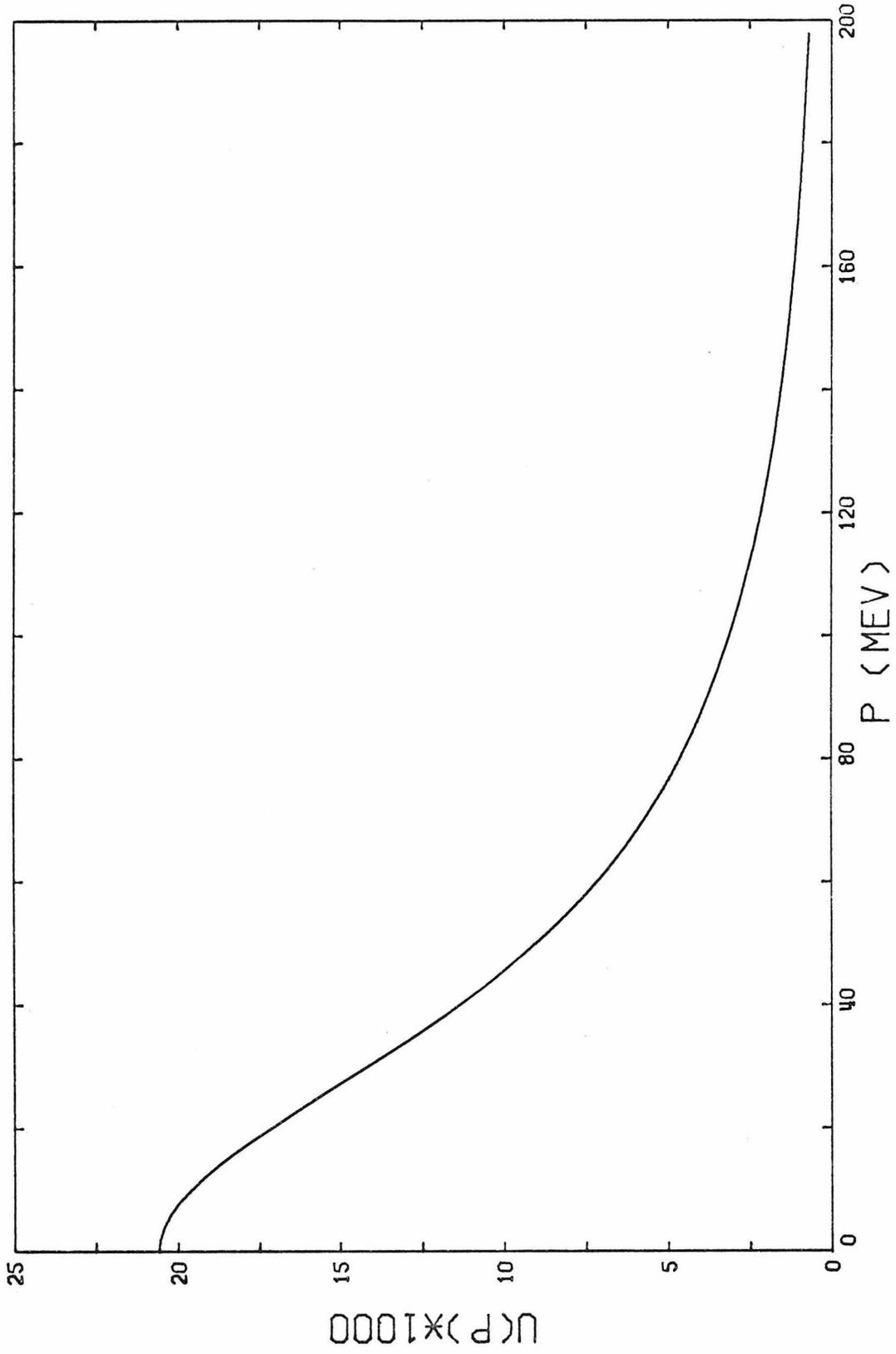


Figure 23. Hulthen Wave Function.

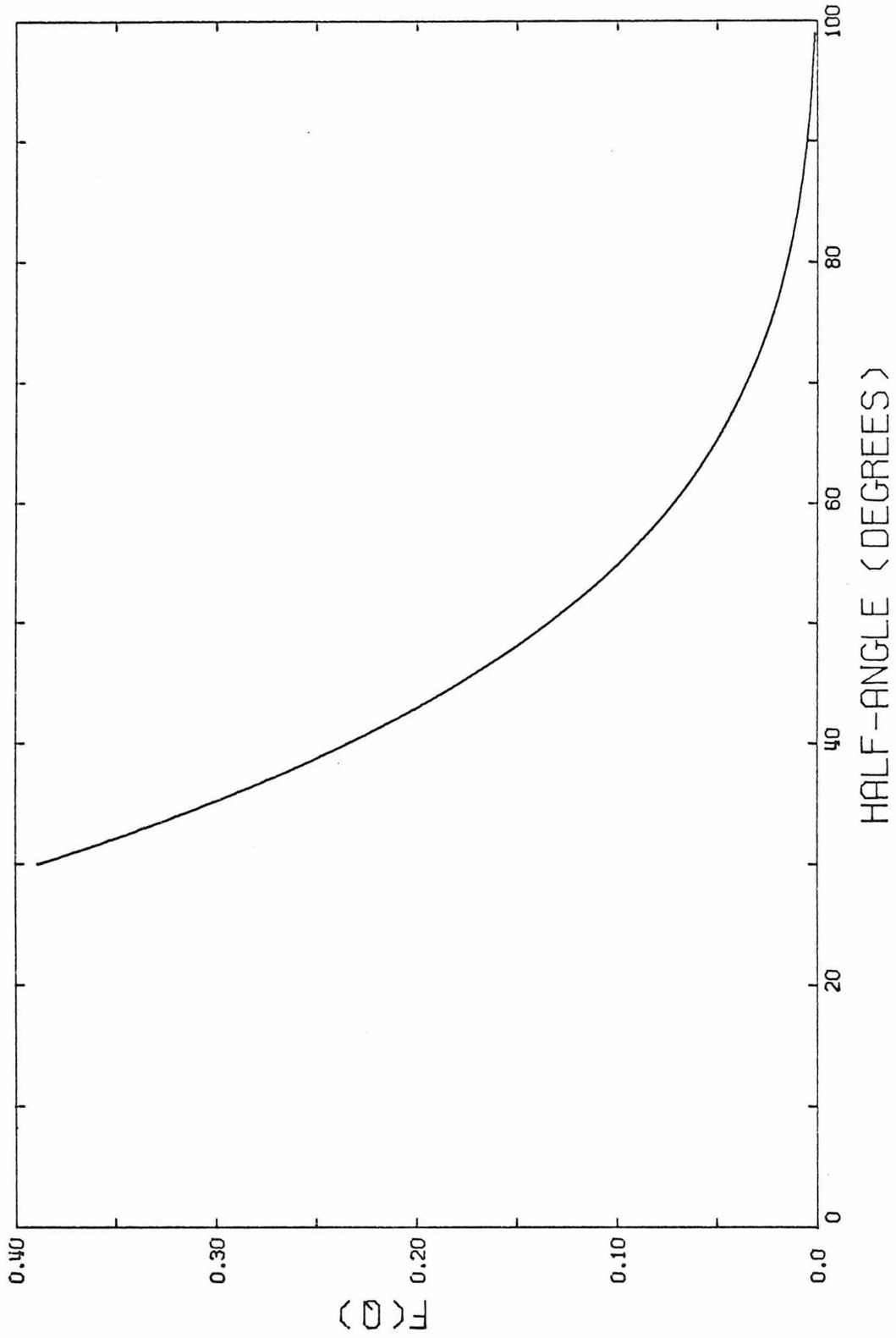


Figure 24. Relationship Between Form Factor and Counter Half-Angle.

## APPENDIX H Details of the Cherenkov Counters

Figure 25 is a diagram of one of the two totally absorbing Cherenkov counters. In each counter there were two lead glass blocks optically joined together. The properties of the glass are summarized in Table VI.1, taken from reference 14. Seven XP1040 phototubes were optically joined to the back of the blocks with RTV compound in a pattern shown in Figure 26. Each phototube was wrapped with several layers of Netic and Conetic magnetic shielding. The lead glass blocks and their phototubes were placed in a  $\frac{1}{4}$  inch thick steel box, which was surrounded by about forty windings of M19 26 gauge hot rolled, non-oriented silicon steel for further shielding against the magnetic field of the synchrotron. All this material surrounding the Cherenkov counters served as radiation shielding as well as magnetic shielding. Each counter with its shielding was placed upon a trolley which rolled with two wheels on a circular track and two wheels on the platform holding the track. Also on the trolley were mounted a veto counter (made from one half inch of scintillator shielded with one half inch of lucite) and a  $1\frac{1}{4}$  inch thick lead wall with an aperture eight inches high. The veto counter was between the lead wall and the Cherenkov counter, and the lead wall was, of course, between the deuterium target and the veto counter. Most of the experiment was performed with the aperture in the wall about seven inches wide, but when the target length was decreased from about 6.5 inches to 3.3 inches in order to improve our kinematic resolution, the aperture was reduced to 3.5 inches width. The distance between the counter and the nearest

side of the lead wall was about five inches, give or take a couple of inches. This distance was not critical because it did not determine the geometrical acceptance of the counter. The acceptance was determined by the size of the aperture in the lead wall and by the distance from the lead wall to the target (27.6 inches from the center of the deuterium target to the side of the lead wall nearest the lead glass).

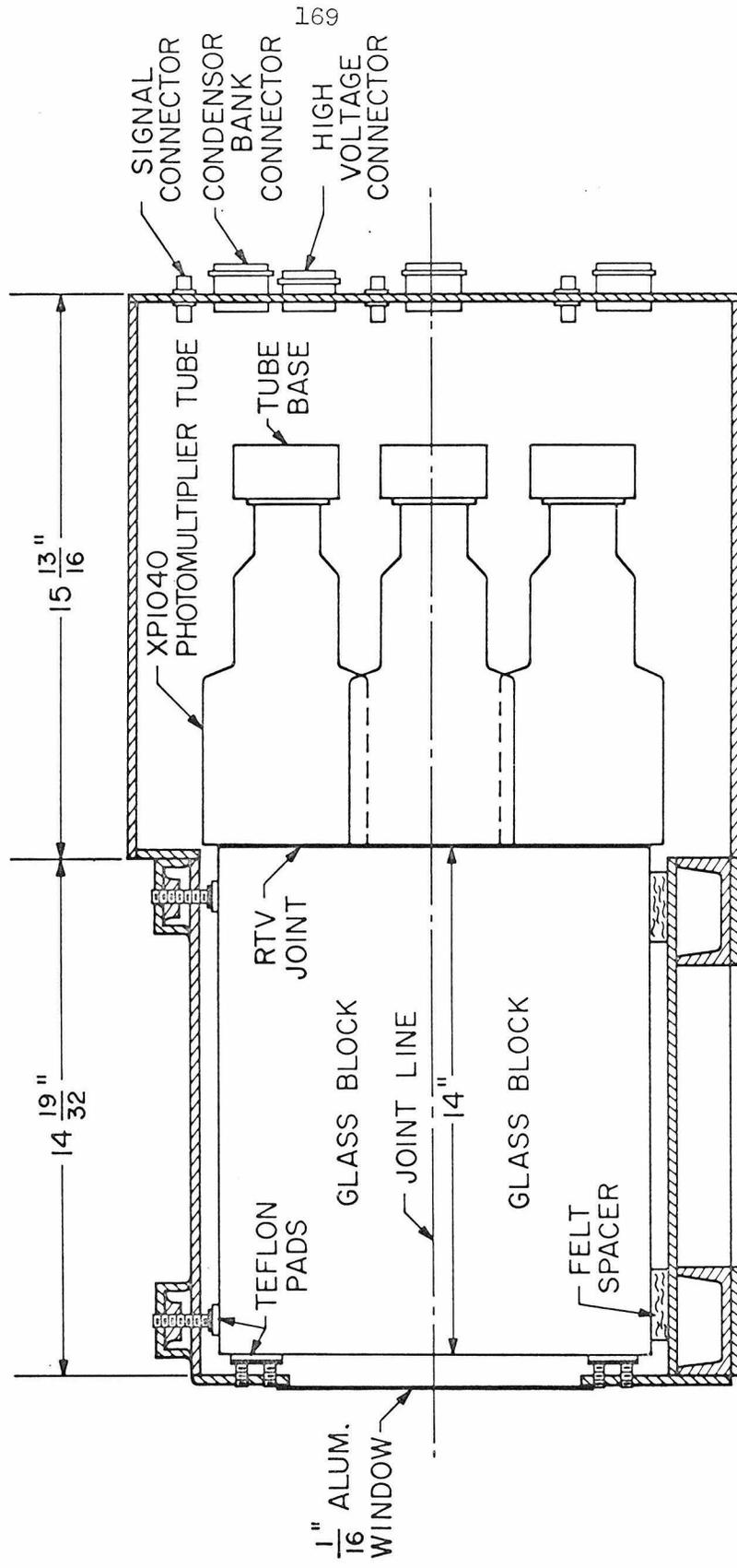


Figure 25. Cherenkov Counter.

TABLE VI.1

## Summary of Lead Glass Characteristics

Glass Type	DF-4	Code	649-339
Density	3.88 gm/cm <sup>3</sup>		
Index of refraction	$n_D = 1.649$	Dispersion	$n_F - n_C = .019$
Radiation length	2.5 cm		
Energy loss for a 500-MeV electron	4.5 MeV/cm		
Composition	K <sub>2</sub> O	6 percent	
	SiO <sub>2</sub>	41 percent	
	PbO	52 percent	

Manufactured by Hayward Scientific Glass Corporation, Whittier,  
California.

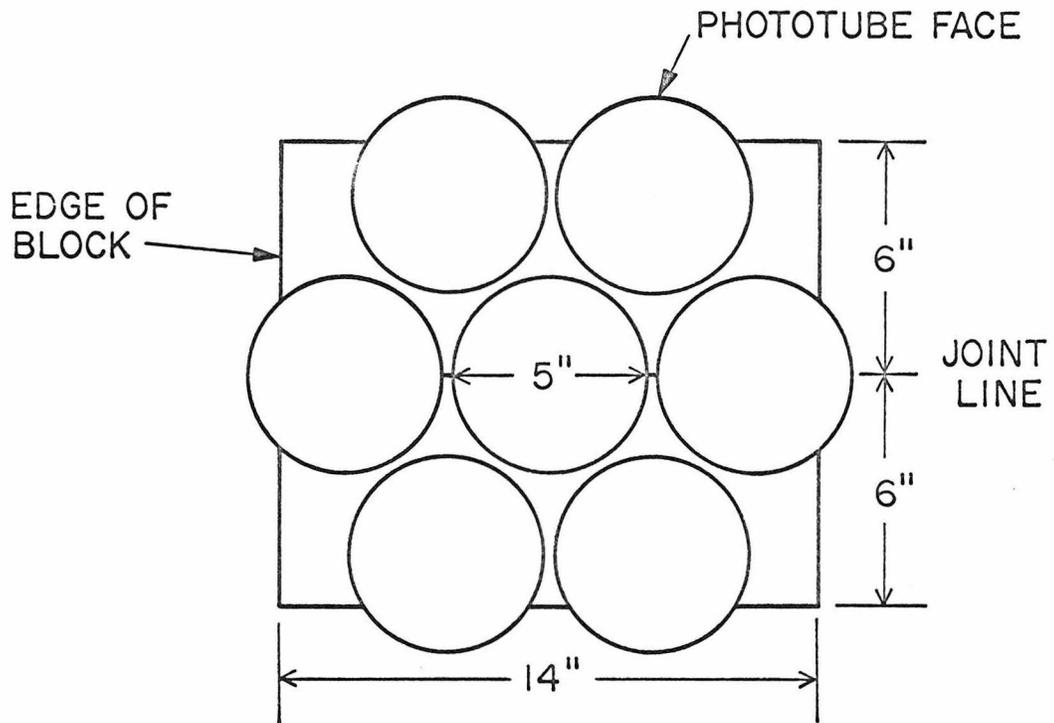


Figure 26. Phototube Placement.

## APPENDIX I Error Analysis

Frequently when discussing the extraction of the yield from the background, we ignored or postponed discussion of errors and ambiguities. The statistical error was calculated by the method derived in Appendix E, but other errors require more discussion. In order to check the effect of the errors and ambiguities, we made a series of tests of our fitting method, the results of which are shown in Table VI.2. This table requires a great deal of explanation. We chose three typical runs -- one corresponding to photoproduction of etas in a forward direction (counter half angle = 40 degrees, synchrotron endpoint energy = 925 MeV), one corresponding to photoproduction near threshold (60 degrees, 750 MeV), and one corresponding to photoproduction in a backward direction (85 degrees, 825 MeV). For each test performed we show the number of etas that resulted from the fitting method (the columns labeled "ETAS"), the number of standard deviations by which the fit deviated from the background (the columns labeled "BG"), and the number of standard deviations by which the fit deviated from the foreground ("FG"). By "background" I mean the background region as defined in Section III C, and by "foreground" I mean the eta region as defined in the same section. Since these definitions of Section III C were somewhat arbitrary and ambiguous, some of the tests were designed to see the effect of varying the definitions of the eta and background regions. To understand how we calculated the number of standard deviations by which the fit to a particular region deviated from perfection, see Appendix E.

In test number one we applied our normal fitting procedure to the three typical runs. The results of all other tests are to be compared with test one. To put the effect of the tests in proper perspective, we now give the statistical errors that we expect from our fitting procedure:  $\pm 42$  etas for 40 degrees, 925 MeV;  $\pm 66$  etas for 60 degrees, 750 MeV;  $\pm 39$  etas for 85 degrees, 825 MeV. These errors are the standard deviations of the number of etas as calculated by the methods of Appendix E.

Test number two was designed to see what effect we could expect from errors in our estimate of the number of accidental coincidences. Recall that in Section III C we tried to correct for the presence of accidentals in the background. For this test we took the extreme case of one hundred percent error in the estimate of the accidentals; we assumed there were no accidentals. As can be seen from the table, the effect of errors in the estimate of the accidentals was negligible.

Test number three was designed to see if we might have made our timing cuts so tight that we lost some etas. We expanded the timing cuts to include all events over a range of about twenty nanoseconds centered at the coincidence peak (which had about three nanoseconds full width at half maximum). This expanded timing cut is to be compared with the approximately eight nanosecond range used normally. The fits showed essentially no change in the number of etas. This result is not surprising, for by breaking up the data into rather narrow timing bins we found that the time width of the eta events was distinctly narrower than the three nanoseconds of the

whole coincidence peak. The eta events, after all, tended to occur with higher photon energies than most of the background, and it is not surprising that higher pulse-heights could be timed more accurately.

The next five tests checked the effect of various ambiguities in our selection of the foreground and background regions. Test number four showed what happened when we expanded the background region to include all of what formerly was in the eta region. As described in Section III C, we corrected for the "leakage" of etas into the background region. In other words, we estimated the number of etas in each bin of the background and corrected for this number before computing the best uncorrelated fit to the background.

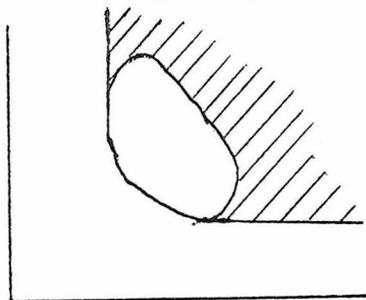
With test number five we examined the effect of expanding the foreground region to include what was formerly the entire background region. We subtracted the calculated background from the distribution of events, then found the maximum likelihood fit to the eta peak in the combined eta and background regions. The background was not reiterated; so the column labeled "BG" is left blank for test number five.

Test number six verified that raising the mass cut from 240 MeV to 280 MeV has no appreciable effect (recall that we excluded all events with measured invariant mass below 240 MeV in order to prevent the correlation at low masses from upsetting our fit).

With test number seven we examined the effect of changing the low energy cutoffs. While the discriminators in the electronic logic rejected all events with pulse-height corresponding to photon energy

below about 100 MeV, for analyzing the data we raised the cutoff to about 140 MeV in order to make the cutoff sharp. But for this test we left the cutoffs at the values determined by the discriminators.

The eighth test was the last of the tests involving modification of the regions used in our fitting procedure. Normally the background region included events whose invariant mass fell above the eta region. But we feared that the small background in that region of the distribution of events might have a significant effect on the fits. Yet this part of the background is highly subject to errors caused by inaccurate assumptions involving the computation of eta leakage into the background. The shaded area in the below diagram shows the region excluded from the background region by the eighth test.



To summarize the effect of changing the background or foreground regions, we see from Table VI.2 that all tests left the number of etas unchanged to within a standard deviation. The worst change occurred with test number five, in which we expanded the eta region.

We turn, now, to other tests of the fitting procedure. Recall that we used a Monte Carlo program to generate an eta peak that was fitted to the observed data. In this Monte Carlo program, statistics were limited by financial considerations. To test the effect of the finite statistics we fluctuated the Monte Carlo results by their

expected error and fitted the slightly modified eta peak to the data. This test, test number nine, showed that our Monte Carlo statistics were quite adequate. As a matter of fact, a programming error led us to use the same background region in this test as for test number eight. So the effect of the Monte Carlo statistics is to be seen in the comparison between tests eight and nine.

Another input to the generation of the Monte Carlo eta peak was the shape of the cross section used. Normally we tried to use a realistic cross section, but for this test we used a cross section that turned on instantly at threshold and thereafter was flat. No significant effect on the number of etas was observed for this tenth test. Note, however, that for the run taken near threshold the foreground fit was too poor to be consistent with chance. This large value of "FG" would have led us to reject this run as unreliable if we had found such a poor fit with a reasonable cross section (we rejected runs with either "FG" or "BG" greater than 2.5). At this point we make a confession. For many of the runs near threshold, the overall fit was rather poor, with "FG" or "BG" often greater than 2.5. We attributed this poor fit to the fact that the cross section used went to zero at the threshold for eta photoproduction from hydrogen, while in reality the threshold for the deuterium case is well below that for hydrogen. To test this explanation of the poor fit, for all runs taken at 725 MeV nominal synchrotron end-point we used for the foreground fit a Monte Carlo eta peak characteristic of an end-point of 750 MeV. Since the maximum likelihood fit was allowed to vary the

energy vs. the pulse-height calibration in each counter, we can expect this change to show some of the effect of the lower deuterium threshold. The fits did indeed improve considerably, but the number of etas calculated changed typically by less than one standard deviation. We used in our data analysis the eta yields corresponding to the improved fits whenever those improved fits had both "FG" and "BG" less than 2.5, as they generally did.

For the eleventh test, we doubled the dimensions of each bin in the background region, but not in the eta region (i.e., each new bin in the background had the area of four old bins). B. G. is not included for this test because an error was made in its calculation. Test eleven showed no significant difference from test one.

The resolution was, in section III C, considered to be characterized by

$$(\mu/\sigma)^2 = \alpha E$$

where  $\mu$  is the average pulse-height for showers of energy  $E$  ( $\mu$  is proportional to  $E$ ) and  $\sigma$  is the standard deviation of the pulse-height for showers of energy  $E$ .  $\alpha$  is proportionality constant which was varied in order to obtain a good fit. For the twelfth test,  $E$  was replaced by  $E + 300$  MeV, and no significant change was seen.

The most unfortunate test was test number thirteen. We computed  $f(E_1, E_2) = \gamma X$  calculated background  $X (E_1 - \bar{E}_1) (E_2 - \bar{E}_2)$ .  $\gamma$  was chosen so that calculated background +  $f$  had a coefficient of correlation between  $E_1$  and  $E_2$  of .1. The thirteenth test consisted of adding  $f$

to the  $E_1 - E_2$  distribution (thereby causing the background to have a ten percent correlation) and extracting the eta yield while assuming in our usual manner that there is no correlation. The result of this test showed large deviations from test number one. For two of the three settings, BG is unreasonably large. But since we did not know how to include the effect on the statistical fluctuations in the  $E_1 - E_2$  distribution from adding correlation, BG is unreliable for test thirteen.

The large effect correlation had in test number thirteen led us to study this effect more closely. We generated artificial runs with no correlation in the background (test number fourteen), then generated similar artificial runs with a correlation term such that there was a .1 coefficient of correlation between the energies measured by the two Cherenkov counters (test number fifteen). Notice how much larger BG became when correlation was added. BG averaged over all the runs of our experiment was about .5. This positive value indicates that our assumption of no correlation was inaccurate, but this indication of a correlation is about one-fourth that which we would expect from a .1 correlation. The effect of correlation is expected to lead to a moderately systematic error of about the same size as the standard deviation resulting from limited statistics.

There were several other sources of error in our results. For example, our quantameter calibration was known to within three percent (see section II B). A three percent error in this constant would contribute a three percent error to the cross section at all energies.

A similar systematic error comes from the uncertain branching ratio

$$\frac{\eta \rightarrow 2\gamma}{\eta \rightarrow \text{anything}} \approx .375 ,$$

taken from the world average as compiled by the Particle Data Group.<sup>(25)</sup>  
This ratio is believed to be known to within about four percent.

Geometrical errors were surprisingly serious. We failed to correct for distortion of the transit's close up lenses when surveying, and we failed to hold the counters in place as rigidly as we should have. As the trolleys holding the counters were pushed along their rails, they could distort. We estimate that our counter positions relative to the target were known to within an error of  $\pm .1''$ , leading to a partially systematic error of about  $\pm 5\%$  in the yields.

We next discuss a source of error that was expected to be large, but turned out to be small. In Appendix G we showed that within the framework of the impulse approximation, there is an interference term contributing to each yield a term proportional to  $F(Q)$  (see equation G.10). But from Figure 24, for a counter setting with a half-angle of forty degrees or less,  $F(Q)$  could exceed .2. In order to see whether the interference term could be expected to contribute significantly, we evaluated G.11 with the approximations discussed in Appendix G and using a Breit-Wigner  $S_{11}$  resonance for  $A_1$  and  $A_2$ . As was pointed out in Appendix G, summation over spins leads to a factor of three decrease in the contribution of  $S_{11}$  resonances to the interference term. Furthermore, in equation G.10, the term  $2\text{Re}(A_1^* A_2)$  has to be smaller than  $|A_1|^2 + |A_2|^2$ . Finally,  $F(Q)$  as shown in

Figure 24 is generally larger than the average  $F(Q)$  that appears in G.11. It is therefore not surprising that the calculated contribution to each yield from the interference term was always less than the statistical errors, and was usually far less than those errors.

TABLE VI.2

## Tests of Background Fitting Procedure

Test	40°, 925 MeV			60°, 750 MeV			85°, 825 MeV		
	ETAS	BG	FG	ETAS	BG	FG	ETAS	BG	FG
1	861	1.04	.50	2574	1.72	1.17	804	2.01	.25
2	870	1.03	.52	2570	1.79	1.19	807	1.87	.28
3	849	.84	.58	2570	1.50	1.07	803	1.88	.63
4	868	.95	1.17	2574	1.57	1.01	799	1.69	.07
5	820		.41	2642		-.17	825		-.26
6	853	.88	.50	2574	2.01	1.17	794	1.54	.25
7	854	.87	.31	2566	2.27	1.09	830	1.64	.31
8	862	1.71	.55	2584	-.54	1.33	811	.99	.34
9	862	1.71	.56	2584	-.54	1.34	811	.99	.30
10	864	1.03	.55	2576	1.73	3.28	804	2.01	.24
11	831		1.13	2549		1.38	804		.97
12									
13	1062	1.49	-.76	2793	4.43	.65	882	3.04	-.02

## VII. REFERENCES

1. Riccardo Levi Setti, "Elementary Particles", University of Chicago Press, p. 6, Part II (1963).
2. J. O. Maloy, CTSL-34 (unpublished).
3. H. A. Thiessen, CTSL-21 (unpublished).
4. F. Wolverton, "Manual for BPAK I, Thick Radiator Bremsstrahlung Computer Program" (unpublished).
5. R. Littauer, Rev. Sci. Instr. 25, 148 (1954).
6. R. R. Wilson, Nuclear Instruments I, 101 (1957).
7. R. Gomez, J. Pine and A. Silverman, Nuclear Instruments and Methods 24, 429 (1963).
8. J. A. Richards and L. W. Nordheim, Phys. Rev. 74, 1106 (1948).
9. D. Ritson, "Techniques of High Energy Physics".
10. H. Ruderman, R. Gomez and A. V. Tollestrup, CTSL-31 (unpublished).
11. D. F. Crawford and H. Messel, Phys. Rev. 128, 2352 (1962).
12. C. A. Heusch and C. Y. Prescott, Phys. Rev. 135, B772 (1964).
13. A. V. Tollestrup and R. Gomez, "A Multi-Channel Pulseheight Analyzer", (unpublished).
14. A. V. Tollestrup, CTSL-31 (unpublished).
15. R. F. Deery, CTSL-11 (unpublished).
16. H. W. Wooley, R. B. Scott, F. G. Brickwedde, J. Research NBS 41, 379 (1948).
17. Y. C. Chau, N. Dombey and R. G. Moorhouse, Phys. Rev. 163, 1632 (1967).

18. R. L. Walker, Phys. Rev. 182, 1729 (1969).
19. R. G. Moorhouse and W. A. Rankin, Nuclear Physics B (to be submitted).
20. W. B. Richards et al., Phys. Rev. D 1, 10 (1970).
21. B. Delcourt et al., Physics Letters 29B, 75 (1969).
22. B. D. Winstein, Ph. D. Thesis, California Institute of Technology (1970).
23. C. Bacci et al., Physics Letters 28B, 687 (1969).
24. K. M. Watson, Phys. Rev. 85, 852 (1952).
25. Particle Data Group, Physics Letters 33B, (1970).
26. H. J. Lipkin, Physics Letters 12, 154 (1964).
27. W. A. McNeely, Ph. D. Thesis, California Institute of Technology (1971).
28. D. Faïman and A. W. Hendry, Phys. Rev. 180, 1572 (1969).
29. L. A. Copley, G. Karl and E. Obryk, Physics Letters 29B, 117 (1969).
30. A. Biètti, Phys. Rev. 142, 2158 (1966).
31. Aachen-Berlin-Bonn-Hamburg-Heidelberg-Munich Collaboration, Nucl. Phys. B8, 535 (1968).
32. See Reference 1, part II, p. 23.
33. W. G. Jones et al., Physics Letters 23, 597 (1966).
34. R. K. Logan and F. Uchiyama-Campbell, Phys. Rev. 149, 1220 (1966).
35. R. K. Logan and F. Uchiyama-Campbell, Phys. Rev. 153, 1634 (1967).
36. S. R. Deans and Wendell G. Holladay, Phys. Rev. 161, 1466 (1967).
37. S. Gasiorowicz, Elementary Particle Physics, John Wiley and Sons, Inc., New York (1966).

38. R. L. Anderson and R. Prepost, *Phys. Rev. Letters* 23, 46 (1969).
39. S. D. Ecklund and R. L. Walker, *Phys. Rev.* 159, 1195 (1967).
40. R. G. Moorhouse, *Phys. Rev. Letters* 16, 772 (1966).
41. A. Dar and V. F. Weisskopf, *Phys. Rev. Letters* 20, 762 (1968).
42. P. S. L. Booth et al., *Lett. al Nuovo Cim.* 2, 66 (1969).
43. C. A. Heusch, C. Y. Prescott and R. F. Dashen, *Phys Rev. Letters* 17, 1019 (1966).
44. G. F. Chew and H. W. Lewis, *Phys. Rev.* 84, 779 (1951).
45. F. T. Hadjioannou, *Phys. Rev.* 125, 1414 (1962).
46. G. F. Chew and G. C. Wick, *Phys. Rev.* 85, 636 (1952).
47. G. F. Chew and M. L. Goldberger, *Phys. Rev.* 87, 778 (1952).
48. P. Morse and H. Feshbach, Methods of Theoretical Physics, Part II, Chapter 9 (McGraw-Hill Book Company, Inc., 1953).
49. A. O. Barut, The Theory of the Scattering Matrix for the Interactions of Fundamental Particles, (The Macmillan Company, 1967).
50. G. F. Chew, F. E. Low, M. L. Goldberger and Y. Nambu, *Phys. Rev.* 106, 1345 (1957).
51. H. Cramer, Mathematical Methods of Statistics, Chapters 30 and 33, (Princeton University Press, 1961).
52. Reference 51, Chapter 27.
53. R. L. Walker, Single Pion Photoproduction in the Resonance Region, CALT-68-219 (Presented in 1969 at the International Symposium on Electron and Photon Interactions at High Energies).
54. L. S. Rochester, Ph.D. Thesis, California Institute of Technology (1968).