PHYSICAL REVIEW D VOLUME 20, NUMBER 11 1 DECEMBER 1979

Pion-nucleon partial-wave analysis

R. E. Cutkosky, R. E. Hendrick,* J. W. Alcock,† Y. A. Chao,‡ R. G. Lipes,§ and J. C. Sandusky‖
Physics Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

R. L. Kelly
Lawrence Berkeley Laboratory, Berkeley, California 94720
(Received 29 May 1979)

We have carried out a partial-wave analysis of amalgamated pion-nucleon scattering data at 35 momenta in the range 0.429–1.995 GeV/c. At each momentum, a large set of fits was obtained, with the amplitude represented as the sum of a fixed background and the fitted terms of an accelerated convergence expansion. The effects of discrete and continuum ambiguities were separated by special statistical and numerical techniques which sort the fits into clusters. Dispersion relations on selected curves within the kinematical region for physical scattering were used to resolve the ambiguities and to derive constraints. These constraints, which ensure that the fitted amplitudes have a relatively smooth energy dependence consistent with analyticity, were used iteratively to arrive at a unique set of amplitudes which fit the scattering data throughout the energy region studied.

I. INTRODUCTION

We report here results from a partial-wave analysis of πp scattering data for pion laboratory momenta between 0.42 and 2.0 GeV/c. The corresponding c.m. energy range is 1320 to 2160 MeV. Some partial results from an earlier stage of this analysis have been reported in previous publications.[1] This earlier work covered a more limited momentum range, 0.8 to 2.0 GeV/c. The present work also includes data which were not available at the time the initial stage was reported.

The purpose of our analysis is to derive information about the pion-nucleon scattering amplitude, and to make this information available for general use. It is necessary to have such information in order to develop and test theories of hadronic scattering. In particular, information about pion-nucleon resonances, as extracted from the partial-wave amplitudes, is an essential ingredient for tests of models of baryon structure.

Our analysis process involves several steps. First, amalgamated pion-nucleon scattering data were prepared in 3° angular intervals at 35 momenta, using techniques described in the preceding paper.[2] At each momentum, the amplitude was represented by an accelerated convergence expansion, and many randomly started fits were obtained, as explained in detail in Sec. II of this paper. Special techniques were used to sort the fits into clusters related by continuum ambiguities. Section III describes the method of hyperbolic dispersion relations (for forward and backward scattering and four interior curves) which were used to select the preferred fit at each momentum. These dispersion relations were also used to derive constraints which were fed back into the fitting at single momenta through an iterative process described in Sec. IV. Some general characteristics of our results are discussed in Sec. V of this paper.

Partial-wave amplitudes obtained from our analysis are presented in the following paper,[3] along with a discussion of some of the implications about resonances. A table of the amplitudes can be obtained on magnetic tape by contacting R.L.K.

The data used in this analysis came from well over 100 different pion-proton scattering experiments. The combination of information from many different sources representing a variety of experimental conditions is an important feature of our analysis. In the amalgamation stage, we need a dense and uniform distribution of data, even if some of them are of relatively low weight, in order to stabilize the interpolating surface. The comparison of data from different experiments, as described in the preceding paper, also allows us to identify discrepant data, and helps to eliminate biases coming from unknown systematic errors.

A complete listing of the input data for the amalgamation stage of our analysis will be included in the forthcoming catalog of the Particle Data Group's reaction-data base.[4] The older data in our collection were taken from previous compilations, particularly that of Lovelace *et al.*,[5] and were also checked against the original sources. Newer data were collected by us during the course of the analysis. A complete annotated bibliography for each experiment will be included with the data in Ref. 4. The data collection and bibliography are also available to interested parties on magnetic tape (contact R.L.K.). Important as the diversity of information represented by this data collection

is, we will have to be content here with a description of only a few of the largest and most significant recent experiments in the data set.

The largest elastic differential-cross-section experiment used was performed by the Bristol-Southampton-Rutherford collaboration at Rutherford Laboratory.[6] Data on π^+p and π^-p elastic scattering were taken at 51 momenta between 400 and 2150 MeV/c. The data are still somewhat preliminary, although it is expected that above 600 MeV/c only the normalization is likely to change significantly in the final results.[7] We allowed generous normalization errors for these data when carrying out the amalgamation, and used only data above 600 MeV/c. For π^+p elastic scattering there are two more quite large cross-section experiments, one from earlier work of the Bristol-Southampton-Rutherford collaboration between 800 and 1594 MeV/c,[8] and one from a Maryland-Argonne group between 1207 and 2300 MeV/c.[9] Turning to elastic polarization, the π^+p measurements are dominated by the Rutherford Laboratory experiment of Martin *et al.*[10] which measured the polarization at 68 momenta between 603 and 2651 MeV/c. π^-p elastic polarization has not been measured so exhaustively, and this remains the weakest part of the available elastic data. Three of the larger experiments are those of Refs. 11–13. The most important charge-exchange cross-section and polarization measurements are those of Rutherford Laboratory[14,15] and LBL.[16,17] The Rutherford measurements cover the range 617 to 2724 MeV/c with 22 momenta each for cross sections and polarizations, while LBL measured cross sections and polarizations at 6 and 5 momenta, respectively, in the range 1030 to 2390 MeV/c. Most of the momenta in each experiment are chosen to coincide for the two types of data.

The 35 laboratory momenta at which amalgamated data sets were prepared covered the range from 429 to 1995 MeV/c. There is approximately 50 MeV/c between successive momenta, but the momenta were chosen to minimize the need for interpolation in momentum during the amalgamation process, especially for the less copiously measured data types. As a result, the momenta are not distributed uniformly throughout the range.

II. ENERGY-INDEPENDENT ANALYSIS

In the first stage of our analysis we make single-energy χ^2 fits to the amalgamated data. The real and imaginary parts of each invariant amplitude are separately parametrized as the sum of a fixed "Born term" containing contributions from nearby singularities in the $\cos\theta$ plane and a polynomial term whose coefficients are the adjustable parameters used in the fitting.[18] The explicit form of each real or imaginary part is

$$F(x) = B(x) + R(x)P(z(x)), \tag{2.1}$$

where $x = \cos\theta$, B is the Born term, P is a polynomial in the variable $z(x)$ constructed according to the accelerated-convergence-expansion (ACE) prescription, and R is a fixed modulating factor. Our fitting procedure is a Monte Carlo method of minimization from randomly generated starting points to search for local minima of our χ^2 function, X^2, which includes not only the conventional χ^2 for the fit to the amalgamated data, but also additional contributions described in Sec. II D. When a sufficient number of these X^2 minima have been found and stored they are machine scanned to identify and combine clusters of statistically indistinguishable local minima. Accurate error matrices are formed for each cluster by combining the error matrices of the cluster's individual solutions with the intracluster dispersion. The clusters formed at each energy are the basic input for the subsequent energy-dependent stages of the analysis.

In this section we describe in detail the parametrization of the Born and polynomial terms in the amplitudes, our procedures for imposing unitarity constraints and for including Coulomb corrections, the Monte Carlo method of locating local minima, and our technique of cluster formation.

A. Born terms

The real and imaginary parts of the invariant amplitudes for πN scattering at fixed s are real-analytic functions of x with t- and u-channel cuts along the real axis, and with a nucleon pole in each B amplitude at $u = M^2$. The Born terms we use include the nucleon pole, the two-pion-exchange (TPE) contribution to the imaginary parts, and contributions from P, ρ, f, and Δ exchange. As described in Sec. II B, the fitted terms are analytic and convergent in enlarged regions corresponding to removal of both the nucleon pole and the discontinuities of the imaginary parts at small t ($4\mu^2 \leq t \lesssim 16\mu^2$) coming completely from TPE. The nucleon and TPE contributions are the most critical parts of the Born terms, while any inaccuracies in the remaining contributions can be more easily compensated by the fitted terms.

In practice, we have found that the quality of χ^2 fits is most sensitive to the nucleon term, which fortunately has a well-known coupling constant so that the pole can be included quite accurately.

The discontinuities of the imaginary parts for $4\mu^2 \leq t \lesssim 16\mu^2$ are not precisely known, but the TPE

contributions to the imaginary parts of high-angular-momentum partial waves calculated by Alcock and Cottingham[19] (AC) are certainly more reliable than any values that could be extracted directly from data by energy-independent fitting. The real parts, on the other hand, are less reliably calculated by AC, and are better determined by the data. We therefore want to constrain our amplitudes to be consistent with the AC imaginary parts for sufficiently large L, and it is convenient to do this by including them in the Born terms and deleting the corresponding TPE discontinuities from the fitted terms. This parametrization forces the imaginary parts of our partial waves to approach the AC values for very large L, and stabilizes the waves at smaller L against erratic behavior coming from statistical fluctuations in the data. In practice, the fitted term is always chosen to be sufficiently flexible so that these effects are confined to L values lying somewhat higher than those of the prominent peripheral resonances.[20]

With the exception of the TPE part, the Born terms are all constructed using simple Regge forms with absorptive corrections, and with some modification to improve their asymptotic behavior at large x. We use conventional trajectories, with residue parameters chosen to be qualitatively consistent with more detailed Regge models from high-energy fits and also chosen to give low-energy Born terms of reasonable magnitude. During the course of the analysis we have tried numerous variations on the Regge exchanges, and have found that the final partial-wave amplitudes are not very sensitive to the precise forms used, as long as the nucleon pole residue is handled correctly. The basic set of Regge exchanges used for our current results is given in Table I.

The amplitudes written in Table I all have essential singularities at $x=\infty$ which we will remove before proceeding. In fact, the existence of s-channel Regge poles implies that the physical amplitudes have a specific power-law dependence on x as $x\to\infty$, and we will later build this feature into the fitted terms by choosing appropriate modulating factors. First, however, we modify the amplitudes in Table I so that they become constant at large x. This is easily done by replacing the Mandelstam variables t and u everywhere by new variables

$$\hat{t}=tf(z_t),\quad z_t=(t_b-t)/s\,, \tag{2.2}$$

and

$$\hat{u}=M^2+(u-M^2)f(z_u),\quad z_u=(u_b-u)/s\,, \tag{2.3}$$

where t_b and u_b are the t- and u-channel branch-point locations and $f(z)$ is a function which is equal to unity at the origin, is approximately equal to unity for $|z|\lesssim 1$, and which behaves like z^{-1} for large z. To maintain the proper analytic structure of the full amplitude we choose $f(z)$ to be analytic in the entire z plane cut along the negative real axis. Note that t_b and u_b, and therefore $\hat{t}$ and $\hat{u}$, are different for the real and imaginary parts of the amplitudes in Table I. For the real parts, $t_b=4\mu^2$ and $u_b=(M+\mu)^2$. For the imaginary parts, $t_b\gtrsim 16\mu^2$ is the double-spectral-function boundary

TABLE I. Regge exchanges.

Exchange	Amplitude	Notation	
Pomeron	$B_P=(61.1i)e^{3t}$		
ρ	$A_\rho=-16\xi_\rho S_M^{\alpha_M}/D_M$ $B_\rho=70\xi_\rho S_M^{\alpha_M-1}/D_M$	$\xi_\rho=1-e^{-i\pi\alpha_M}$	$\alpha_M=0.5+t$ $S_M=1+\frac{s-(M+\mu)^2}{0.47}$
f	$A_f=-3.8\xi_f S_M^{\alpha_M}/D_M$ $B_f=-58\xi_f S_M^{\alpha_M-1}/D_M$	$\xi_f=1+e^{-i\pi\alpha_M}$	$D_M=(1-\alpha_M)(2-\alpha_M)$
N	$B_N=[6\pi\alpha_B' g^2(u/M^2)e^{0.67(u-M^2)}$ $+210(1-u/M^2)e^{0.67u}]\xi_N S_B^{\bar{\alpha}_N}/D_N$	$\bar{\alpha}_N=\alpha_B'(u-M^2)$ $\xi_N=1+e^{-i\pi\bar{\alpha}_N}$ $D_N=-\bar{\alpha}_N(1-\bar{\alpha}_N)$ $g^2=14.3$	$\alpha_B'=0.97$
Δ	$A_\Delta=(12.4+5.0u)e^u\xi_\Delta S_B^{\bar{\alpha}_\Delta}/D_\Delta$ $B_\Delta=(11.6-3.6u)e^u\xi_\Delta S_B^{\bar{\alpha}_\Delta}/D_\Delta$	$\bar{\alpha}_\Delta=-0.42+\alpha_B' u$ $\xi_\Delta=1-e^{-i\pi\bar{\alpha}_\Delta}$ $D_\Delta=-\bar{\alpha}_\Delta(1-\bar{\alpha}_\Delta)$	$S_B=s-M^2-\mu^2$

for four-pion exchange, and $u_b \gtrsim (M+\mu)^2$ is the double-spectral-function boundary for πN exchange. When using these new variables we also multiply $\mathrm{Re}B_N$ by $f[(u_b - M^2)/s]$ to maintain the correct residue at the nucleon pole.

The function $f(z)$ we use is

$$f(z) = \frac{1+2\sqrt{z}+3z/2}{(1+\sqrt{z}/2)^4} . \tag{2.4}$$

The coefficients have been chosen so that $f(z) \sim 1$ for $|z| \lesssim 1$. In particular,

$$\lim_{z\to 0} f(z) = 1 - \frac{z^{3/2}}{2} + O(z^2) , \qquad f(1) = \tfrac{8}{9} . \tag{2.5}$$

The function approaches $24/z$ as $z \to \infty$, and its only singularity other than the branch point at $z=0$ is a pole at $z=4$ on the second sheet.

Having introduced the new variables, we form s-channel isospin amplitudes from the exchange amplitudes in Table I, keeping Pomeron and Reggeon exchange separate for the present:

$$\begin{aligned} B_3^P &= B_1^P = B_P , \\ A_3^R &= -A_\rho + A_f + \tfrac{1}{3}A_\Delta , \\ B_3^R &= -B_\rho + B_f - \tfrac{2}{3}B_N - \tfrac{1}{3}B_\Delta , \\ A_1^R &= 2A_\rho + A_f + \tfrac{4}{3}A_\Delta , \\ B_1^R &= 2B_\rho + B_f + \tfrac{1}{3}B_N - \tfrac{4}{3}B_\Delta . \end{aligned} \tag{2.6}$$

Reggeon- and Pomeron-exchange partial-wave amplitudes, $b^R_{IJ\pm}$ and $ib^P_{IJ\pm}$, are now calculated by numerical integration. We note that TPE can proceed only by inelastic intermediate states, so that the AC imaginary parts $b^{AC}_{IJ\pm}$ are the "shadow of inelastic one-pion exchange (OPE)" in the same sense that the Pomeron is generally supposed to be the "shadow of diffractive production." Thus, we add the Pomeron and AC amplitudes to form purely imaginary absorptive amplitudes $ib^A_{IJ\pm}$:

$$ib^A_{IJ\pm} = ib^P_{IJ\pm} + ib^{AC}_{IJ\pm} . \tag{2.7}$$

For large J the amplitudes $ib^A_{IJ\pm}$ are dominated by the AC part and are relatively accurate, but for small J they may become unrealistically large, necessitating a cutoff. We use an energy-dependent cutoff function $iH(W)/2$, where H is chosen to approximate the average value of $\sigma_{\text{inelastic}}/\sigma_{\text{total}}$ for $\pi^\pm p$ scattering when W is in the vicinity of the $\pi\Delta$ and ρN thresholds and then to approach unity like $1 + aW^{-1}$ as $W \to \infty$. The function used is

$$H(W) = \begin{cases} 1.48(W-1.21)^2, & 1.21 < W < 1.54 \\ (W-1.4)/(W-0.65), & W > 1.54 . \end{cases} \tag{2.8}$$

We implement the cutoff smoothly in J by constructing new amplitudes $i\tilde{b}^A_{IJ\pm}$ such that $\tilde{b}^A_{IJ\pm}$ is bounded above by both $b^A_{IJ\pm}$ and by $H/2$, and such that $\tilde{b}^A_{IJ\pm} \simeq b^A_{IJ\pm}$ if $b^A_{IJ\pm} \ll H/2$. Using the fact that $b^A_{IJ\pm}$ is a monotonically decreasing function of J, we can accomplish this with the following recursive algorithm:

$$\begin{aligned} &\tilde{b}^A_{IJ\pm} = b^A_{IJ\pm}, \quad J \geq J_0 \gg 1 , \\ &\tilde{b}^A_{IJ\pm} = \tilde{b}^A_{I(J+1)\pm} + \frac{(b^A_{IJ\pm} - \tilde{b}^A_{I(J+1)\pm})(H/2 - \tilde{b}^A_{I(J+1)\pm})}{(b^A_{IJ\pm} - \tilde{b}^A_{I(J+1)\pm}) + (H/2 - \tilde{b}^A_{I(J+1)\pm})} , \\ &\qquad J < J_0 . \end{aligned} \tag{2.9}$$

The Reggeon-exchange and absorptive amplitudes are now combined by the simple absorptive prescription of multiplying together the corresponding S-matrix elements. The full Born term T-matrix elements are then

$$b_{IJ\pm} = b^R_{IJ\pm}(1 - 2\tilde{b}^A_{IJ\pm}) + i\tilde{b}^A_{IJ\pm} . \tag{2.10}$$

B. Fitted terms and unitarization

The fitted term for each real or imaginary part is of the form $R(x)P(z(x))$, where R is a modulating factor and P is a polynomial with adjustable coefficients. The variable $z(x)$ is chosen according to the prescription of Cutkosky and Deo[18]; $z(x)$ is analytic in the cut x plane and maps the cut x plane onto the interior of a unifocal ellipse, the physical region being mapped onto $-1 \leq z \leq 1$. This mapping extends the formal domain of convergence of our parametrization into the full cut x plane. As discussed previously, the x-plane branch points are taken to lie at $t = 4\mu^2$ and $u = (M+\mu)^2$ for the real parts, and at the 4π-exchange and πN-exchange double-spectral-function boundaries for the imaginary parts. We refer the reader to Ref. 18 for a precise mathematical definition of $z(x)$ and for a description of efficient numerical techniques for the calculation of $z(x)$.

The polynomial coefficients of $P(z)$ are not suitable for direct fitting to the data because this would allow no simple way to impose unitarity constraints. Instead, we formulate the fitting procedure in terms of "low" and "high" partial waves, where the low waves vary freely, subject only to unitarity constraints, and the form of the parametrization is used to generate a linearly related set of high partial waves by extrapolation in J. This procedure was suggested by Cutkosky and Deo, and was described explicitly by them for the spinless case. In this section we give the details of the extrapolation procedure for meson-nucleon scattering. It will be convenient to formulate the extrapolation procedure in terms of the amplitudes f_1 and f_2, because of the simple relation between their Legendre projections and the physical partial-wave amplitudes, Eq. (A20). In most of this section we temporarily suppress isospin indices;

the same extrapolation procedure is used for $I=\frac{1}{2}$ and $I=\frac{3}{2}$.

Before discussing the extrapolation procedure itself, we will specify the modulating factor R which multiplies each polynomial. We use R to take into account, at least roughly, the influence of distant singularities on the general shapes of the amplitudes. R is naturally parametrized in terms of the s-channel Regge behavior of the amplitudes as $x\to\infty$. Examination of the partial-wave expansions of f_1 and f_2, Eq. (A18), shows that at large x, $f_1\sim x^{\alpha_+-1/2}$ and $f_2\sim x^{\alpha_--1/2}$ where $\alpha_\pm$ are the leading s-channel $J=L\pm\frac{1}{2}$ trajectories. For $I=\frac{3}{2}$, $\alpha_+=\alpha_\Delta$ and $\alpha_-\simeq\alpha_\Delta-1$ while for $I=\frac{1}{2}$, $\alpha_+\simeq\alpha_-=\alpha_N$. Thus, there is a variation of about 1 in the asymptotic exponents of x for the various amplitudes. However, the characteristic energy dependence of the exponents is more important for our purposes than this small variation in the absolute level. So we use essentially the same modulating factor for all the amplitudes and choose it to grow like $x^{\bar\alpha}$ where $\bar\alpha\simeq\alpha_N-\frac{1}{2}$. (There is a small difference between the modulating factors used for real and imaginary parts.) A further clue to the parametrization of R is provided by the behavior of $P'_{\alpha+1}(x)$ at nonasymptotic values of x. For moderately large $\bar\alpha$, appreciable deviations from the asymptotic $x^{\bar\alpha}$ behavior develop when $|x|\simeq\sqrt{\bar\alpha}/2$. This gives a natural distance scale in the x plane for appreciable variation to develop in R. Finally, the general form of the parametrization $B+RP$ requires that R be real analytic and nonvanishing in the cut x plane. A simple form embodying all these features is

$$R^p(x)=\left[1+\left(\frac{(x_t^p-x)(x-x_u^p)}{\bar\alpha/4}\right)^{1/2}\right]^{\bar\alpha}, \tag{2.11}$$

where $\bar\alpha=s-1$, x_t^p and x_u^p are the appropriate t- and u-channel branch points in the x plane, and p is an index to distinguish the real ($p=1$) and imaginary ($p=2$) part modulating factors.

Having chosen the modulating factors, the extrapolation procedure is now developed as follows. We define

$$\tilde f_j(x)=f_j(x)-f_{j,\mathrm{Born}}(x),\quad j=1,2 \tag{2.12}$$

where $f_{j,\mathrm{Born}}$ is constructed from the Born-term partial waves given in Eq. (2.10); $\tilde f_j$ is parametrized as

$$\tilde f_j(x)=\sum_{k=0}^{n_j-1}\left[a_{jk}^1\mathcal{T}_k^1(x)+ia_{jk}^2\mathcal{T}_k^2(x)\right], \tag{2.13}$$

where the a_{jk}^p are real coefficients,

$$\mathcal{T}_k^p(x)=R^p(x)T_k(z^p), \tag{2.14}$$

and $T_k(z^p)$ is a kth-order polynomial in $z^p(x)$. The specific choice of the set of polynomials $\{T_k\}$ can be left open for the present; it is irrelevant to the extrapolation procedure. The partial-wave amplitudes following from this parametrization are [see Eq. (A20)]

$$f_{Js}=b_{Js}+\tilde f_{Js}, \tag{2.15}$$

where

$$\tilde f_{J\pm}=q\sum_{k=0}^{n_1-1}\left(a_{1k}^1Z^1_{J\pm1/2,k}+ia_{1k}^2Z^2_{J\pm1/2,k}\right) +q\sum_{k=0}^{n_2-1}\left(a_{2k}^1Z^1_{J\mp1/2,k}+ia_{2k}^2Z^2_{J\mp1/2,k}\right) \tag{2.16}$$

and

$$Z^p_{lk}=\tfrac{1}{2}\int_{-1}^{1}dx\,P_l(x)T_k^p(x). \tag{2.17}$$

For clarity we will temporarily index the $\tilde f_{Js}$ amplitudes in serial order as

$$\left.\begin{aligned}\tilde f_{2J}&\equiv\tilde f_{J-}\\ \tilde f_{2J+1}&\equiv\tilde f_{J+}\end{aligned}\right\},\quad J=\tfrac{1}{2},\tfrac{3}{2},\ldots, \tag{2.18}$$

and similarly for the coefficients a_{ik}^p,

$$\begin{aligned}a_{2k+1}^p&\equiv a_{1k}^p,\quad k=0,1,\ldots,n_1,\\ a_{2k+2}^p&\equiv a_{2k}^p,\quad k=0,1,\ldots,n_2.\end{aligned} \tag{2.19}$$

In practice, it is convenient when varying the number of parameters to add or delete parameters alternately from f_1 and f_2 so that n_1-n_2 is always 0 or 1. We will specialize to this case so that the index of a_m^p runs from 1 to $n=n_1+n_2$. The real or imaginary part of Eq. (2.16) can now be rewritten in the compact form

$$\tilde f_m^p=q\sum_{m'=1}^{n}B_{mm'}^pa_{m'}^p, \tag{2.20}$$

where

$$\begin{aligned}B^p_{2J,2k+1}&=Z^p_{J-1/2,k},\\ B^p_{2J+1,2k+1}&=Z^p_{J+1/2,k},\\ B^p_{2J,2k+2}&=Z^p_{J+1/2,k},\\ B^p_{2J+1,2k+2}&=Z^p_{J-1/2,k}.\end{aligned} \tag{2.21}$$

To eliminate the coefficients a_m^p from Eq. (2.20) we define S^p to be the submatrix of B^p corresponding to its first n rows

$$S^p_{mm'}=B^p_{mm'},\quad m,m'=1,2,\ldots,n \tag{2.22}$$

so that

$$a_m^p=\frac{1}{q}\sum_{m'=1}^{n}(S^p)^{-1}_{mm'}\tilde f^p_{m'} \tag{2.23}$$

and

$$\tilde f_m^p=\sum_{m'=1}^{n}C^p_{mm'}\tilde f^p_{m'}, \tag{2.24}$$

where

$$C^{\flat}_{mm'} = \sum_{m''=1}^{n} B^{\flat}_{mm''}(S^{\flat})^{-1}_{m''m'} . \tag{2.25}$$

Equations (2.24) and (2.25) define the extrapolation procedure. The coefficients $a^{\flat}_{jk}$ have been replaced by a set of relations between the low ($m' \leq n$) and high ($m > n$) partial-wave amplitudes. For $m \leq n$, $C^{\flat}_{mm'} = \delta_{mm'}$ and Eq. (2.24) reduces to an identity. We will refer to $C^{1}_{mm'}$ and $C^{2}_{mm'}$ as "extrapolation matrices," and in the rest of this paper will go back to using the usual Js form for the indices.

The low waves can now be parametrized in explicitly unitary form, but there is no guarantee that the extrapolated high waves will satisfy unitarity. We include several safeguards in our analysis to eliminate this potential problem. We always use enough low waves so that all the high waves correspond to impact parameters lying somewhat outside the peripheral region. The high waves are thus fairly small, and the main effect of unitarity is just to require that they have positive imaginary parts. Our parametrization is biased towards having small positive imaginary parts for large J, by the way in which we incorporate the AC contributions. In spite of this, small deviations from unitarity do occur in some fits if the high waves are completely unconstrained. To provide a constraint we include a term

$$U = C \sum_{IJs}^{\text{high}} (J+1/2)(\eta_{IJs}{}^{2}-1)^{2}\theta(\eta_{IJs}{}^{2}-1) \tag{2.26}$$

in our χ^2 function. The sum runs over high waves of both isospins, and the constant C is adjusted so that U is typically $\lesssim 1$. This has turned out to be quite an effective constraint on our analysis; there are no statistically significant violations of unitarity in the high waves when U is included in X^2. As a final safeguard, however, at the very end of the fitting and solution-sorting procedure, when a reliable covariance matrix has been constructed, a small adjustment is made to the low partial waves to guarantee that the resulting high waves are explicitly unitary.

In some stages of the fitting procedure it is convenient to be able to reconstruct the full scattering amplitudes directly from the low waves without extrapolating in J or reconstructing the $a^{\flat}_{jk}$ parameters. For this we use the f and g amplitudes because they are particularly suitable for incorporating electromagnetic corrections (see Sec. II C). Omitting electromagnetic corrections for the present, we can write out Eq. (2.15) for the real and imaginary parts of the complete partial-wave amplitude for each Js as

$$f^{\flat}_{Js} = b^{\flat}_{Js} + \sum_{J's'}^{\text{low}} C^{\flat}_{Js,J's'}(f^{\flat}_{J's'} - b^{\flat}_{J's'}) . \tag{2.27}$$

Substituting in Eq. (A17) for f and g and rearranging terms, one obtains

$$\begin{aligned} f^{\flat}(x) &= B^{\flat}_{f}(x) + \frac{1}{q}\sum_{Js}^{\text{low}} f^{\flat}_{Js} Q^{\flat}_{f,Js}(x) , \\ g^{\flat}(x) &= B^{\flat}_{g}(x) + \frac{\sin\theta}{q}\sum_{Js}^{\text{low}} f^{\flat}_{Js} Q^{\flat}_{g,Js}(x) , \end{aligned} \tag{2.28}$$

where

$$\begin{aligned} B^{\flat}_{f} &= f^{\flat}_{\text{Born}} - \frac{1}{q}\sum_{Js}^{\text{low}} b^{\flat}_{Js} Q^{\flat}_{f,Js} , \\ B^{\flat}_{g} &= g^{\flat}_{\text{Born}} - \frac{\sin\theta}{q}\sum_{Js}^{\text{low}} b^{\flat}_{Js} Q^{\flat}_{g,Js} . \end{aligned} \tag{2.29}$$

Here f_{Born} and g_{Born} are the pure Born-term f and g amplitudes and

$$\begin{aligned} Q^{\flat}_{f,Js} = \sum_{J'} (J'+\tfrac{1}{2})(&C^{\flat}_{J'-,Js} P_{J'-1/2} \\ &+ C^{\flat}_{J'+,Js} P_{J'+1/2}) , \end{aligned} \tag{2.30}$$

$$Q^{\flat}_{g,Js} = \sum_{J'} (C^{\flat}_{J'-,Js} P'_{J'-1/2} - C^{\flat}_{J'+,Js} P'_{J'+1/2}) .$$

Note that the sums in the Q functions run over all $J' \geq \frac{1}{2}$.

C. Electromagnetic corrections

The f and g amplitudes from which we actually construct observables include electromagnetic (em) corrections in the standard way:

$$\begin{aligned} f &= f_{\text{em}} + f_{N} , \\ g &= g_{\text{em}} + g_{N} , \end{aligned} \tag{2.31}$$

where f_{em} and g_{em} are purely em amplitudes, and

$$\begin{aligned} f_{N} = \frac{1}{q}\sum_{J} (J+\tfrac{1}{2})(&e^{2i\Sigma_J} f_{J-} P_{J-1/2} \\ &+ e^{2i\Sigma_J} f_{J+} P_{J+1/2}) , \end{aligned} \tag{2.32}$$

$$g_{N} = \frac{\sin\theta}{q}\sum_{J} (e^{2i\Sigma_J} f_{J-} P'_{J-1/2} - e^{2i\Sigma_J} f_{J+} P'_{J+1/2}) .$$

Here f_{Js} is the nuclear amplitude for wave Js in the channel under consideration and Σ_{Js} is the em phase shift. The em corrections are different for each channel; for definiteness we will describe the treatment of $\pi^{+}p$ elastic scattering in detail, and then give the modifications necessary for $\pi^{-}p$ elastic and charge-exchange (CEX) scattering.

We use the em amplitudes of Tromborg *et al.*,[21]

$$\begin{aligned} f_{\text{em}} &= f^{\gamma} F \exp(2i\sigma_0 + i\theta + i\Xi) , \\ g_{\text{em}} &= g^{\gamma} F \exp(2i\sigma_0) . \end{aligned} \tag{2.33}$$

f^γ and g^γ are relativistic single-photon-exchange amplitudes constructed from the invariant amplitudes[21,22]

$$A^\gamma = 2\pi\alpha(\mu_p - 1)\left(\frac{E_{\rm lab}}{q^2\sin^2(\theta/2)} - \frac{1}{M}\right),$$
$$B^\gamma = \frac{-2\pi\alpha\mu_p}{q^2\sin^2(\theta/2)}, \tag{2.34}$$

where α is the fine-structure constant and $\mu_p = 2.793$ is the proton magnetic moment. F is the product of the pion[23] and proton form factors which we parametrize as

$$F = (1 - t/0.471)^{-1}(1 - t/0.71)^{-2}, \tag{2.35}$$

where t is in GeV^2. The various contributions to the phases are

$$\sigma_0 = \arg\Gamma(1 + i\eta),$$
$$\theta = -\eta\ln[\sin^2(\theta/2)], \tag{2.36}$$
$$\Xi = \eta\int_{-4q^2}^{0}(dt/t)[1 - F(t)],$$

where $\eta = \alpha/v_{\rm lab}$ and $v_{\rm lab}$ is the laboratory velocity of the pion. We use an approximation for Ξ obtained by carrying out the integral with F set to $(1 - t/\bar{m}^2)^{-3}$ and with $\bar{m}^2$ chosen so that the approximation has the correct asymptotic form as $q^2 \to \infty$. This gives

$$\Xi \approx -\eta(\ln\xi + \tfrac{3}{2} - 1/\xi - 1/2\xi^2), \tag{2.37}$$

where $\xi = 1 + 4q^2/\bar{m}^2$ and $\bar{m}^2 = 0.619$. In fact, this is also a good approximation for small q^2; if we were to require instead that the approximation become exact as $q^2 \to 0$ it would only be necessary to change $\bar{m}^2$ to 0.607.

We can now obtain the em phase shifts Σ_{Js} by numerical partial-wave projection of $f_{\rm em}$ and $g_{\rm em}$. This cannot be done by direct numerical integration because of the singular behavior of $f_{\rm em}$ at 0°. To handle this problem we introduce the nonrelativistic Coulomb amplitude

$$f_C = -\eta[2q\sin^2(\theta/2)]^{-1}\exp(2i\sigma_0 + i\theta) \tag{2.38}$$

and rewrite $f_{\rm em}$ as

$$f_{\rm em} = f_C + \tilde{f}_{\rm em}. \tag{2.39}$$

The partial-wave amplitudes corresponding to f_C are well known, and the residual amplitude $\tilde{f}_{\rm em}$ is sufficiently well behaved at 0° to allow numerical integration. Our prescription for calculating Σ_{Js} is thus to calculate numerically the partial-wave amplitude Δ_{Js} corresponding to $\tilde{f}_{\rm em}$ and $g_{\rm em}$, add this to the Coulomb amplitude corresponding to f_C, and then set Σ_{Js} equal to half the phase of the resulting S-matrix element, i.e.,

$$\Sigma_{J\pm} = \tfrac{1}{2}\arctan\left(\frac{\sin(2\sigma_{J\pm1/2}) + 2\,\mathrm{Re}\,\Delta_{J\pm}}{\cos(2\sigma_{J\pm1/2}) - 2\,\mathrm{Im}\,\Delta_{J\pm}}\right), \tag{2.40}$$

where $\sigma_{J\pm1/2} = \arg\Gamma(J \pm \frac{1}{2} + 1 + i\eta)$.

It is now straightforward to reexpress f and g completely in terms of low partial waves, as was done in Sec. II B in the absence of em corrections. One obtains the expressions[24]

$$f^p(x) = B_f^p(x) + \frac{1}{q}\sum_{Js}^{\rm low}[f_{Js}^p Q_{f,Js}^p(x) + \xi_p f_{Js}^{\bar{p}} P_{f,Js}^{\bar{p}}(x)],$$
$$g^p(x) = B_g^p(x) + \frac{\sin\theta}{q}\sum_{Js}^{\rm low}[f_{Js}^p Q_{g,Js}^p(x) + \xi_p f_{Js}^{\bar{p}} P_{g,Js}^{\bar{p}}(x)], \tag{2.41}$$

where $\xi_1 = -1$, $\xi_2 = 1$, $\bar{p} = 3 - p$ (i.e., $\bar{p}$ and p correspond to opposite parts), and

$$B_f^p = f_{\rm em}^p + f_{N,\rm Born}^p - \frac{1}{q}\sum_{Js}^{\rm low}(b_{Js}^p Q_{f,Js}^p + \xi_p b_{Js}^{\bar{p}} P_{f,Js}^{\bar{p}}),$$
$$B_g^p = g_{\rm em}^p + g_{N,\rm Born}^p - \frac{\sin\theta}{q}\sum_{Js}^{\rm low}(b_{Js}^p Q_{g,Js}^p + \xi_p b_{Js}^{\bar{p}} P_{g,Js}^{\bar{p}}), \tag{2.42}$$

$$Q_{f,Js}^p = \sum_{J'}(J' + \tfrac{1}{2})[\cos(2\Sigma_{J'-})C_{J'-,Js}^p P_{J'-1/2} + \cos(2\Sigma_{J'+})C_{J'+,Js}^p P_{J'+1/2}],$$
$$P_{f,Js}^p = \sum_{J'}(J' + \tfrac{1}{2})[\sin(2\Sigma_{J'-})C_{J'-,Js}^p P_{J'-1/2} + \sin(2\Sigma_{J'+})C_{J'+,Js}^p P_{J'+1/2}], \tag{2.43}$$

$$Q_{g,Js}^p = \sum_{J'}[\cos(2\Sigma_{J'-})C_{J'-,Js}^p P'_{J'-1/2} - \cos(2\Sigma_{J'+})C_{J'+,Js}^p P'_{J'+1/2}],$$
$$P_{g,Js}^p = \sum_{J'}[\sin(2\Sigma_{J'-})C_{J'-,Js}^p P'_{J'-1/2} - \sin(2\Sigma_{J'+})C_{J'+,Js}^p P'_{J'+1/2}].$$

Equations (2.41)–(2.43) apply to π^+p elastic scattering only. The corresponding equations for π^-p scattering are obtained by making the substitutions

$$f_{\rm em} \to -f_{\rm em}^*,$$
$$g_{\rm em} \to -g_{\rm em}^*, \tag{2.44}$$
$$\Sigma_{Js} \to -\Sigma_{Js}.$$

Since Σ_{Js} is an odd function of the pion charge, we can use the same Q and P functions for π^-p elastic scattering as for π^+p, except that we must change the sign of the P functions. For π^-p CEX scattering we make a rough approximation which is justified by the absence of Coulomb-nuclear interference in CEX, and by the fact that the statistical accuracy of CEX data is generally considerably less than that of elastic data. We use $-\Sigma_{Js}/2$ for the CEX em phase, and only require

that the corresponding Q and P functions be correct to first order in $\Sigma_{J\delta}$. Thus we can use the same Q functions, and can use P functions which differ from the above only by having an additional overall factor of $-\frac{1}{2}$.

D. Single-energy fits

The parametrization described above has been used to fit data by searching for local minima of the χ^2 function

$$X^2=\chi^2+\Phi+U+D\,, \tag{2.45}$$

where χ^2 is the contribution from experimental data, Φ is a truncation function, U is the unitarity constraint defined in Eq. (2.26), and D contains constraints from the hyperbolic-dispersion-relation fits described in the following sections. In this section we describe further the χ^2 and Φ contributions to X^2, and our Monte Carlo searching procedure.

In χ^2 we include elastic and CEX differential-cross-section (DCS) and polarization (POL) data, $\pi^{\pm}p$ total-cross-section data, and the real parts of the forward elastic amplitudes as determined from dispersion relations. We use DCS and POL data which have been amalgamated[2] at the momentum of the fit. As described in Ref. 2, the amalgamation procedure produces DCS and POL data for all three reactions which have been interpolated to the same energy and angles, which include corrections for systematic effects, and which have had the errors of discrepant data enlarged. Unlike ordinary experimental data, the amalgamated data are highly correlated and it is necessary to take this into account in χ^2 using the method of "correlation vectors."[2]

χ^2 also includes constraints on the forward elastic amplitudes from the total-cross-section data of Carter *et al.*[25] interpolated to the momentum of the fit, and from the forward real parts of Engelmann and Hendrick[26] calculated at the momentum of the fit. The question of what em corrections to apply to $f(0^\circ)$ when comparing to $\sigma_{\rm tot}$ and the forward real-part calculations is somewhat ambiguous. As an example, consider a case where elastic unitarity holds and the total cross section is measured in such a way that $\sigma_{\rm tot}=\int d\Omega(|f_N|^2+|g_N|^2)$. It is then easy to see that $\sigma_{\rm tot}=(4\pi/q)\mathrm{Im}f_0(0^\circ)$ where f_0 has no em corrections at all. We have followed the suggestion of Tromborg *et al.*[21] and used this prescription throughout the analysis. We also use $\mathrm{Re}f_0(0^\circ)$ to compare with the forward real parts of Engelmann and Hendrick.

At 687 MeV/c, χ^2 includes constraints on the f amplitude obtained from the shape of the cusp, associated with the ηn threshold, in the $\pi^- p$ differential cross section. We have determined Ref and Imf at 15 angles by analysis of preliminary data from Binnie and Sarma.[27] We also use the results of Bhandari and Chao[28] for the phase of the ηn production term in the S_{11} partial-wave amplitude and for the phase of the f amplitude at 180°, which they obtained by analysis of the data of Debenham *et al.*[29]

The truncation function (TF) Φ is used to effect a smooth cutoff in the number of free parameters. We wish to use a conservatively large number of parameters, but to avoid the erratic behavior that could result from fitting high-order polynomials with no *a priori* constraints on the coefficients. To accomplish this we use a TF of the general form

$$\Phi=\sum_{jI\flat}K^{\flat}_{jI}\oint_{E^{\flat}}|dz|W(z)\left(\frac{\tilde f^{\flat}_{jI}(x^{\flat}(z))}{R^{\flat}(x^{\flat}(z))}\right)^2, \tag{2.46}$$

where $\tilde f^{\flat}_{jI}$ is defined in Eq. (2.12). The integral runs around the periphery of the unifocal ellipse $E^{\flat}$ onto which the cuts of $\tilde f^{\flat}_{jI}$ are mapped by $z^{\flat}(x)$. Since $|z^{\flat}|>1$ on $E^{\flat}$, high powers of $z^{\flat}$ in $\tilde f^{\flat}_{jI}/R^{\flat}$ contribute more to Φ than low powers. The coefficients $K^{\flat}_{jI}$ are adjusted so that the contribution of $(z^{\flat})^n$ to Φ becomes appreciable for n in the range where it is desired to effect a smooth cutoff.

For convenience we choose the weight function $W(z)$ to be

$$W(z)=\frac{2}{\pi}|z^2-1|^{-1/2} \tag{2.47}$$

and the polynomials introduced in Eq. (2.14) to be Chebyshev polynomials. This simplifies Φ because with this weight function the Chebyshev polynomials are orthogonal on a unifocal ellipse of arbitrary size:

$$\frac{2}{\pi}\oint_E|dz|\,|z^2-1|^{-1/2}T_k(z)T_l(z)=\delta_{kl}N_k\,, \tag{2.48}$$

where

$$N_k=2\delta_{k0}+R^{2k}+R^{-2k} \tag{2.49}$$

and R is the sum of the semiaxes of the ellipse E. Thus, Φ reduces to

$$\Phi=\sum_{jI\flat k}K^{\flat}_{jI}N^{\flat}_k|a^{\flat}_{jIk}|^2\,. \tag{2.50}$$

For fixed $k>0$, $N^{\flat}_k$ decreases with increasing energy (the ellipse shrinks) so there is a natural tendency for Φ to allow more parameters to vary freely as the energy increases. The rate of increase in the number of free parameters could be further adjusted by allowing the scale parameters $K^{\flat}_{jI}$ to be energy dependent, but we have not found

this to be necessary. We have used $K_{jI}^1 = 0.05$ and $K_{jI}^2 = 0.02$ throughout, and find that this gives a satisfactory cutoff over our entire momentum range.

It is useful for later applications to have a quantitative estimate at each local minimum of X^2 of the effective number of parameters which are actually used in fitting the data and the physical constraints contained in U and D, i.e., the effective number of free parameters which we call ν. For purely quadratic fitting of the type considered by Cutkosky[30] (and used in the hyperbolic-dispersion-relation fitting described in Section III and Appendix B), it is easy to show that the contributions of the individual coefficients to the TF tend to peak in the vicinity of the cutoff. The coefficient which gives the maximum contribution to the TF is about equally constrained by the TF itself and by the other constraints in X^2. A similar behavior is observed in our energy-independent fits, and we use this behavior to estimate the effective number of free parameters. At each local minimum, we find the eight indices $k = \kappa_{jI}^p$ for which $N_k^p |a_{jIk}^p|^2$ takes its maximum values (with fixed values of p, j, and I). Since the cutoff generally extends over a range of k and can be continuously adjusted by adjusting the scale parameters, it is more appropriate to think of k as a continuous variable than as a discrete variable. We therefore improve the estimates of κ_{jI}^p by fitting quadratic functions of k to the values of $N_k^p |a_{jIk}^p|^2$ near $k = \kappa_{jI}^p$ and finding the points at which these functions assume their maxima, say $k = \bar{\kappa}_{jI}^p$. These are to be identified as the values of k at which Φ and the rest of X^2 constrain a_{jIk}^p with equal strength. An equally constrained parameter of this sort we count as being $\frac{1}{2}$ of an effective free parameter. In general, some parameters with smaller values of k will also be partially constrained by Φ, and some with larger values of k will be partially constrained by the rest of X^2. We assume that the cutoff is sufficiently symmetric so that these effects approximately cancel in the count of effective free parameters and are then left with the estimate

$$\nu = \sum_{pjI} (\bar{\kappa}_{jI}^p + \tfrac{1}{2}) . \tag{2.51}$$

The effective number of degrees of freedom referred to later is equal to the number of data points minus ν.

In practice, ν varies from about 25 to 45 over the momentum range considered here. The k range of the smooth cutoff also varies with momentum, being sharper at low momentum where R is large [cf. Eq. (2.49)]. At our lowest momenta the cutoff is essentially the same as a sharp truncation, while at our highest momenta it extends over a range of about three partial waves of each isospin.

In the initial single-energy fitting with $D = 0$, a Monte Carlo method was used to find numerous local minima of X^2 in parameter space. Referring to the results of previous analyses[31,32] and to our own results at neighboring energies, we defined a region in parameter space corresponding to large areas in the unitarity circle of each low partial wave where solutions could reasonably be expected to be found. Random starting points were generated within this region, and the CERN minimizer MIGRAD[33] was used to move from each starting point to a local minimum. In our use of MIGRAD we supplied the minimizer with analytically calculated derivatives of X^2 with respect to the search parameters, and MIGRAD returned an estimate of the covariance matrix of the search parameters at each local minimum using the variable metric method. If local minima lying outside the search region were persistently found, the region was enlarged and/or moved to include the location of the new solutions in subsequent runs. At energies where the amount of experimental data was particularly small and the number of ambiguous solutions correspondingly large, it was useful to also use local minima from a more highly constrained neighboring energy as starting points. This increased the chances of finding the correct solution at the poorly constrained energy. The searching was continued until we stopped finding significant new solutions, and the number of solutions obtained was large enough for carrying out the cluster recognition and construction procedures described in Sec. II E. This required locating 60 to 100 local minima at each energy.

These original single-energy fits initiated an iterative procedure described later, in which consistency between the experimental data at a given energy and the dispersion-relation constraints contained in D was achieved. Modifications made in subsequent rounds of single-energy fitting included, beside the iterative adjustment of the dispersion relation constraints, the addition of new experimental data as they became available and refinement of the Born terms, em corrections, etc. In most of the subsequent runs the local minima from the original runs were used as starting points, in addition to a smaller amount of running with randomly generated starting points.

E. Ambiguities and cluster formation

The final stage of the energy-independent part

of our program is an analysis of the distribution of local minima of X^2. It is convenient to think of the local minima at a particular energy as points in the multidimensional space of search parameters (low partial waves). Discrete ambiguities are characterized by well-separated clusters of points with similar values of X^2. The size and shape of each cluster is influenced by continuum (overall phase and relative phase) ambiguities. Some degree of ambiguity is always present, but the seriousness of this problem depends greatly on the quantity and quality of the available data. Unitarity is also effective in reducing the degree of ambiguity, as are the analyticity constraints implicit in our parametrization of the amplitudes. The remaining ambiguities are the topic of this subsection in which we describe our techniques for identifying distinct clusters of local minima, and for determining their size and shape. These clusters are used as the input data for the dispersion-relation fitting programs described later. Although the cluster analysis was carried out both for the initial fits with $D=0$ and for the subsequent fits with dispersion-relation constraints, most of the serious discrete ambiguities were resolved when the constraints were applied. Thus, the clustering techniques are really tailored to the $D=0$ situation, where we sometimes found as many as 20 clusters, rather than the constrained situation where we seldom found more than one cluster.

At the beginning of a cluster scan we are presented with a set of individual points (local minima) f^n_{pJs} in an N-dimensional parameter space. With each point there are associated values of X^2, χ^2, etc., and a covariance matrix $U^n_{pJs,p'J's'}$ obtained from the minimizer. Before initiating the scan we make three modifications to these matrices. The matrix U^n returned by the minimizer is an approximation to the standard expression for the covariance matrix corresponding to a local minimum of a quadratic χ^2 function, i.e., an approximation to twice the inverse of the second-derivative matrix of X^2. If all the approximations involved are accurate, and X^2 is really locally quadratic, then U^n corresponds to a unit increment in X^2. We wish to be more conservative than this, and to enlarge U^n by an overall scale factor so that it will agree better with realistic cluster sizes as observed in distributions of local minima. For a quadratic χ^2 function this would correspond to scaling up the standard increment by the same factor, and it is useful for later discussion to denote the scale factor as δX^2. The scale factor is chosen empirically and for most of the fitting described here we have used $\delta X^2=5$. The second modification of U^n is to multiply by another scale factor equal to χ^2 divided by the effective number of degrees of freedom, if this quantity is greater than unity. This is the same prescription which has long been used by the Particle Data Group[34] to allow for biases due to inconsistencies among the data used in a fit. These scale factors are typically 1.4 or less, and in the worst cases may be as large as 2.0. The last modification is to replace the smallest $N/10$ diagonal matrix elements of U^n by the next largest one. This is done because the covariance matrix obtained from the minimizer for a particular local minimum depends somewhat upon the details of the search in which the minimum was found. Matrix elements may fluctuate by as much as a factor of 2 among several different searches all of which reach the same minimum. (In fact, one of the main points of cluster formation is to use a large sample of local minima to build up more accurate covariance matrices than those obtained from the minimizer.) The above replacement is generally confined to small waves with fairly high J, and reduces the chances that anomalously small diagonal matrix elements will cause two local minima to appear to be statistically distinguishable when they should actually be assigned to the same cluster. In the following we continue to use the notation U^n with the understanding that it now denotes the modified covariance matrix.

These modifications of U^n exemplify a general feature of our approach to error propagation in this analysis. Modern πN data is sufficiently precise and abundant that the purely statistical errors in a single fit are usually quite small. Systematic errors of various kinds are larger and tend to dominate the true uncertainties in the results. We take systematic uncertainties into account as best we can, using the experience we have built up by observing the results of many fits, both energy independent and energy dependent. However, quantitative estimates, such as the exact size of δX^2 or the exact number of small matrix elements of U^n which are enlarged, are necessarily somewhat subjective. In these estimates we have attempted to be rather conservative and to avoid underestimating the magnitudes of any significant sources of systematic error. Our aim is to produce final covariance matrices that are realistic estimates of the true uncertainties in the determination of the πN amplitudes from existing data, although there will certainly be some instances in which our error estimates fall short of this goal. In particular, we expect that there are significant non-Gaussian tails associated with our error estimates. It is commonly observed in distributions of experimental data that fluctuations of several standard deviations

occur more frequently than would be expected on the basis of Gaussian statistics. We suspect that our amplitudes and quantities derived from them will also have this property, and that it will be particularly severe for quantities corresponding to directions in parameter space which are poorly constrained by the input data (such as the overall phase, spin rotation parameters, etc.).

Proceeding now to the main stage of the cluster analysis, our goal is to replace the relatively large number of individual local minima by a smaller number of clusters with central values and covariance matrices which accurately describe their position, size, and shape. This is done by scanning over the points and building up the clusters during the scan. The scan is done in two passes, the first of which builds up the central part of each cluster and the second of which builds up the "tails." For clarity of exposition we describe here a simplified procedure in which the points are scanned once, assigned to clusters, and then the cluster characteristics are calculated after the scan, using the assigned points. A description of the complete double-scan procedure is given in Appendix C. In the simplified procedure, the first step is to examine pairs of points and decide whether they are close enough to belong to the same cluster. For this comparison we wish to use the degree to which points are statistically indistinguishable as a measure of closeness, rather than some absolute Euclidean distance in parameter space. Let u^n_{Js} be the 2×2 submatrix of U^n corresponding to wave Js. Then our basic definition of the distance between two points is

$$D_{nm}{}^2 = \sum_{Js}^{\text{low}} d_{nm,Js}{}^2 , \tag{2.52}$$

$$d_{nm,Js}{}^2 = \sum_{pp'} (f^n_{pJs} - f^m_{pJs})(2u^n_{Js} + 2u^m_{Js})^{-1}_{pp'} \times (f^n_{p'Js} - f^m_{p'Js}) . \tag{2.53}$$

The distance between two waves must be modified if the minimizer's estimates of u^n_{Js} and/or u^m_{Js} are sufficiently inaccurate that their sum is not positive definite. In that case, we drop the off-diagonal matrix elements of the sum and define $d_{nm,Js}{}^2$ to be

$$d_{nm,Js}{}^2 = \sum_{p} \frac{(f^n_{pJs} - f^m_{pJs})^2}{(2u^n_{Js} + 2u^m_{Js})_{pp}} . \tag{2.54}$$

As a further safeguard against the effects of anomalously large separations of points which are actually in the same cluster, we replace the largest two values of $d_{nm,Js}{}^2$ by the third largest value in the sum over Js in Eq. (2.52). This modification is in the same spirit as the previous enlargement of the smallest $N/10$ diagonal elements of U^n, but unlike the previous case it is usually confined to poorly determined partial waves with small J, typically S and P waves.

Two points, n and m, are now assigned to the same cluster if D_{nm} is less than some comparison distance. For small X^2 separation $|X_n{}^2 - X_m{}^2|$ we use a fixed comparison distance $D_{nm}{}^2 < D_0{}^2$. If the X^2 separation is sufficiently large we relax the criterion by letting the comparison distance increase linearly with $|X_n{}^2 - X_m{}^2|$ in order to include the "tails" of the cluster. Except in a few special cases, the X^2 separation at which we begin to relax the distance criterion is chosen to be the same δX^2 that was used to scale up U^n, and the general criterion used is

$$\frac{D_{nm}{}^2}{D_0{}^2} < 1 + \left(\frac{|X_n{}^2 - X_m{}^2| - \delta X^2}{2\delta X^2} \right) \theta(|X_n{}^2 - X_m{}^2| - \delta X^2) . \tag{2.55}$$

It can be seen from Eq. (2.53) that $D_0{}^2$ should lie somewhere between a lower limit of $\sim \nu/4$ (if U^n has been properly scaled) and an upper limit determined by the minimum separation between clusters. Within these rough guidelines there are large variations from energy to energy, depending on the quality of the data, the number and distribution of local minima found, etc. It is usually necessary to try several values of $D_0{}^2$ and observe the results before settling on a final value. Too large a value will result in all the local minima coalescing into one or two very large clusters; too small a value results in no clustering at all. After assigning points to clusters, we construct the central value and the covariance matrix associated with a cluster as averages over its points. The formulas used in constructing these averages are explained in detail in Appendix C.

At the conclusion of the solution-sorting procedure, the central value obtained by the averaging procedure for each cluster is replaced by the result of a new, brief minimization search consisting of a few steps of a modified steepest-descent minimization utilizing the final covariance matrix (which is not altered in this final minimization) to help determine the direction of the steps. This is done to ensure that the final central value will be very close to the best local minimum in the cluster. At this stage, small adjustments in the low partial-wave amplitudes are also made, using the final covariance matrix, to ensure that all high partial waves will exactly satisfy unitarity.

Through a series of tests, we have checked

that the procedure we use gives reasonable results and is stable. With reasonable choices for parameters such as D_0 and δX^2, the number of distinct clusters with acceptable X^2 values is not changed when new randomly started fits are added. Furthermore, if the total weight of a cluster is large enough, that is, if it already includes several fits with X^2 values within δX^2 of the minimum, then the central value and the covariance matrix are also stable.

III. SELECTION AND SMOOTHING OF ENERGY-INDEPENDENT FITS

At each energy, the fitting and solution-sorting procedures described above identify a number of regions in the parameter space; within each region the experimental data are adequately represented. The size of a region is determined not only by data uncertainties, but also reflects the existence of continuum (or phase) ambiguities. These ambiguities arise because the amplitudes are complex, so that there are more real quantities to be determined at each angle and energy than there are types of experimental data. The arbitrariness in phases is controlled and reduced to a manageable extent by the fitting techniques described earlier. Nevertheless, continuum ambiguities persist somewhat, as of course they should, since they do correspond to a real lack of information from experiments.

Identification of the correct amplitudes requires first, resolution of the discrete ambiguity, and second, after determination of the correct fit region, reduction of its size by further restriction of the continuum ambiguities. Finally, we must try to identify and correct possible biases in the original fitting which might arise from erroneous data, improper resolution of ambiguities in single-energy fits, insufficient searching for minima, etc. To help identify the correct amplitudes at each energy, we use the fact that the complex amplitudes $F(E)$ which describe the scattering are analytic functions of the energy E, i.e., the $F(E)$ satisfy dispersion relations.

In most applications of dispersion relations, $\mathrm{Re}F(E)$ at a given energy E is obtained as an integral over $\mathrm{Im}F$ at all energies. This approach is applicable to the forward C amplitudes, and we include calculated values of $\mathrm{Re}C(t=0)$ (Ref. 26) among the data used in single-energy fits. For other amplitudes F, however, $\mathrm{Re}F$ and $\mathrm{Im}F$ are known with similar accuracy and over the same limited energy region, so that the usual dispersion-relation approach is not suitable. We use general methods of functional analysis to develop an expansion for F in terms of functions whose energy dependence is consistent with the dispersion relations. This expansion is used to construct a smoothed fit to $F(E)$ which then provides the starting point, as we describe below, for solving the problems mentioned in the preceding paragraph.

There are many suitable candidate amplitudes $F(E)$ for application of dispersion relations. We choose to employ dispersion relations for the amplitudes $B^{(\pm)}$ and $C^{(\pm)}$ along hyperbolic curves in the Mandelstam plane (HDR). We have previously used the $C^{(+)}$ amplitudes along four such curves in a determination of the σ term for πN scattering.[35]

We avoid the use of partial-wave dispersion relations (PWDR) in this stage of the analysis because of our general philosophy of separating the partial-wave analysis itself from the study of the properties of individual partial waves. The analyticity properties of the HDR amplitudes are also slightly simpler than those of partial waves; in particular, they automatically incorporate s-u crossing symmetry. Another important reason for preferring HDR is that because of practical limitations of computer memory size, it is impossible to take exact account of all the error correlations in a simultaneous dispersion-relation fit to a large number of amplitudes. However, in HDR fits along curves which are fairly well separated in the Mandelstam plane, we can reasonably assert that at least the correlations between amplitudes on different curves are weak. This assertion is the basis of the iterative scheme described in Sec. IV for handling these correlations in an approximate way. In PWDR the correlations between all the fitted amplitudes could be large from the outset.

We also avoid the use of fixed-t dispersion relations (FTDR). These have been used in photoproduction[36-38] and in πN partial-wave analysis.[39,40] Compared with HDR, FTDR avoid having a left-hand cut with a complicated, unknown discontinuity; instead, there is a right-hand cut region where the amplitude can only be obtained by extrapolation of the partial-wave expansion. As discussed for photoproduction by Wessel[37] and Noelle,[38] use of our expansion would allow this extrapolation to be valid in principle, but the convergence would be slow and the extrapolation errors would be large. These problems with FTDR have their most serious effect in the backward-scattering region, where the fixed-t lines escape from the physical region. Since in this region the data are known to vary rapidly (and sometimes also erratically) as the energy and angle are changed, accurate treatment of this region is important. In the HDR method, the

curves stay inside the physical region, so the dispersion-relation input from nearby energies is derived as directly as possible from experimental information.

In the HDR fits, data from our energy-independent analyses have been supplemented by results from Bugg[41] (10 momenta, at 0.427 GeV/c and below) and from Ayed[32] (we used the results from 35 momenta, at 0.490 GeV/c and above). As the momentum range of our analysis has been progressively extended, the overlapping results from these groups have been dropped. Errors on the amplitudes from Ayed above our momenta were multiplied by 4 (and by a larger factor on curves I and II) so that their results would give only a very weak constraint on the extrapolation of our results to higher energies. We also use threshold values of F and $-idF/dq$ obtained from Bugg *et al.*[42] scattering lengths.

A. Hyperbolic dispersion relations

The curves on which we consider dispersion relations include five s-u-symmetric hyperbolas which pass through the threshold point and are completely defined by t_v, the value of t at their vertex ($s=u$), and by x_∞, the cosine of the asymptotic scattering angle. The values of these parameters are listed in Table II along with $x_1 = \cos\theta_{\text{c.m.}}$ at $P_{\text{lab}} = 1$ GeV/c. Curve I is actually a degenerate hyperbola, a pair of straight lines:

$$\nu = \pm 4M(t-\mu^2)/\mu\,, \quad \nu \equiv s-u\,. \tag{3.1a}$$

The other curves are described by the general formula

$$t^2 - (t_v^2 + b^2\nu_{\text{th}}^2)t/t_v = b^2(\nu^2 - \nu_{\text{th}}^2)\,, \tag{3.1b}$$

where

$$b = (1-x_\infty)/(3+x_\infty)\,, \quad \nu_{\text{th}} = 4M\mu\,.$$

Along with these five curves, we include as a "curve zero" the forward-scattering curve ($t=0$). Since the $C^{(\pm)}(0)$ amplitudes have been included directly, only the $B^{(\pm)}(0)$ amplitudes are treated by the methods of this section.

On the five nonforward curves, the amplitudes $C^{(+)}$, $B^{(-)}$, $C^{(-)}/(s-u)$, and $B^{(+)}/(s-u)$ are analytic functions of t with branch points on the real axis. The physical πN threshold is at $t=0$ and the $\pi\pi$ threshold is at $t_{\pi\pi}=4\mu^2$. On curve I, the second intersection with the πN threshold occurs at $t_{\pi N} = 8M\mu^2/(4M-\mu)$; on the other curves it is more distant than $t_{\pi\pi}$. There are no kinematical singularities associated with the vertices of the hyperbolas because we consider even functions of $(s-u)$. For

$$0 < t < t_2 = \min(t_{\pi\pi}, t_{\pi N})\,,$$

curve I intersects the nucleon pole twice, the other curves intersect it once (for $s>u$). On the other curves there is also a second intersection, but with $t > t_{\pi\pi}$, except for curve V (the backward-scattering curve), on which the second intersection is at $t=\infty$. The amplitudes are thus all analytic in the cut t plane, with a cut along the physical region $t \leq 0$, a second unphysical cut for $t \geq t_2$, and either one or two poles for $0 < t < t_2$.

Each of the 20 nonforward amplitudes $F(t)$ is represented in the form

$$F(t) = f_0(t) + \Sigma a_n h_n(t)\,, \tag{3.2}$$

where $f_0(t)$ is a "Born term" which includes nucleon and delta [i.e., $\Delta(1230$ MeV)] pole terms as well as the leading ηN threshold singularity. The expansion functions $h_n(t)$ are generated from the kernel function $H(t,t')$ associated with a Hilbert space of analytic functions, and the coefficients a_n are determined by minimizing the sum of χ^2 and the norm of the expansion as described in Appendix B. The construction of f_0 and H is outlined in the following subsections; the same construction method, except for minor differences, was used by Chao *et al.*[35]

Although each of the amplitudes is considered as a function of t, for the development of convenient representations for f_0 and H we first define auxiliary functions of t which approximate the Mandelstam variables. We define a function $\mathfrak{s}(t)$ which resembles $s(t)$ in the s-channel physical region but in which the kinematical singularity is pushed to $t > t_2$:

$$\mathfrak{s}(t) = [\tfrac{1}{2} + (a_s - b_s t)^{1/2}]^2 + c_s\,, \tag{3.3}$$

where a_s, b_s, and c_s are constants on a given hyperbola chosen so that $\mathfrak{s}(t) \approx s(t)$ in our energy region. For constructing the Regge factors for t-channel exchange terms, we introduce

$$\tilde{t}\,(t) = t f(\tfrac{1}{3} - \tfrac{2}{3}t)\,, \tag{3.4}$$

where $f(z)$ is defined in Eq. (2.4). For $\tilde{u}(t)$ we use a rather complicated expression to match well to $u(t)$ on the backward curves and in addition to arrange that $\tilde{u}$ will approach a constant for large t. With $\bar{u} = a_u - b_u t$, we have

$$\tilde{u}(t) = u_0 + (\bar{u} - u_0) f(c_u - d_u \bar{u}). \tag{3.5}$$

TABLE II. Hyperbolas used for dispersion relations.

Curve	1	2	3	4	5
t_v	μ^2	$2\mu^2$	$2.5\mu^2$	$3\mu^2$	$4\mu^2$
x_∞	0.857	0.4	−0.1	−0.7	−1.0
x(1 GeV/c)	0.814	0.263	−0.231	−0.687	−1.0

[A large number of adjustable constants were used because it was convenient to build in the desirable properties sequentially, not because they were all needed to obtain a sufficiently good match to $u(t)$.]

On curve 0 (forward scattering) there is no unphysical region, so the treatment can be simplified. We use the variable $\bar{t} = 16\,M^2\mu^2 - (s-u)^2$. The two amplitudes $F(\bar{t})$ [either $B^{(-)}$ or $B^{(+)}/(s-u)$] are represented in the same general form of Eq. (3.2), but with some minor differences in the construction of the functions $f_0(\bar{t})$ and $H(\bar{t},\bar{t}')$. The $F(\bar{t})$ on curve 0 have a physical cut for $\bar{t} < 0$, but no right-hand cut. Thus, we do not use the functions $\tilde{s}$, $\tilde{t}$, and $\tilde{u}$ in the forward direction.

B. Construction of the Born term

The nucleon contribution to f_0 has the general form $N(t)\,R(t)$ + corrections. Here $N(t)$ is the elementary pole contribution with an added constant to match a form appropriate to axial-vector coupling at the threshold $t = 0$:

$$\begin{aligned} N(0,B^{(+)}) &= -4Mg^2/[\mu(4M^2-\mu^2)], \\ N(0,B^{(-)}) &= -\mu N(0,B^{(+)})/2M, \\ N(0,C^{(+)}) &= \mu^3 N(0,B^{(+)})/4M^2, \\ N(0,C^{(-)}) &= \mu N(0,B^{(-)}) . \end{aligned} \tag{3.6}$$

On the five nonforward curves, the factor $R(t)$ is a Regge factor which has the general form given in Table I, with trajectory and residue functions constructed using $\tilde{s}(t)$ and $\tilde{u}(t)$; it is normalized to unity at the intersection of the curve with the locus of the nucleon pole which occurs at the smaller value of t. The signature factor is modified so that it will be real analytic and so that the leading S-wave threshold imaginary part will not be present:

$$\xi(t) = 1 + \cos\tilde{\alpha}\,\pi - i\,\tilde{W}^{3/2}\sin\tilde{\alpha}\,\pi\,, \tag{3.7}$$

where $\tilde{\alpha} = \alpha(\tilde{u})$ and

$$\tilde{W} = [\tilde{s}(t) - \tilde{s}(0)]/[\tilde{s}(t) + \tilde{s}(0)]. \tag{3.8}$$

The corrections are added terms dropping faster than $\tilde{s}^{-4}$ which compensate for the fact that $R(t)$ is not constant for $t > 0$. We construct these correction terms so that on curves I and II the correct residue will also be obtained at the second intersection of these curves with the nucleon pole locus, and so that the threshold values given by Eq. (3.6) will be reproduced exactly. On curve I we also match this threshold value at $t = t_2$.

The delta contribution to $f_0(t)$ has a Breit-Wigner resonance form written in terms of $\tilde{s}(t)$, which has P-wave threshold behavior and a pole corresponding to $s = (1.21 - 0.50i)^2$ GeV^2 and the partial-wave residue $(0.035 - 0.038i)$ GeV. Also, the respective values for $B^{(+)}$ and $B^{(-)}$ of 0 and 200 GeV^{-2} were matched at $t = 0$ on all curves, and at $t = t_2$ on curve I. These parameters, along with the partial-wave residue, were chosen to give a good fit to the Bugg amplitudes.[41] The entire contribution falls as $\tilde{s}^{-6}$ at high energies.

The leading ηN-threshold cusp term [proportional to $(t - t_\eta)^{1/2}$, where t_η is the value of t at which the hyperbola crosses the ηN threshold] is also included in f_0. This term has a coefficient whose modulus is determined from the η-production cross-section slope measurement of Binnie *et al.*[43] and whose phase has been determined by Bhandari and Chao.[28] The cusp term is made to vanish at the πN threshold so it will not contribute strongly to S-wave scattering near threshold, and the term is made manifestly real analytic by inclusion of an extra πN threshold factor in the imaginary part, as in Eq. (3.7) for the nucleon signature factor ξ. The contribution of the ηN term is made to fall off as $\tilde{s}^{-4}$ at high energies. This term is not significant in the HDR fits, except for $C^{(-)}$ within about 100 MeV/c of the ηN threshold, and for $C^{(+)}$ over a somewhat smaller range.

Since the forward amplitudes decrease weakly as s increases, we use the simple nucleon pole without an additional damping factor. Then $N(\bar{t})$ automatically satisfies Eq. (3.6) and drops off as s^{-1} for $B^{(+)}$ and as s^{-2} for $B^{(-)}$. The Δ-pole and ηN-threshold terms are similar to those used for the nonforward amplitudes except that weaker damping factors are used; the asymptotic form is given the same power of s as the nucleon pole. To provide damping, the factors must have a cut, and since there is no unphysical cut, they must be complex on the physical cut. We use an inverse power of

$$w = 1 - i\,(-\bar{t}/10)^{1/2}$$

and subtract a compensating term so that the threshold behavior of $B^{(\pm)}$ will correspond to P waves.

C. Construction of the kernel function

The expansion term in the expression (3.2) for $F(t)$ is assumed to lie in a certain Hilbert space of analytic functions. This function space is defined, first, by the domain in which the functions are required to be analytic, and second, by the norm in the space, which depends on how large the functions are within the domain and on its

boundary. The choice of Hilbert space defines a natural sequence of expansion functions and also determines the way the expansion is to be truncated, as in our use of the truncation function, Eq. (2.46), for the energy-independent fitting. A Hilbert space of analytic functions is also defined uniquely by the reproducing kernel function. When the data are related linearly to the fitting function it is easy to exploit the power and flexibility of the kernel function method[30,44] (see Appendix B). This is the situation for the HDR data.

The kernel used in fitting the nonforward amplitudes is written in the form

$$H(t,t')=[H_1(t,t')+H_2(t,t')]\left[\sum_{i=1}^{3} G_i(t)G_i(t')^*\right], \tag{3.9}$$

where the individual terms are chosen to provide various features which are desirable on empirical or theoretical grounds. As a technical aid in constructing the term $H_1(t,t')$, we map the cut physical sheet (including part of the $t > t_2$ cut) as well as a part of the second sheet (up to the curve drawn in Fig. 1) onto the interior of a strip. The second-sheet region is chosen so that it contains the Δ pole but no poles from other known resonances. Inclusion of this region implies neglect of the singularities associated with physical production thresholds (except for ηN), which is justified for the present work because these effects are weakly energy dependent and small in comparison with uncertainties in the input data. The use of a kernel associated with a strip is based on the assumption that as s increases, similar intervals of $\ln s$ correspond to a similar amount of variation in F.[44] The second term $H_2(t,t')$ is added in order to account in a simple way for the cut $t > t_2$. Finally, the modulating factors $G_i(t)$ are chosen to accommodate significant variations in the scale of the various amplitudes on the different curves as t is varied.

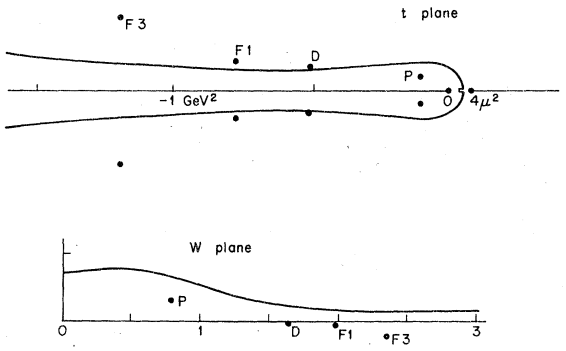

FIG. 1. Conformal mapping for curve 3 used in defining $H_1(t,t')$. The curve in the t plane is on the second sheet. This curve is mapped onto the real axis in the w plane. The curve in the w plane is the locus of the physical region for πN scattering, $t<0$. The location of the poles for several prominent resonances is shown on both planes. The notation used is the following: $P = P_{33}$ (1.23 GeV), $D = D_{13}$ (1.51 GeV), $F1 = F_{15}$ (1.69 GeV), $F3 = F_{37}$ (1.95 GeV).

The construction of the strip mapping is obtained in several stages. First, we introduce a variable $\sigma(t)$ which varies approximately like s on the curves, but which has a simpler dependence on t:

$$\sigma(t)=\alpha t+\beta[(\alpha t+s_0)^{1/2}-s_0^{1/2}], \tag{3.10}$$

where $s_0=(M+\mu)^2$ is the threshold value of s (used here to set a scale for the distance of the artificial singularity introduced) and where $\alpha=-2/(1-x_\infty)$ and β (see Table III) are chosen so that on the curves

$$\left.\frac{d\sigma}{ds}\right|_{t=0}=\left.\frac{d\sigma}{ds}\right|_{t=-\infty}=1 .$$

Next, the right-hand cut in the σ plane is opened up with a square-root mapping, and the further transformations

$$\begin{aligned} z &= e^{i\pi\gamma/2}\{(\sigma^{1/2}+\tfrac{3}{2}iM)^{1-\gamma} \\ &\quad +[(\sigma^{1/2}+i\rho)^2-\sigma_0]^{(1-\gamma)/2}\}, \\ \zeta &= a(z-ic)^2+b \end{aligned} \tag{3.11}$$

are applied. We use $\gamma=0.05$; this corresponds to requiring resonance poles to satisfy asymptotically $|\mathrm{Im}s| \gtrsim 0.05\ \mathrm{Re}s$, but this relation applies only for Res far beyond the region covered in our analysis (see Fig. 1). The parameter σ_0 governs the value of Res at which the second-sheet region is narrowest; with $\sigma_0=0.9M^2$ for curve I and $\sigma_0=M^2$ for the others, this occurs for Re$s\approx 3M^2$ on all curves. The parameter ρ is taken to be $M/2$.

The other curve-dependent parameters are listed in Table III. The value of c was determined by requiring the point

TABLE III. Mapping parameters for Eqs. (3.10) and (3.11).

Curve	β	a	b	c
0		0.161	0.036	2.07
1	0	0.214	0.006	1.94
2	0.56	0.211	0.021	1.93
3	1.11	0.208	0.024	1.91
4	1.53	0.207	0.025	1.90
5	1.42	0.207	0.023	1.91

$$s = s_c = (2.75 - 0.15i)\ \mathrm{GeV}^2$$

to map onto the real ζ axis, b was adjusted so that the second-sheet nucleon pole would be located at a positive value of ζ, and a was chosen so that $\sigma = -s_0$ would map into $\zeta = -1$.

The location of the point s_c was based on an estimate of the position, obtained after the first round of energy-independent fitting and HDR smoothing, of the resonance pole likely to be nearest the real axis. The second D_{13} resonance was used for this estimate; it was known to be the most critical, but its width was not known very well. However, a small mistake in the value of $\mathrm{Im}\, s_c$ is not as serious as it might seem at first. In fact, the first round of HDR fits employed different mapping parameters from those quoted here, so that the mapping brought the D_{13}' well into the upper half ζ plane. Since the HDR results are used indirectly and our studies of partial-wave amplitudes are based on the results of refitting to observables, as described later, any bias against narrow resonances produced by oversmoothing is reduced. As a result, we were not prevented from finding a resonance that was narrower than expected.

The mapping onto a strip of width π is given by

$$w = 2\ln[\zeta^{1/2} + (1+\zeta)^{1/2}] \tag{3.12}$$

(see Fig. 1). In the strip we use the elementary kernel function[44]

$$H_0(w, w') = \frac{i}{\sinh\frac{1}{2}(w - w'^*)} \tag{3.13}$$

modified as follows:

$$H_1(w, w') = H_0(w, w') - (1-\Delta)K_1(w)K_1(w')^*/K_2\ , \tag{3.14}$$

where

$$K_1(w') = \left(\frac{dH_0(w, w')}{dw}\bigg|_{w=w_0}\right)^*\ ,$$
$$K_2 = \frac{dK_1(w)}{dw}\bigg|_{w=w_0}\ , \tag{3.15}$$

and $w_0 = w(t=0)$. This modification reduces the threshold derivatives of H by a factor Δ, and compensates for the fact that dw/dq is rather large at threshold. We use $\Delta = 0.3$ for C and $\Delta = 0.05$ for B; the extra reduction for B is introduced because the scale for dB/dq is very small in comparison to the scale for B near threshold. (Only S waves contribute to the derivative, while P waves contribute to the threshold value of B with large coefficients.)

The cut $t > t_2$ is generated by a kernel function which corresponds to square-root singularities on a unit circle[44]:

$$\bar{H}(v, v') = vv'^* + \tfrac{1}{2}(1 - vv'^*)^2 \ln\frac{3 - vv'^*}{1 - vv'^*}\ . \tag{3.16}$$

To make the nearby part of the cut somewhat weaker than the more distant part (which includes the ρ and the reflected Δ), we add two terms

$$H_2(t, t') = A_2[\bar{H}(v_c, v_c') + \bar{H}(v_d, v_d')]\ , \tag{3.17}$$

where

$$v_x(t) = \frac{(t_x - t)^{1/2} - M}{(t_x - t)^{1/2} + M} \tag{3.18}$$

and $t_c = t_2$, $t_d = 4t_2$. The relative scales of H_1 and H_2 are determined by A_2; we take $A_2 = 4$ so that H_1 and H_2 give similar contributions for $0 < t \lesssim t_2$.

We use three modulating factors $G_i(t)$ rather than a single one in order to have greater flexibility in accommodating the quite different variation with energy associated respectively with t- and u-channel Regge exchanges and with more central effects such as given by parton interchange. These three functions are nonvanishing real-analytic functions of t with branch cuts for $t \gg t_2$. The forward term $G_t(t)$ is taken to be a Regge-type function of $\tilde{t}$ and $\tilde{s}$ which matches the scale (but not oscillations or phase variations) of the Pomeron term (for $B^{(+)}$ and $C^{(+)}$) or the ρ-exchange term (for $B^{(-)}$ and $C^{(-)}$) in the fit of Alcock *et al.*[45] The backward term $G_u(t)$ is a Regge-type function of $\tilde{u}$ and $\tilde{s}$ which matches the scale of Δ exchange in the same fit. Finally, $G_s(t)$ is a function of $\tilde{s}$ which drops off as $\tilde{s}^{-4}$ for large $\tilde{s}$. For the crossing-odd functions $C^{(-)}$ and $B^{(+)}$, an extra factor $(s-u)/\tilde{s}(t)$ is included in all three modulating functions.

The kernel for the forward amplitudes is taken to have the form

$$H(\bar{t}, \bar{t}') = H_1(\bar{t}, \bar{t}')G(\bar{t})G(\bar{t}')^*\ . \tag{3.19}$$

The term H_2 is omitted because there is no cut for $\bar{t} > 0$. To construct H_1 we again use Eqs. (3.13)–(3.15); the mapping to the strip is obtained by use of Eqs. (3.11) and (3.12) with $\sigma = -\bar{t}/(8M\mu)$ and with the parameters $\sigma_0 = M^2$, $\rho = M$, and $\gamma = 0.025$.

Since the forward amplitudes do not decrease rapidly at high energies, a strong damping factor is not required. For $B^{(-)}$, we take, simply, $G = 1$. For $B^{(+)}$, to compensate the symmetry factor $(s-u)$, we use

$$G(\bar{t}) = 3(s-u)(y + 2\bar{t}^{1/2})/(y + \bar{t}^{1/2})^2 \tag{3.20}$$

with $y = 12\ M\mu$. Note that this function is complex in the physical region $\bar{t} < 0$. It approaches a constant at high energies.

D. Resolution of discrete ambiguities

To resolve the discrete ambiguity at each energy, we must choose the correct sequence of fit regions at successive energies—the correct "path." We do this partly by eye, and partly by an automated comparison with a previous set of HDR fits.

For each of the amplitudes $F(t)$, we drop the input data at a set of $n (n = 1, \ldots, 5)$ contiguous momenta, and obtain, from the remaining input data to the HDR fit, predictions for F at these momenta. These predictions are then compared with the single-energy fits, as described below. The predictions for each amplitude are independent, but for a given amplitude the predictions at nearby momenta are strongly correlated. The correlation distance (in terms of s) is about twice the distance to the curve in Fig. 1 from the real axis, which corresponds to a range of about 3–4 neighboring momenta.

For each initial momentum i and $n \leq 5$ we consider sequences $S_{i,n} \equiv (K_i, \ldots, K_{i+n-1})$, where K_j denotes the Kth fit region found at the jth momentum. We then compute $\chi^2(S_{i,n})$ for agreement of the sequence of fits with the predictions. This χ^2 takes into account, for each K_j, the correlations between different amplitudes on the same curve, and for each amplitude, the correlations between the predictions at different momenta. To save computer time, if $n > 2$ we do not examine all possible sequences $S_{i,n}$, we consider only those sequences $(K_i, \ldots, K_{i+n-1})$ in which the subsequence $(K_i, \ldots, K_{i+n-2})$ is among the better $S_{i,n-1}$ and the subsequence $(K_{i+1}, \ldots, K_{i+n-1})$ is among the better $S_{i+1,n-1}$ found in previous calculations for sets of $n-1$ contiguous momenta. In general, we limit the $S_{i,n-1}$ to about 20 having the lowest values of χ^2, but also require that their χ^2 not differ from the best value by more than about 500 [the expected value of $\chi^2(S_{i,n})$ is $44n$]. To guard against the possibility that the results might be distorted by a poor estimate of the covariance, we also make calculations of χ^2 in which the fit covariances and prediction covariances are not weighted equally.

Longer-range correlations from the HDR fits are checked by making an initial selection of a path through fit regions at all energies, and using these as data for new HDR fits from which predictions are derived. The final path is one which is stable as this procedure is repeated.

The selection of the best path was most difficult at the earliest stage of our analysis, when we used energy-independent fits made without input constraints from HDR, because then there were many diffuse fit regions at each energy. Using continuity and using comparison with HDR fits to the amplitudes of Ayed as a weak constraint, two quite distinct plausible paths were chosen in the region $0.82 \leq P_{\text{lab}} \leq 1.437$ GeV/c. HDR fits to both of these paths gave predictions which led, after a few rounds of reselection and fitting, to a unique path which was different from both initial guesses.

Results from the alternate-path-comparison program are usually unambiguous—for most momenta (j), the best $S_{i,n}$ for various i and n usually include consistently one particular fit region K_j. The K_j selected in this way have been, at almost all momenta, either the fit which gave the lowest χ^2 in the energy-independent fitting or else the second lowest. If the alternate path comparison was ambiguous, which in practice meant that $\chi^2(S_{i,n})$ calculated with two different K_j were nearly equal, we examined the behavior of partial waves with prominent, noncontroversial structure. In almost all cases of ambiguity a satisfactory resolution was to combine the two fit regions into one larger one; our general approach, in cases of difficulty, has been to enlarge the errors, so that HDR fits at neighboring momenta would not be strongly biased. Most ambiguities disappeared after the cycle of HDR fitting-path selection was repeated a few times.

At a few momenta HDR fitting alone was not enough to resolve the situation. At these momenta new energy-independent fits, using predictions from the HDR as additional input data, gave results which were more consistent with the behavior at neighboring momenta.

E. HDR fits and continuum ambiguities

As examples of HDR fits, we show in Fig. 2 the $C^{(\pm)}$ and $B^{(\pm)}$ amplitudes on a typical interior curve, number III. The $B^{(\pm)}$ amplitudes on the forward curve are shown in Fig. 3. We also show, in Fig. 4, the $C(\pi^* p)$ amplitudes on the backward curve. The moduli of these amplitudes are determined by the cross sections for backward elastic scattering of π^+ and π^-, while the phases are determined by our fitting procedure. A relation among the amplitudes is also obtained from the cross section for backward charge-exchange scattering. However, the 180° cross sections for elastic and charge-exchange scattering usually involve a significant extrapolation from angles at which measurements were made. The HDR fits on the backward curve therefore provide a check on the extrapolated 180° cross sections derived

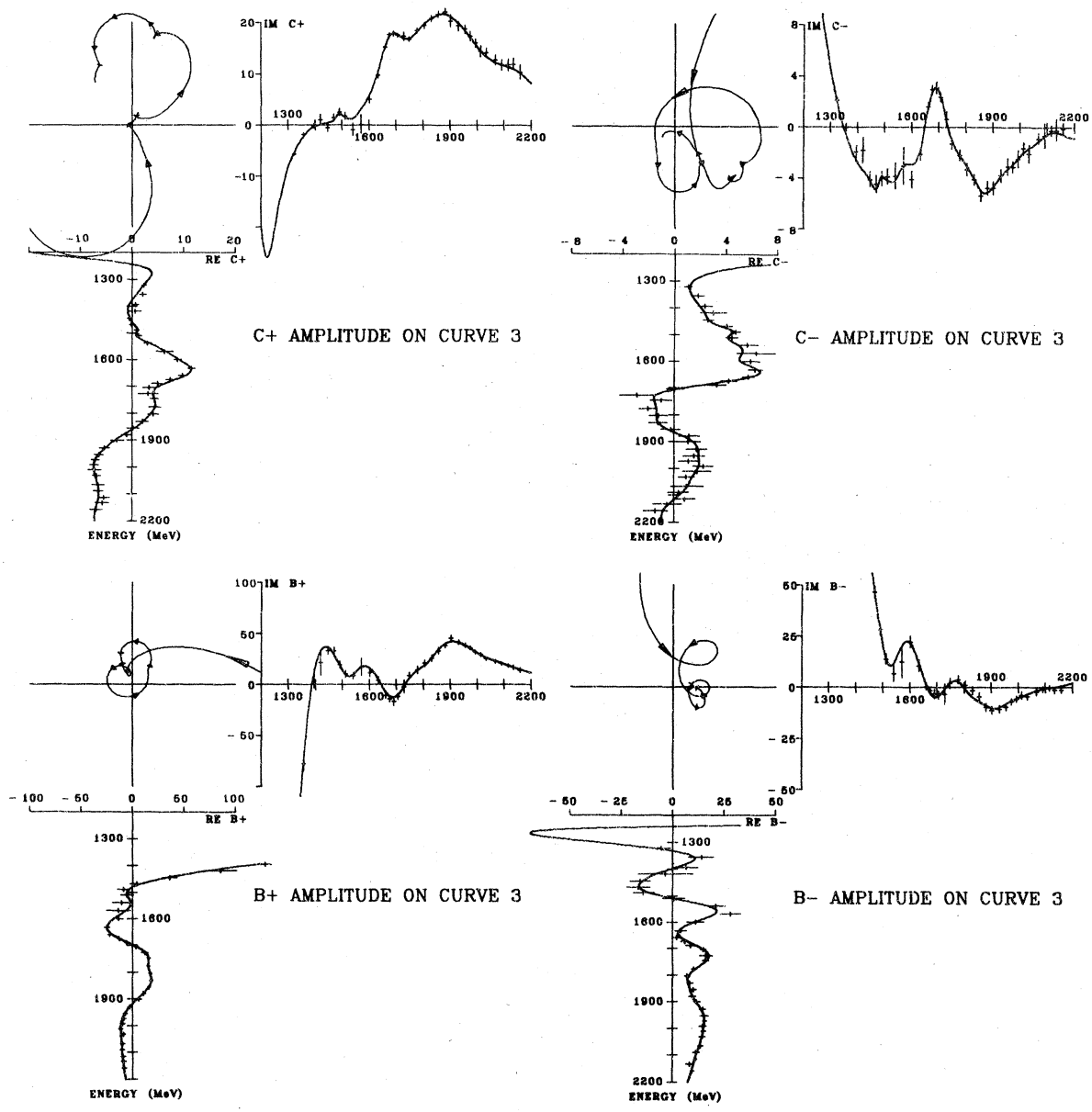

FIG. 2. The $C^{(\pm)}$ and $B^{(\pm)}$ amplitudes on curve 3. The energy scale is labeled in MeV, but the scales for C and B are in GeV^{-1} and GeV^{-2}, respectively (see Appendix A). The data points for real and imaginary parts are the values obtained from the single-energy fits. The errors shown are the diagonal errors after compensation for the effect of HDR constraints [Eqs. (4.8)]. However, at several energies where this procedure led to errors which were large and highly correlated to an unusual degree, we have plotted instead the uncompensated error multiplied by two. The fits shown take into account the effects of error correlations among the amplitudes. On the projections of the fits onto the Argand diagram, the energy dependence is indicated by arrows drawn at 100 MeV intervals, with the length of the arrow proportional to the rate of change.

from the fitted partial-wave expansion in energy-independent analyses, as well as on the fitted phases. This point is discussed further in Sec. V.

To a large extent, fits such as those shown in Figs. 2–4 are determined by simple continuity and smoothness requirements on the real and imaginary parts separately, but the cut-plane analyticity also allows both parts to change rapidly as the energy is increased, if the changes are related. Sharp turns to the left (anticlockwise loops) are tolerated, while turns to the right are inhibited. These restrictions on the energy de-

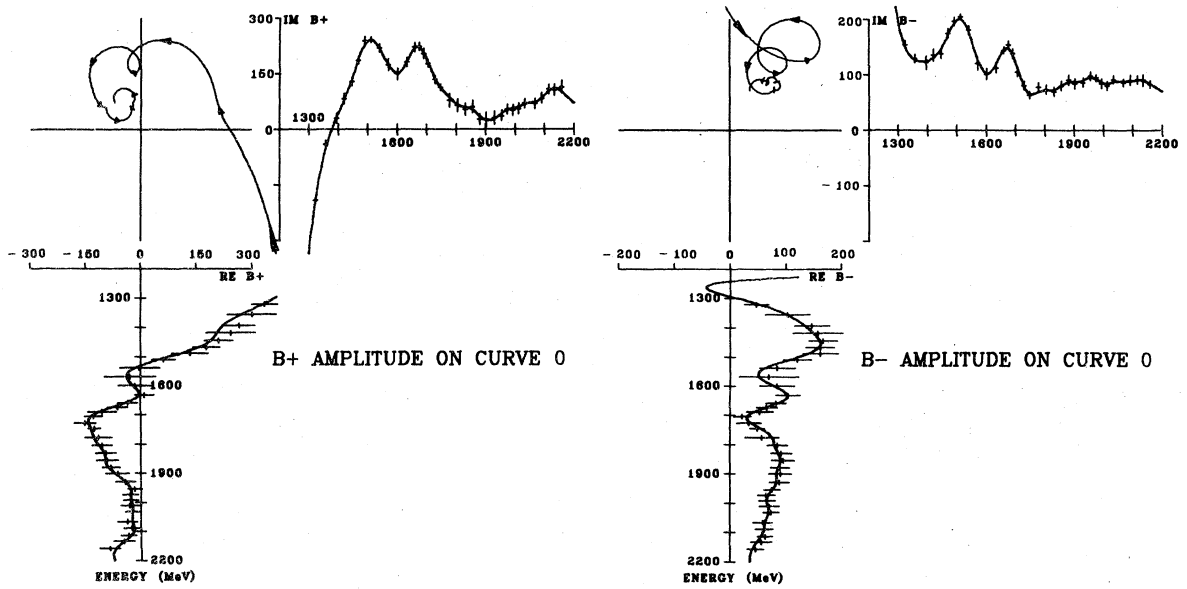

FIG. 3. The $B^{(\pm)}$ amplitudes in the forward direction. (See also the caption for Fig. 2.)

pendence of amplitudes serve as additional information, provided by the dispersion relations, which constrain the continuum ambiguities.

At interior scattering angles, the relations between the amplitudes, the observables, and the continuum ambiguities are more complicated than they are at 180°. However, it is possible to understand in a general way how the HDR constrain the ambiguities without examining the explicit relations.

Continuum ambiguities show up as large, highly correlated uncertainties in the amplitudes, obtained in the energy-independent fitting, which the HDR use as input data. We treat the correlations between amplitudes on the same curve by an iterative scheme described in Sec. IV. Correlations between amplitudes on different curves are ignored; this is equivalent to giving the continuum ambiguities an artificially enlarged scope, and is reasonable if $\Delta\theta$, the angular separation between

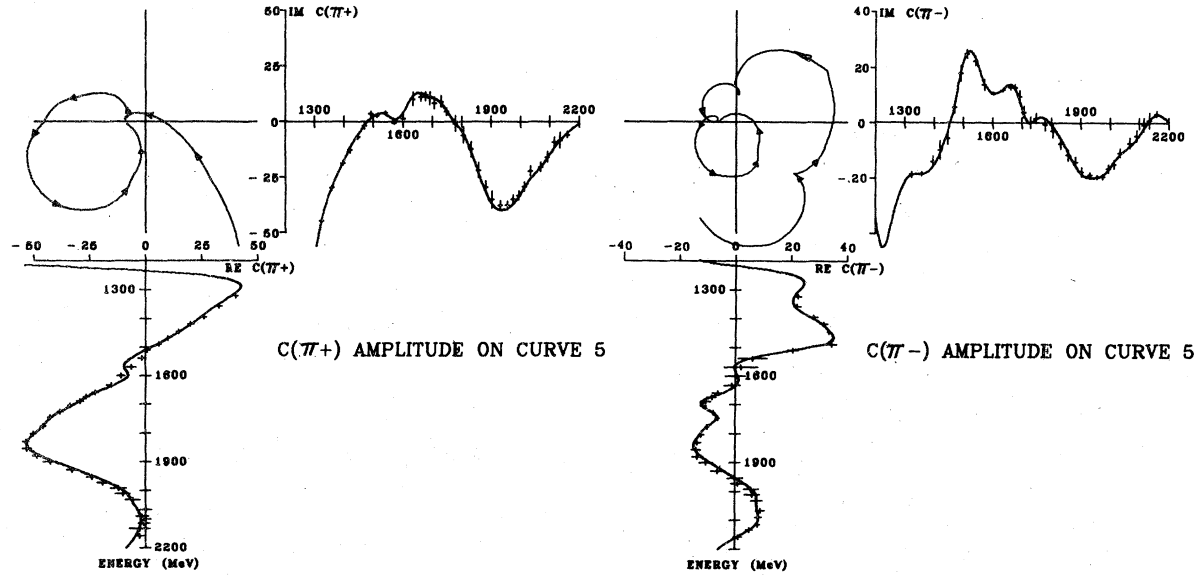

FIG. 4. The $C(\pi^{\pm}p)$ amplitudes for backward scattering. The square moduli of these amplitudes are proportional to the backward elastic DCS. (See also the caption for Fig. 2.)

curves, satisfies $\Delta\theta \gtrsim \pi/J$, where J is the highest angular momentum of a partial wave with a significant uncertainty.

A quantitative measure of the degree of correlation between the amplitudes on a given curve can be obtained from the eigenvalues of the 8×8 correlation matrix. Typically, the largest eigenvalue is about 20 times the smallest eigenvalue. The large eigenvalues correspond to continuum ambiguities, while the small eigenvalues correspond to combinations of amplitudes which are directly constrained by well-measured observables.

Our HDR fits to a given amplitude require about $\frac{1}{3}$ as many independent coefficients a_n as there are data. On the average, therefore, the errors are reduced by about $1/\sqrt{3}$ by fitting, and the HDR predictions at a given momentum have errors about $1/\sqrt{2}$ times the errors on data at neighboring momenta.[46] (At the higher momenta, these predictions have greater relative uncertainties than at lower momenta.) In the iterative fitting, use of the predictions for three amplitudes also reduces the errors on the fourth by a similar factor. Since we fit each amplitude separately, we do not obtain correlated predictions for different amplitudes. For each amplitude, we find that predicted real and imaginary parts at the same momentum generally have about the same error and are weakly correlated. This is an indication that each component depends, although in a somewhat complicated way, on both the real and imaginary parts at the other energies.

Since the prediction errors for real and imaginary parts of a given amplitude are nearly equal and uncorrelated and also somewhat smaller than the input errors, the HDR predictions effectively constrain the long axes of the error ellipses from the energy-independent analyses, but not the short axes. In other words, the HDR predictions have more weight than the combination of other effects which limit continuum ambiguities in fitting at a single energy, but less weight than the actual experimental data. This would be a basic requirement of any reasonable analysis procedure. Although the HDR constraints were not, in general, strong enough to pull the single-energy fits away from the data, the HDR predictions have been useful as an indication of possible systematic errors in some of the data.

IV. ITERATIVE FITTING PROCEDURE

We use an iterative procedure to approximate the effect of simultaneously fitting to all data by amplitudes which incorporate Mandelstam analyticity. The main part of the iteration cycles between single-energy fits (Sec. II) and single-curve fits (Sec. III). However, even on a single curve it is not practical to fit all amplitudes simultaneously, so we take the correlations between the amplitudes into account by another iteration method, in which each amplitude is fitted separately as described in Sec. III.

A. Amplitude correlations in HDR

The correlations between the amplitudes on a given curve are treated approximately. The data for a given amplitude are adjusted to take into account the results of HDR fits to the other amplitudes. This adjustment is made separately at each momentum.

For any given curve and momentum, let X denote the vector of amplitudes obtained from the single-energy fits (X has K real components) and let W be their $K\times K$ weight matrix (inverse covariance matrix). Here $K=8$ for the nonforward curves, but for the forward B amplitudes $K=4$. Also, let P denote the vector of predicted amplitudes and V the corresponding weight matrix obtained by using the data for the curve at all other momenta in the HDR. Since the HDR treat each amplitude separately, V has a block diagonal form with 2×2 blocks. The vector of fitted amplitudes F minimizes the quantity, written in matrix notation,

$$\chi^2=(F-X)^{\dagger}\,W(F-X)+(F-P)^{\dagger}V(F-P)\ . \quad (4.1)$$

Considering any of the four complex amplitudes $C^{(\pm)}$ or $B^{(\pm)}$, we use the subscript a to denote the corresponding 2-dimensional subspace, and use the subscript b to denote the orthogonal $(K-2)$-dimensional subspace. Thus, X_a, P_a, and F_a are vectors with 2 components and refer to the given amplitude, while X_b, P_b, and F_b have $K-2$ components, comprising the other amplitudes. Likewise, W_{aa} and V_{aa} are 2×2 submatrices of W and V, W_{bb} and V_{bb} are $(K-2)\times(K-2)$ submatrices, and W_{ab} is a $2\times(K-2)$ submatrix, etc. ($V_{ab}\equiv 0$). Using this notation, we can write the equations for minimization of (4.1) as

$$W_{aa}(F_a-X_a)+W_{ab}(F_b-X_b)+V_{aa}(F_a-P_a)=0\ , \quad (4.2a)$$

$$W_{bb}(F_b-X_b)+W_{ba}(F_a-X_a)+V_{bb}(F_b-P_b)=0\ . \quad (4.2b)$$

The HDR fit to the amplitude a is made independently of the other amplitudes, and the information from other momenta is incorporated into P_a and V_{aa}. Therefore, at the given momentum, F_a minimizes a quantity

$$\chi_a^2 = (F_a - Y_a)^\dagger S_{aa}(F_a - Y_a) + (F_a - P_a)^\dagger V_{aa}(F_a - P_a), \tag{4.3}$$

where S_{aa} and Y_a are to be chosen so that (4.3) will be consistent with (4.1). Apart from a constant, $\chi^2{}_a$ is the minimum of χ^2 (4.1) with respect to F_b. Thus, solving (4.2b) for F_b, we obtain

$$F_b - X_b = -R_{bb}[W_{ba}(F_a - X_a) + V_{bb}(X_b - P_b)], \tag{4.4}$$

where

$$R_{bb} = (W_{bb} + V_{bb})^{-1}. \tag{4.5}$$

Substitution into (4.2a) and comparison with Eq. (4.3) gives the result

$$\begin{aligned} S_{aa} &= W_{aa} - W_{ab} R_{bb} W_{ba}, \\ Y_a &= X_a - S_{aa}^{-1} W_{ab} R_{bb} V_{bb}(P_b - X_b). \end{aligned} \tag{4.6}$$

To start the iteration, we take $V = 0$; then $Y_a = X_a$ and S_{aa}^{-1} is the 2×2 covariance matrix for X_a. Using these input data, we obtain from the HDR values for P and V; these, when used in Eqs. (4.5–6), give new values for Y and S, and so on. We find that three or four steps of this iteration, when safeguarded against oscillation, are adequate for convergence.

In cases where there is a discrepancy between the data and the HDR predictions, the iteration is modified by an error-stretching prescription. For a particular energy and for the amplitude a, the 2×2 weight matrix V_{aa} for the prediction is reduced when the combined χ^2 for (fit) – (prediction) at the two neighboring momenta exceeds a critical value χ_0^2. These values of χ^2 are obtained from Eq. (4.3), minimized with respect to F_a; the reduction is by a factor $r_a = \chi_0^2/\chi^2$, and we use $\chi_0^2 = 8$. Then for a given energy and curve, if $\chi^2 > K$ [where χ^2 is given by Eq. (4.1), minimized with respect to F] the $K\times K$ matrix V is multiplied by $r = (K/\chi^2)^{1/2}$ and also the four 2×2 matrices S_{aa} as given by Eq. (4.6) are multiplied by r. Thus, the factor r is used both to reduce the effect of correlations and also to enlarge the errors directly. Note that after the iteration converges, $r \approx (K/\hat{\chi}^2)^{1/3}$, where $\hat{\chi}^2$ is the value that would have been obtained without any error stretching.

B. Dispersion-relation input to single-energy fits

In cycling between HDR fits and single-energy fits, the principle is similar to the one used in Eqs. (4.1)–(4.6). However, a modification is made to compensate for the fact that the χ^2 surface for the energy-independent fitting is, in general, very poorly approximated by a quadratic function. Instead of the "predicted amplitudes" obtained by removing the entire effect of the input data, we use the "fitted amplitudes," because these provide more stability against the effect of statistically insignificant local irregularities in the χ^2 surface.

The term D in Eq. (2.45) is given by

$$D = \sum_a (\mathfrak{F}_a - \tilde{P}_a)^\dagger \tilde{V}_{aa}(\mathfrak{F}_a - \tilde{P}_a), \tag{4.7}$$

where the sum extends over all 22 (complex) HDR amplitudes, and where the $\mathfrak{F}_a$ are the fitted amplitudes evaluated on the curves by summing the partial-wave expansion without electromagnetic corrections. As remarked above, we take $\tilde{P}_a = F_a$, where F_a is the HDR fitted amplitude, rather than $\tilde{P}_a = P_a$. To ensure that the predictions for observables would be given a small weight in comparison with experimental data on the observables, and also to compensate for the approximate treatment of correlations in the dispersion-relation fits, the weight of the prediction is reduced by a factor: $\tilde{V}_{aa} = \alpha V_{aa}$ (in general, we use $\alpha = 0.3$–0.5).

Single-energy fits, with the term D included in X^2, were made as described in Sec. II. Searches for minima of X^2 were started from positions of minima found previously, in which the term D had been absent or had been obtained from earlier estimates of $\tilde{P}$ and $\tilde{V}$. Also, new randomly chosen points were chosen to replace earlier fits at which the starting values of D were more than 10 times the expected value. Searches which started from recycled fits usually terminated quickly, with the presence of the new information being accommodated by a minor adjustment of the partial-wave amplitudes. In addition, however, many fit regions would disappear completely when HDR constraints were included in X^2.

The new single-energy fits, incorporating the effects of the constraints, are sorted into fit regions by the procedure described in Sec. IIE. Then the method of Sec. IIID is used to select the correct path through the fit regions at successive energies. The set of best fits (lowest X^2 value), one from each of these selected regions, constitutes our results—our values for the partial-wave amplitudes. The errors on these amplitudes are given by the size of the fit region, estimated by the covariance matrix U as determined by the method of Sec. IIE.

In the next cycle of HDR fitting we use as input amplitudes in Eq. (4.1) the best-fit quantities: $X = \mathfrak{F}$. This ensures that X corresponds to an adequate fit to the experimental data; because of the irregular nature of the χ^2 surface, this might not otherwise be true. Although the biases cancel on

the average when best-fit amplitudes are used as input to the energy-independent and dispersion-relation fitting, the final results at nearby energies may be somewhat more correlated then they would be if another prescription were adopted. The amplitudes $\mathfrak{F}$ have a covariance matrix $\mathfrak{U}$ derived from the partial-wave covariance matrix U. The weight matrix W used in Eq. (4.1) is corrected for the effect of the input HDR information as follows

$$W = W_0 - \tilde{V}/\delta\chi^2 , \tag{4.8}$$

where W_0 is the inverse of the $K\times K$ submatrix of $\mathfrak{U}$ for the amplitudes on a given curve, and $\delta\chi^2$ is the scale factor defined in Sec. II E.

There is no guarantee that $X^2 - D$ will have a positive-definite second-derivative matrix at the point where X^2 is minimum; hence W is sometimes not positive definite. This could conceivably cause difficulties in Eqs. (4.1)–(4.6), but since the instability is clearly a very local property ($X^2 - D$ would certainly rise again after moving some distance in the unstable direction) it is physically reasonable to impose a positivity constraint on W. Our constraint is that the smallest eigenvalue of W should be at least as large as 0.001 times the largest eigenvalue (the eigenvectors are not altered). The computation of eigenvalues is made using the diagonal errors of the components of $\mathfrak{F}$ to define the metric.

V. RESULTS AND CONCLUSIONS

In this section we discuss the quality of our single-energy fits and show some sample comparisons with data. In the following paper[3] we discuss the partial-wave amplitudes and the determination of resonance parameters. Graphical display of the individual partial-wave amplitudes is deferred to that paper, where the amplitudes are shown along with the fits used for resonance parameter extraction. Our partial-wave amplitudes for $J \leq 39/2$, along with the error matrices, are available to interested users on magnetic tape; inquiries should be directed to R.L.K. Inquiries about use of our computer programs may

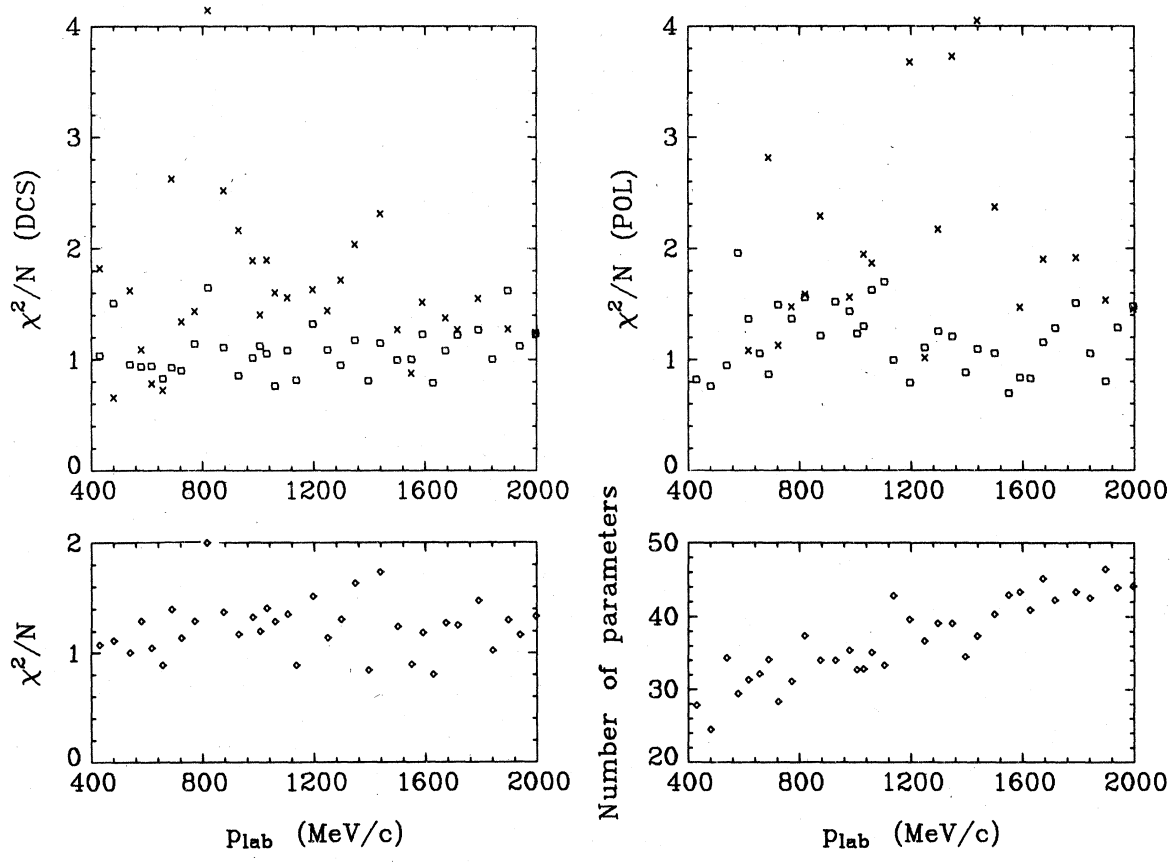

FIG. 5. The upper plots show the values of χ^2 per datum for our best fits to amalgamated elastic (□) and charge-exchange (×) cross-section and polarization data at each of our 35 momenta. The lower plots show the overall value of χ^2 per datum and the effective number of parameters at each of the momenta.

be directed to R.L.K. or R.E.C.[47]

The quality of our single-energy fits to amalgamated data is illustrated in Fig. 5. The upper plots show that the fits to elastic data are generally satisfactory, but that the charge-exchange data are difficult to fit at many of the momenta. This is in spite of the fact that the errors of the charge-exchange data are often significantly enlarged in the amalgamation procedure. The effects of the "pull" of the charge-exchange data on the elastic fits are indicated in Table IV. The problems with charge exchange arise primarily from the data of Brown *et al.*[14,15] and are similar to the difficulties experienced by Pietarinen.[48] If the data of Brown *et al.* are correct, one is led to the conclusion that there is significant structure in the charge-exchange amplitudes which is absent from the elastic amplitudes. This might be due to a breakdown of isospin conservation, but the data of Brown *et al.* are also difficult to fit in the data-amalgamation phase of our analysis where no isospin constraints are present. Another possibility is that the errors of the data are underestimated due to the presence of unknown experimental biases. We have looked for specific problem regions in the data that could be blamed on unknown biases, and have been unable to find anything very distinctive. It appears that if the difficulties are due to biases they are of a rather erratic character and result in an overall level of scatter in the data which is larger than would be expected on the basis of the quoted errors.

The overall values of χ^2 per datum shown in Fig. 5 are not much larger than those for the elastic data alone, because there are generally many more elastic data than charge-exchange data. Overall values of χ^2 per degree of freedom are typically about 20% larger than χ^2 per datum. The effective number of free parameters ν, defined by Eq. (2.51), is plotted vs momentum in Fig. 5, and increases from approximately 25 to 45 over the range studied. As discussed in Sec. IID, this variation in ν is obtained with fixed scale parameters in the TF.

Figures 6–9 show our best fits to all the amalgamated data at 820, 1030, 1437, and 1995 MeV/c. We have chosen to show four data sets which include both charge-exchange cross-section and polarization data, so the quality of the fits in these figures is somewhat worse than average. The overall value of χ^2 per datum is 2.0 at 820 MeV/c, which is the worst of any of our 35 data sets, while at 1030, 1437, and 1995 MeV/c the values are 1.4, 1.7, and 1.3, respectively. The relative precision of the six different types of data shown in the figures is typical of the whole resonance region.

An important aspect of our use of amalgamated data is the allowance for highly correlated statistical fluctuations in the data. In the course of a single-energy fit one obtains fitted amplitudes for these fluctuations which can be used to correct the data for their effects, in a manner similar to the renormalization of experimental data using a fitted scale factor. This procedure is described in Sec. IX of the preceding paper[2] where the correction is defined in Eq. (9.34). The data shown in Figs. 6–9 have been corrected in this way, and the χ^2 contributions arising from the constraints on the correction amplitudes were included in the values of χ^2 per datum discussed above. The correction amplitudes are often quite significant, indicating the importance of allowing for this kind of error. In the case of quantitatively poor fits (χ^2 per datum $\gtrsim 1.5$) it is a fairly common occurrence for a large fraction of the χ^2 excess to arise from the presence of a large amplitude collective fluctuation. Such fits would, of course, be still worse if one did not allow for the possibility of collective fluctuations.

An interesting example of this kind of effect occurs in the fit to charge-exchange polarization data at 1030 MeV/c shown in Fig. 7. The amalgamated charge-exchange polarization data at this momentum comes from measurements of Brown *et al.*[15] at 1027 MeV/c and Shannon *et al.*[17] at 1030 MeV/c. This is the one momentum where a

TABLE IV. χ^2 per datum for each of the six data types summed over subsets of the 35 data sets analyzed. The top line includes all the data sets. The second line includes only the 20 data sets which contained data on both charge-exchange cross sections and polarizations. The bottom line includes only the 5 data sets which contained no data on either charge-exchange cross sections or polarizations.

Subset	π^+p DCS	π^+p POL	π^-p DCS	π^-p POL	CEX DCS	CEX POL
All	1.08	1.12	1.08	1.25	1.65	2.09
CEX	1.14	1.14	1.14	1.30	1.75	2.09
$\overline{\text{CEX}}$	0.88	1.03	0.93	0.94	···	···

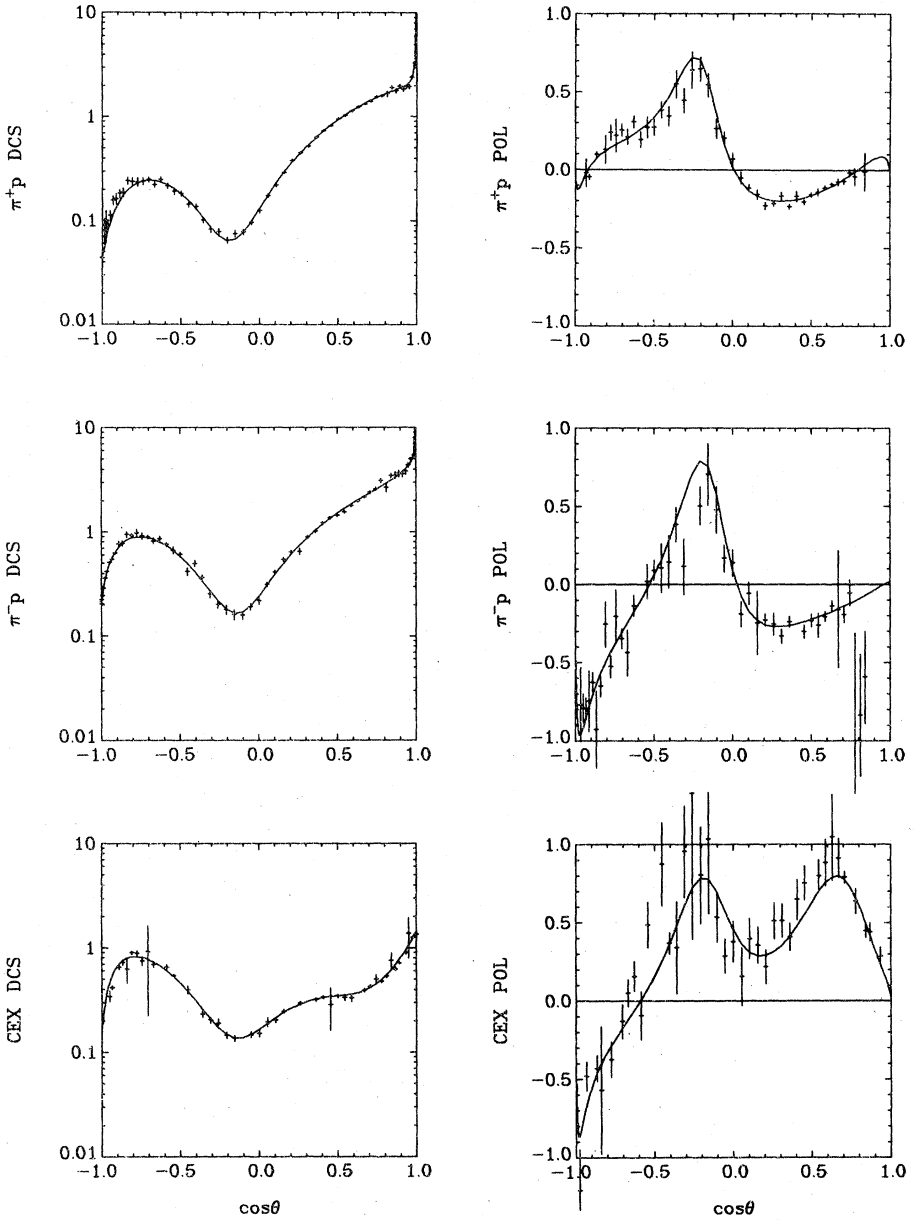

FIG. 6. Single-energy fit to amalgamated data at 820 MeV/c. DCS units are mb/sr.

clear systematic disagreement exists between these two experiments. The disagreement is in the angular range $-0.2 > \cos\theta > -0.8$, where Brown *et al.* obtain a positive polarization and Shannon *et al.* obtain a negative polarization. As discussed by Brown *et al.*, a likely explanation for the discrepancy is that this momentum was incorrectly calibrated in one or both of the experiments. The energy dependence of the polarization is such that either a downward shift of the Brown *et al.* momentum by 30 MeV/c or an upward shift of the Shannon *et al.* momentum by 30 MeV/c would make the two data sets more compatible. Our amalgamation program does in fact resolve the discrepancy along these lines, producing amalgamated data which lie between the two experiments in the contested angular region, and obtaining recalibrated momenta of 1010 MeV/c and 1050 MeV/c for Brown *et al.* and Shannon *et al.*, respectively. However, the amalgamated data also have a correlated error structure which allows for a collective correction reflecting the sensitivity of the data to the momentum calibrations. When these data are fit simultaneously with other types of data at

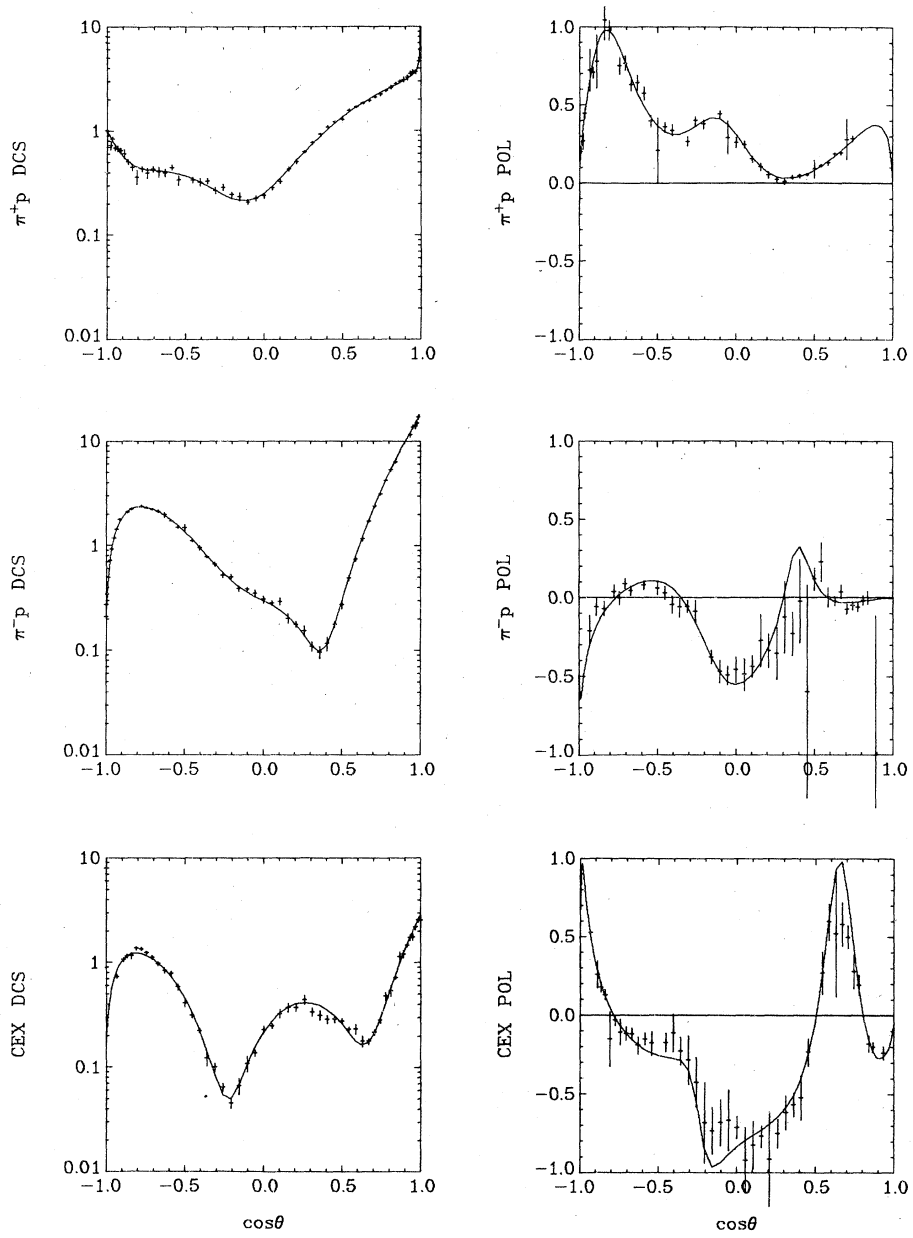

FIG. 7. Single-energy fit to amalgamated data at 1030 MeV/c. DCS units are mb/sr.

the same momentum, a large amplitude collective shift develops which contributes 19 to χ^2 and distorts the data in the backward hemisphere so that it becomes similar to the original data of Shannon *et al.* This version of the data is shown in Fig. 7. One must regard the details of this particular fit with some caution because χ^2 per datum for the charge-exchange polarization is 1.9 and the variance of the correction amplitude obtained by the amalgamation program was evidently too small, but it does illustrate rather clearly the importance of allowing for collective effects.

In Ref. 3 our partial-wave amplitudes are compared with results from other analyses. Here we only comment on a few points pertaining to the invariant amplitudes and to observables. At 0.4 GeV/c, our amplitudes match smoothly to those of Bugg.[41] In our previous publications[1] some differences with the results of Ayed[32] were noted, and these differences persist in the work reported here. The differences at the higher energies have become even more pronounced; there is a sharp discontinuity (about 5 times our estimated error) between the Ayed C amplitudes and ours at 2 GeV/c,

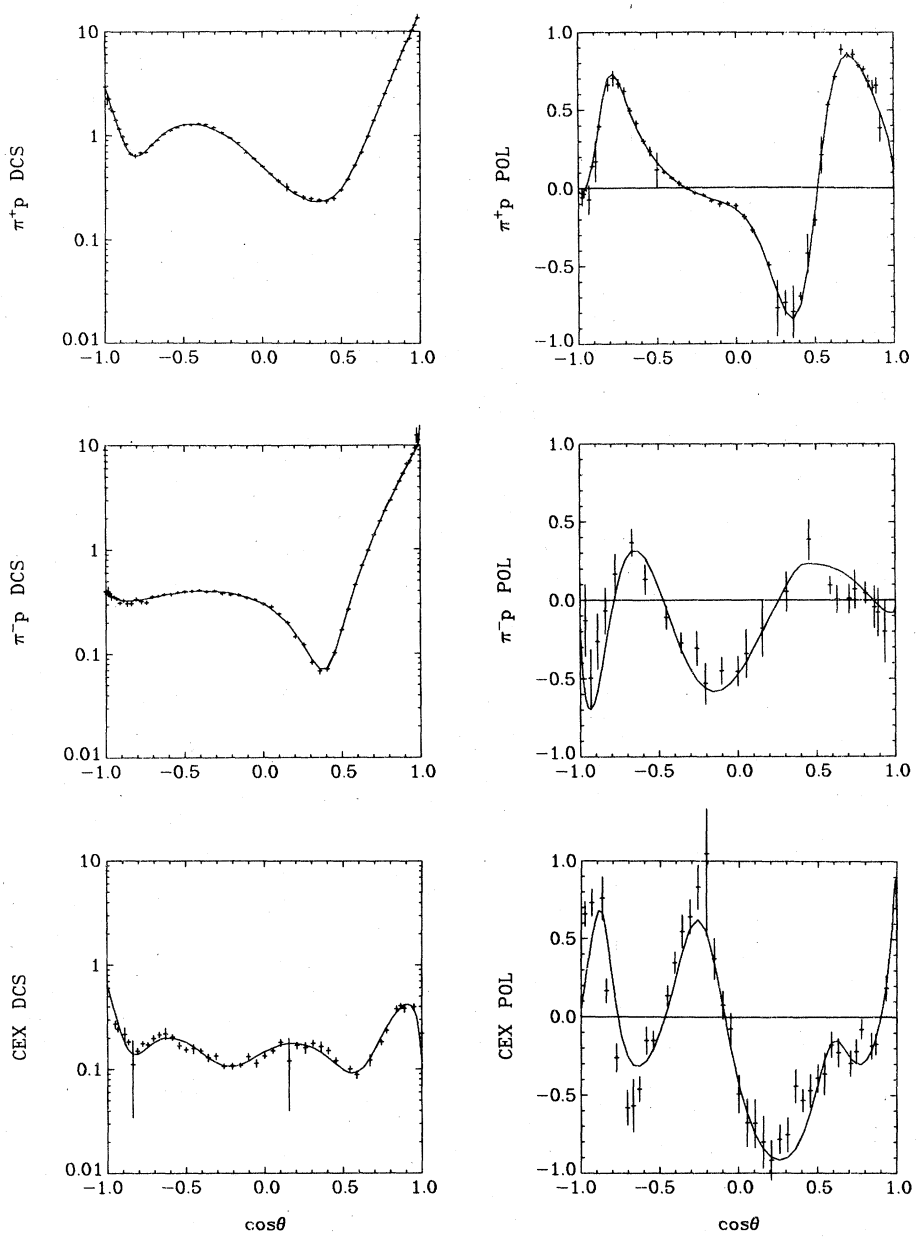

FIG. 8. Single-energy fit to amalgamated data at 1437 MeV/c. DCS units are mb/sr.

for $\cos\theta \sim 0.8$. We are in general agreement, in this respect, with Höhler *et al.*,[40] who have noted that the amplitudes of Ref. 32 are not consistent with fixed-t dispersion relations.

Our amplitudes seem generally to be in qualitative agreement with the Karlsruhe-Helsinki amplitudes.[40] In the backward direction, however, which Pietarinen has examined in detail and where we are able to make an explicit comparison with his results,[49] there are some sizable differences which appear to be related to different ways of treating the data. In Fig. 10 we show our fitted DCS at 180°, along with the DCS reconstructed from Pietarinen's backward $C^{(\pm)}$ amplitudes. The fit shown in Fig. 4 has been used to smooth and interpolate between our single-energy fits. From a study of the HDR fits, we have indications that for $P_{\rm lab} \lesssim 0.7$ GeV/c some of our fitted 180° π^+p and π^-p DCS may be too large, perhaps by as much as 0.1 mb/sr.[50] A discrepancy of this magnitude is not inconsistent with the errors in the data, at least if the effects of systematic errors are taken into account. However, the differences with Pietarinen at higher momenta are

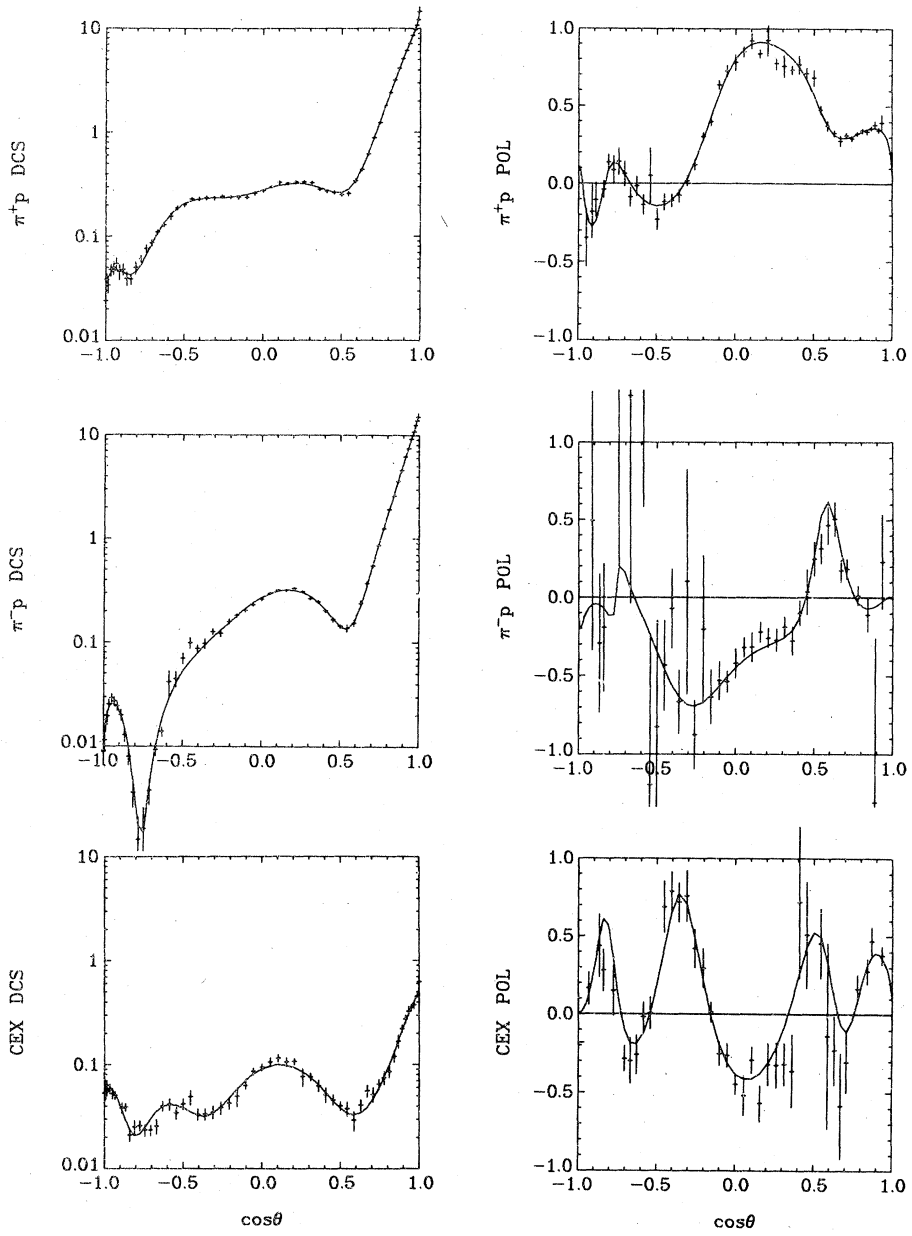

FIG. 9. Single-energy fit to amalgamated data at 1995 MeV/c. DCS units are mb/sr.

more sizable and also more serious.

The differences in the CEX DCS at $P_{\text{lab}} \sim 1$ GeV/c appear to be related to the backward dip in the DCS which can be seen in Fig. 7. These differences are therefore probably related to different ways of extrapolating from interior angles to 180°.

The differences in the π^+p DCS appear to be related, in part, to selection of different data, and in part, to a different treatment of the normalization and momentum errors in the data along with somewhat different methods of extrapolation. Pietarinen seems to have relied primarily on the 180° DCS of Rothschild *et al.*,[51] and he has not used the recent Bristol-Southampton-Rutherford (BSR) data.[6] In forming the amalgamated data, we have found that the $\pi^{\pm}p$ data of Rothschild *et al.* are not consistent with the BSR data or with most other data near the backward direction,[8, 9, 52-54] but can be made consistent if an upward shift of several percent is allowed for in both the π^+ and π^- beam momenta. As we discussed above in connection with CEX POL data at

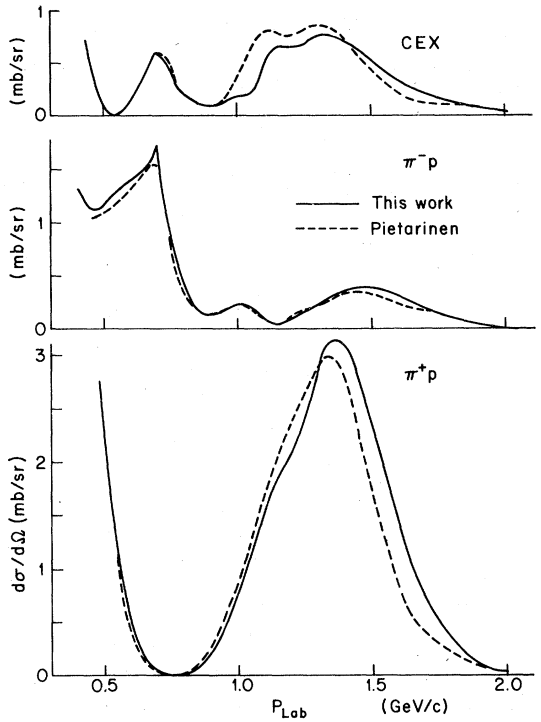

FIG. 10. Fitted differential cross section ($d\sigma/d\Omega$) at 180° for charge exchange and elastic scattering. Our results are shown as solid curves, those of Pietarinen (Ref. 49) are shown as dashed curves.

1030 MeV/c, such shifts can be accommodated automatically by our fitting procedure. If a momentum shift is applied to the Pietarinen curve and the result also renormalized upward by a few percent, most of the disagreement disappears.

Throughout this energy region, the π^+p DCS near 180° shows a great deal of structure; this is illustrated in Figs. 8 and 9. Our use of the new BSR data may contribute to the differences we have found in this structure. Different methods of fitting can also lead to different shapes for the fitted structure, which can have a substantial effect on the value obtained at 180°. The near-backward region of π^+p and CEX scattering is sensitive to the presence of the u-channel nucleon pole, which we have included explicitly in our single-energy fits. This is a possible reason why, in comparison with Pietarinen, our extrapolations seem to be more different for π^+p and CEX than for π^-p.

Pietarinen's 180° amplitudes are also quite different from ours in phase. Given the differences shown in Fig. 10, this is to be expected, as we have both used dispersion relations which have the effect of relating moduli and phases.

The most significant conclusion we are able to draw from the quality of our single-energy fits is that it is quantitatively difficult to fit the charge-exchange cross-section and polarization data of Brown *et al.*, and that it is difficult to reconcile these data with existing elastic-scattering data. Although these data represent a tremendous advance in our knowledge of the charge-exchange process, we believe that more data of comparable precision are needed to clarify the present situation. Concerning elastic measurements, the weakest area in the existing data is π^-p polarization which has not yet been subjected to an exhaustive modern measurement of the sort now available for $\pi^\pm p$ cross sections and π^+p polarization. Finally, one should not lose sight of the continuing need for spin-rotation parameter measurements[55] which have not yet been attempted anywhere in the resonance region.

ACKNOWLEDGMENTS

We thank Michael Levine for his invaluable assistance in developing a communication link between CMU and the computers at LBL, and for his helpful advice on many computational physics problems. We gratefully acknowledge the help given by Bhibuti Deo, Richard Hicks, and Chia-Chiang Shih in developing (for K^+p analysis) earlier versions of the ACE techniques and computer programs. This work was supported by the Division of High Energy Physics of the U. S. Department of Energy under Contracts Nos. EY-76-02-3066 and W-7405-ENG-48 and by the Elementary Particle Physics Program of the U. S. National Science Foundation under Grant No. PHY76-21097.

APPENDIX A: NOTATION AND CONVENTIONS

Except where otherwise specified, we use units in which $\hbar = c = 1$ and express all quantities in GeV units. The nucleon mass is denoted by M and the pion mass by μ.

The general form of the πN scattering amplitude in terms of the invariant amplitudes A and B is[56]

$$M_{fi} = \bar{u}_f(p_2)(A + B\gamma \cdot Q)u_i(p_1), \tag{A1}$$

where $Q = \frac{1}{2}(q_1 + q_2)$ and p_1, p_2, q_1, q_2 are the initial and final nucleon and pion four-momenta. We also use in this paper the invariant amplitude C (sometimes called A'):

$$C = A + \frac{M(s-u)}{4M^2 - t}B. \tag{A2}$$

For the moment, we refer only to scattering in a single elastic channel, or in a state of definite s-channel isospin.

Evaluation of Eq. (A1) with Dirac spinors of definite helicity yields the following helicity amplitudes (we choose $\phi = 0$):

$$M_{++}=\left[A+\left(\frac{s-M^2-\mu^2}{2M}\right)B\right]\cos(\theta/2),$$
$$M_{+-}=\left[\left(\frac{s+M^2-\mu^2}{2M\sqrt{s}}\right)A+\left(\frac{s-M^2+\mu^2}{2\sqrt{s}}\right)B\right]\sin(\theta/2). \tag{A3}$$

These are related to the helicity amplitudes f_{++} of Jacob and Wick[57] by

$$f_{fi}=\frac{M}{4\pi\sqrt{s}}M_{fi}. \tag{A4}$$

Equation (A1) may be reduced to two-component form by expressing the Dirac spinors in terms of rest-frame Pauli spinors χ. In this paper we use two sets of two-component amplitudes given by

$$f_{fi}=\chi_f^\dagger(f+i\vec{\sigma}\cdot\hat{n}g)\chi_i, \tag{A5}$$

where $\hat{n}=(\vec{q}_1\times\vec{q}_2)/|\vec{q}_1\times\vec{q}_2|$, and by

$$f_{fi}=\chi_f^\dagger(f_1+\vec{\sigma}\cdot\hat{q}_2\vec{\sigma}\cdot\hat{q}_1 f_2)\chi_i. \tag{A6}$$

The relations between the various sets of amplitudes are

$$f_{++}=f\cos(\theta/2)-g\sin(\theta/2)$$
$$=(f_1+f_2)\cos(\theta/2), \tag{A7}$$
$$f_{+-}=f\sin(\theta/2)+g\cos(\theta/2)$$
$$=(f_1-f_2)\sin(\theta/2), \tag{A8}$$
$$f=f_1+f_2\cos\theta,$$
$$g=-f_2\sin\theta. \tag{A9}$$

Convenient expressions for observables in terms of these amplitudes are

$$\sigma_{\text{tot}}=\frac{\text{Im}C(0^\circ)}{p_{\text{lab}}}=\frac{\text{Im}M_{++}(0^\circ)}{p_{\text{lab}}}=\frac{4\pi}{q}\text{Im}f(0^\circ), \tag{A10}$$

$$\frac{d\sigma}{d\Omega}=|f_{++}|^2+|f_{+-}|^2=|f|^2+|g|^2, \tag{A11}$$

$$P\frac{d\sigma}{d\Omega}=2\,\text{Im}(f_{++}f_{+-}^*)=2\,\text{Im}(fg^*). \tag{A12}$$

The partial-wave expansion of the helicity amplitudes is[57]

$$f_{++}=\frac{1}{q}\sum_J(2J+1)f_{++}^J d_{++}^J, \tag{A13}$$

where

$$f_{++}^J=\frac{1}{2}(f_{J-}\pm f_{J+}), \tag{A14}$$

and $f_{J\pm}$ is the partial-wave amplitude for scattering with total angular momentum J and parity $(-1)^{L+1}$, where $L=J\pm\frac{1}{2}$. The normalization of f_{Js} is

$$f_{Js}=\frac{\eta_{Js}\exp 2i\delta_{Js}-1}{2i}. \tag{A15}$$

From Eqs. (A7) and (A8) and the relations

$$d_{++}^J=\frac{\cos(\theta/2)}{J+\frac{1}{2}}(P'_{J+1/2}-P'_{J-1/2})=\frac{P_{J-1/2}+P_{J+1/2}}{2\cos(\theta/2)},$$
$$d_{-+}^J=\frac{\sin(\theta/2)}{J+\frac{1}{2}}(P'_{J+1/2}+P'_{J-1/2})=\frac{P_{J-1/2}-P_{J+1/2}}{2\sin(\theta/2)} \tag{A16}$$

one obtains

$$f=\frac{1}{q}\sum_J(J+1/2)(f_{J-}P_{J-1/2}+f_{J+}P_{J+1/2}), \tag{A17}$$
$$g=\frac{\sin\theta}{q}\sum_J(f_{J-}P'_{J-1/2}-f_{J+}P'_{J+1/2}),$$
$$f_1=\frac{1}{q}\sum_J(f_{J-}P'_{J+1/2}-f_{J+}P'_{J-1/2}),$$
$$f_2=\frac{1}{q}\sum_J(f_{J+}P'_{J+1/2}-f_{J-}P'_{J-1/2}). \tag{A18}$$

The partial-wave projections corresponding to Eq. (A13) are

$$f_{++}^J=\frac{q}{2}\int_{-1}^{1}d(\cos\theta)f_{++}d_{++}^J. \tag{A19}$$

From this follows

$$f_{J\pm}=\frac{q}{2}\int_{-1}^{1}d(\cos\theta)(f_1P_{J\pm1/2}+f_2P_{J\mp1/2}). \tag{A20}$$

The relations between physical s-channel reaction amplitudes and s-channel isospin amplitudes for any of the amplitudes discussed above are

$$F(\pi^+p\to\pi^+p)=F_{3/2},$$
$$F(\pi^-p\to\pi^-p)=\tfrac{1}{3}F_{3/2}+\tfrac{2}{3}F_{1/2}, \tag{A21}$$
$$F(\pi^-p\to\pi^0n)=(\sqrt{2}/3)(F_{3/2}-F_{1/2}).$$

The relations between s-channel isospin amplitudes and crossed-channel isospin amplitudes used in this paper are all contained in Eqs. (2.6). In the HDR fits we use s-u crossing-symmetric amplitudes defined by

$$F(\pi^\pm p\to\pi^\pm p)=F^{(+)}\mp F^{(-)},$$
$$F(\pi^-p\to\pi^0n)=-\sqrt{2}F^{(-)}, \tag{A22}$$

where F is now A, B, or C. These amplitudes obey the crossing relations

$$A^{(\pm)}(s,u)=\pm A^{(\pm)}(u,s),$$
$$B^{(\pm)}(s,u)=\mp B^{(\pm)}(u,s), \tag{A23}$$
$$C^{(\pm)}(s,u)=\pm C^{(\pm)}(u,s).$$

Finally, we give the complete expressions for the nucleon-exchange amplitudes referred to in Sec. III B before Eqs. (3.6):

$$A^{(+)}=\frac{4Mg^2}{4M^2-t},$$
$$A^{(-)}=0,$$
$$B^{(+)}=\frac{g^2}{M^2-s}-\frac{g^2}{M^2-u},$$
$$B^{(-)}=\frac{g^2}{M^2-s}+\frac{g^2}{M^2-u}, \tag{A24}$$

where $g^2/4\pi = 14.3$.

APPENDIX B: FITS TO DATA GENERATED BY A KERNEL FUNCTION

We summarize in this appendix a number of formulas, mostly taken from previous work,[30,44,58] which pertain to the applications of reproducing kernel techniques to the fitting of experimental data. The reproducing kernel $H(t,t')$ of a Hilbert space A is a Hermitian kernel which belongs to A when considered as a function of t at fixed t'. With respect to the inner product of the Hilbert space A, the kernel function H possesses the reproducing property

$$(H(\cdot,t),f)=f(t), \quad f\in A. \tag{B1}$$

For an arbitrary linear functional $f_x \equiv I_x f$ the associated function

$$H_x(t)=[I_x H(\cdot,t)]^* \equiv H(t,\cdot)I_x^\dagger \tag{B2}$$

generates the value f_x:

$$(H_x,f)=I_x f. \tag{B3}$$

An example of H_x is given by the function $K_1(t)$, associated with the threshold derivative operator d/dq, which has been used in Eq. (3.14).

The data correspond to real and imaginary parts of invariant amplitudes F evaluated at real points $\tau<0$ on the lower side of the physical cut; we denote by τ^* the corresponding point on the opposite side of the cut. All of the functions $f\in A$ are real analytic, i.e., they take on conjugate values at τ and τ^*. [The kernel $H(t,t')$ is a real-analytic function of t for real t'; more generally, its analytic continuation satisfies $H(t,t')^*=H(t^*,t'^*)$.] The two functions

$$H_{\tau 1}(t)=\tfrac{1}{2}[H(t,\tau)+H(t,\tau^*)],$$
$$H_{\tau 2}(t)=\tfrac{1}{2}i[H(t,\tau)-H(t,\tau^*)] \tag{B4}$$

are real-analytic functions which generate the real and imaginary parts of a function $f\in A$ at the point τ. We shall henceforth label by a single index a the N data d_a. This single index a is also used to label the corresponding functions $H_a(t)$.

We choose the function $f(t)$, which is interpreted as the "best fit in A" to the data, by minimizing the quantity

$$X^2=\chi^2+\lambda\|f\|^2, \tag{B5}$$

where

$$\chi^2=\sum_{ab}(d_a-f_a)W_{ab}(d_b-f_b), \tag{B6}$$

where the truncation function is proportional to the norm

$$\|f\|^2\equiv(f,f) \tag{B7}$$

and where λ is a scale factor which we shall determine by a maximum-likelihood prescription. The function $f(t)$ which minimizes (B5) is necessarily a sum of the form

$$f(t)=\sum_a H_a(t)g_a, \tag{B8}$$

because any other function $\tilde{f}(t)$ which took on the same values at the data points (and hence led to the same χ^2) would, as a consequence of Eq. (B3), differ from $f(t)$ as given by Eq. (B8) by a function $\Delta f\equiv\tilde{f}-f$ which was orthogonal to f. We would therefore have $\|\tilde{f}\|^2>\|f\|^2$. By using (B8) and (B3), we obtain

$$\|f\|^2=\sum_{ab} g_a H_{ab} g_b, \tag{B9}$$

where

$$H_{ab}=I_a H_b(t) \tag{B10}$$

is a real symmetric matrix.

A standard technique of least squares involves expanding in functions which are orthogonalized with respect to the weights of the data. An especially convenient set of such orthogonal functions is given by the linear combinations

$$F_\alpha(t)=\sum_a H_a(t)h_{a\alpha} \tag{B11}$$

for which the values at the data points

$$F_{a\alpha}=\sum_b H_{ab}h_{b\alpha} \tag{B12}$$

satisfy the eigenvalue equation

$$\sum_{bc} H_{ab}W_{bc}F_{c\alpha}=\gamma^\alpha F_{a\alpha}. \tag{B13}$$

We normalize these functions so that

$$\sum_{ab} F_{a\alpha}W_{ab}F_{b\beta}=\delta_{\alpha\beta}. \tag{B14}$$

Their norms are then given by

$$(F_\alpha,F_\beta)=(\gamma^\alpha)^{-1}\delta_{\alpha\beta}. \tag{B15}$$

In practical applications the numerical effort can be reduced by noting that it is sufficient to calculate a limited number of the eigenvectors $F_{a\alpha}$ de-

fined by Eq. (B13), and furthermore, it is really more important in practice to satisfy the orthogonality relation (B14) than it is to have precise values for the eigenvalues γ^α.

The best-fit function is

$$f(t)=\sum_\alpha C_\alpha F_\alpha(t), \tag{B16}$$

where, by substitution into Eqs. (B5)–(B7) the C_α are required to minimize

$$X^2=\sum_\alpha (D_\alpha - C_\alpha)^2+\lambda\sum_\alpha C_\alpha{}^2/\gamma_\alpha \tag{B17}$$

with

$$D_\alpha=\sum_{ab} d_a W_{ab} F_{b\alpha}. \tag{B18}$$

This gives

$$C_\alpha=\frac{D_\alpha\gamma_\alpha}{\lambda+\gamma_\alpha}. \tag{B19}$$

Another way to obtain Eq. (B19) is to minimize the mean-square deviation from the true value

$$\bar f_a=\sum_\alpha \bar D_\alpha F_{a\alpha} \tag{B20}$$

at an arbitrary point. We consider the covariance matrix

$$\begin{aligned} E_{ab}&=\langle(f_a-\bar f_a)(f_b-\bar f_b)\rangle \\ &=\sum_{\alpha\beta} F_{a\alpha}F_{b\beta}\langle(\bar D_\alpha - C_\alpha)(\bar D_\beta - C_\beta)\rangle. \end{aligned} \tag{B21}$$

We require that the expansion coefficients C_α depend linearily on the data:

$$\begin{aligned} C_\alpha&=\sum_{\alpha'}\Gamma_{\alpha\alpha'}D_{\alpha'} \\ &=\sum_{\alpha'}\Gamma_{\alpha\alpha'}(\bar D_{\alpha'}+\Delta_{\alpha'}), \end{aligned} \tag{B22}$$

where

$$\langle\Delta_\alpha\Delta_\beta\rangle=\delta_{\alpha\beta}. \tag{B23}$$

The true expansion coefficients are of course unknown, but we may assume that in some average sense the $\bar D_\alpha$ are independent and that the individual norms also have the same scale:

$$\begin{aligned} &\langle\|\bar D_\alpha F_\alpha\|^2\rangle=\lambda^{-1}, \\ &\langle\bar D_\alpha\bar D_\beta\rangle=\delta_{\alpha\beta}\gamma_\alpha/\lambda. \end{aligned} \tag{B24}$$

(If E is averaged over the data points, the assumption that the $\bar D_\alpha$ are independent can be omitted.) Substituting Eqs. (B22)–(B24) into Eq. (B21) we obtain

$$\begin{aligned} E_{ab}=\sum_{\alpha\alpha'\beta'} &F_{a\alpha}F_{b\alpha} \\ &\times[\Gamma_{\alpha\alpha'}\Gamma_{\alpha\beta'} \\ &+(\delta_{\alpha\alpha'}-\Gamma_{\alpha\alpha'})(\delta_{\alpha\beta'}-\Gamma_{\alpha\beta'})\gamma_\alpha/\lambda]. \end{aligned} \tag{B25}$$

To minimize any diagonal value E_{aa} we require

$$\Gamma_{\alpha\alpha'}=\delta_{\alpha\alpha'}\gamma_\alpha/(\lambda+\gamma_\alpha), \tag{B26}$$

which is equivalent to Eq. (B19).

Note that our formula for the covariance matrix

$$E_{ab}=\sum_\alpha F_{a\alpha}F_{b\alpha}\gamma_\alpha/(\lambda+\gamma_\alpha) \tag{B27}$$

includes more than the effect of fluctuations in the data; if we just included these fluctuations we would have obtained $[\gamma_\alpha/(\lambda+\gamma_\alpha)]^2$ as the coefficient in (B27). The extra contributions represent an estimate of the truncation error, that is, an estimate of the degree to which Eq. (B19) has reduced the contribution of the true coefficients $\bar D_\alpha$ to the expansion.

The scale factor λ determines the effective number ν of terms used in fitting; this is, roughly, the number of terms for which $\gamma_\alpha \gtrsim \lambda$. A more precise estimate of ν is given by the expected value of χ^2:

$$\begin{aligned} \nu&=N-\langle\chi^2\rangle=N-\left\langle\sum_\alpha (D_\alpha - C_\alpha)^2\right\rangle \\ &=N-\sum_\alpha \langle D_\alpha{}^2\rangle\lambda^2/(\lambda+\gamma_\alpha)^2. \end{aligned} \tag{B28}$$

As in Eqs. (B23)–(B24),

$$\langle D_\alpha{}^2\rangle=\langle\bar D_\alpha{}^2\rangle+\langle\Delta_\alpha{}^2\rangle=(\gamma_\alpha+\lambda)/\lambda \tag{B29}$$

and hence

$$\nu=\sum_\alpha[1-\lambda/(\lambda+\gamma_\alpha)]=\sum_\alpha \gamma_\alpha/(\lambda+\gamma_\alpha). \tag{B30}$$

The value of the truncation function is obtained by using Eqs. (B15) and (B20):

$$\Phi\equiv\lambda\|f\|^2=\sum_\alpha D_\alpha{}^2\lambda\gamma_\alpha/(\lambda+\gamma_\alpha)^2. \tag{B31}$$

Using also (B29), we obtain

$$\langle\Phi\rangle=\sum_\alpha \gamma_\alpha/(\lambda+\gamma_\alpha)=\nu. \tag{B32}$$

If the γ_α are ordered in a decreasing sequence, then the terms in the sum (B32) will also decrease steadily. Because of fluctuations around the average, however, especially if the Born approximation is really a qualitatively good representation of the data, the terms of (B31) would not usually decrease steadily until $\gamma_\alpha \lesssim \lambda$. Also, if the orthogonality condition (B14) is not satisfied (as in the energy-independent fitting described in

Sec. II) we do not have such a good estimate for the effective number of terms needed to fit the data as is given by Eq. (B30). In this case we interpret ν as being, roughly, the point where the terms in the sum for Φ begin to decrease strongly, as in Eq. (2.51).

To determine the value of λ used in the HDR fits, we use the maximum-likelihood procedure suggested by Chao and Pietarinen.[59] We assume that D_α may be considered a Gaussian variable with $\langle D_\alpha{}^2\rangle$ given by Eq. (B29). The quantity $L = -2\ln$ (likelihood) is then given by

$$L=\sum_\alpha\left(\frac{D_\alpha{}^2\lambda}{\lambda+\gamma_\alpha}+\ln\frac{\lambda+\gamma_\alpha}{\gamma_\alpha}\right)+L_0\,, \tag{B33}$$

where L_0 is a constant. Sometimes, to increase λ and hence reduce ν in fitting an amplitude which seems subject to systematic errors, we subtract $\nu_0\ln\lambda$ from this expression for L, where ν_0 is a parameter characterizing the likelihood function for λ. The equation for λ, obtained by minimizing L, is

$$\lambda\sum_\alpha D_\alpha{}^2\gamma_\alpha/(\lambda+\gamma_\alpha)^2=\nu+\nu_0\,, \tag{B34}$$

where ν is given by Eq. (B30).

As an alternative to minimization of X^2 as given by (B5), one can consider two alternative definitions of the "best-fit" function F. The first alternative would be to minimize χ^2, subject to a constraint $\|f\|^2\leq M_0$. The second alternative would be to minimize $\|f\|^2$, subject to a constraint $\chi^2\leq\chi_0^2$ where χ_0^2 might be defined as corresponding to some given confidence level. These two alternatives are formally identical to minimization of X^2, with λ (in the first case) and λ^{-1} (in the second case) being a Lagrange multiplier. (Note that χ^2 and $\|f\|^2$ are monotone functions of λ.) The problem of choosing λ is then replaced by the problem of choosing M_0 (or χ_0^2). When no convincing *a priori* arguments can be given for the choice of M_0 or χ_0^2, we prefer the symmetrical formulation implied by the use of Eq. (B5) and the use of the maximum-likelihood principle.

In the single-energy fitting of Sec. II, we have, in effect, actually used a hybrid approach in choosing the scale factors in the truncation function Φ. By examining fits at a number of energies where the data were plentiful and reliable, we chose scale factors which gave reasonable confidence levels. Then we required that the scale factors be slowly varying with energy, and in fact, were able to use constant scale factors [see the discussion following Eq. (2.50)]. This is almost the same as choosing the Φ's to be slowly varying with energy, except that the Φ's are allowed to fluctuate somewhat in response to the characteristics of the data.

APPENDIX C: FORMATION OF CLUSTERS

As outlined in Sec. II E, our solution-sorting procedure identifies the distinct regions of the parameter space within which the data are well represented by the parameters; this is done by forming clusters of points which correspond to individual fits. The sorting procedure associated with a simplified single-scan technique was outlined in Sec. II E. We describe here how we characterize the location, size, and shape of each of these regions, or clusters, by constructing a central value and a covariance matrix as a weighted average over its points. Finally, we describe the more efficient double-scan procedure which is actually used in practice.

Having assigned points to clusters, we next construct the central value and covariance matrix associated with each cluster as a weighted average over its points. As a first guide toward an appropriate weighted average, consider a statistical model in which X^2 is quadratic and the parameter values within a cluster are normally distributed with a covariance matrix equal to twice the inverse of the second-derivative matrix of X^2. Let us also model our sampling procedure as the generation of uniformly distributed random points in parameter space. Estimates of the mean values and covariance matrix of the distribution are then given by

$$\bar f_{pJs}=\frac{1}{W}\sum_n w_n f^n_{pJs} \tag{C1}$$

and

$$V_{pJs,p'J's'}=\frac{1}{W}\sum_n w_n(f^n_{pJs}-\bar f_{pJs})\times(f^n_{p'J's'}-\bar f_{p'J's'})\,, \tag{C2}$$

where

$$w_n=\exp[-\tfrac{1}{2}(X_n{}^2-X^2{}_{\min})] \tag{C3}$$

and

$$W=\sum_n w_n\,. \tag{C4}$$

This simple model differs from the actual situation in several important respects. First we consider effects which can be included by modifying the weight functions w_n. An actual cluster spans a region of parameter space in which X^2 has complicated nonquadratic behavior, in fact, X^2 generally has many local minima within a cluster. As stated previously, one result of this is that the size of a cluster is not associated with a unit increment in X^2, but with a larger increment, δX^2. This scale factor can be built into the weight functions by using an exponential similar to Eq.

(C3) but with an extra factor of $(\delta X^2)^{-1}$ in the exponent. We also need to adjust the weights to take account of the Particle Data Group (PDG) type of scale factor, $s = \chi^2$ per degree of freedom. Within the limited range of X^2 significant for a single cluster, the main dependence of X^2 on s is one of simple proportionality, $X^2 \simeq cs$. To scale the cluster variance matrix by s we need to include a factor $f(s)$ in the exponent of w_n such that the variation of X^2 within the cluster $\Delta X^2 = c\Delta s$ is replaced by $\Delta(X^2 f(s)) = c\Delta s/s$. The factor f is only to be applied for $s > 1$, so we also require that $f(s) = 1$ for $s \leq 1$ and obtain

$$f(s) = \left(\frac{1+\ln s}{s}\right)\theta(s-1). \tag{C5}$$

The modified weight function which includes both the δX^2 and PDG scale factors is then

$$w_n = \exp\left\{\frac{-1}{2\delta X^2}\left[X_n^2 f(s_n) - X^2{}_{\min} f(s_{\min})\right]\right\}. \tag{C6}$$

It is also necessary to reconsider the model of the sampling procedure that lead to Eqs. (C1) and (C2). We do not simply generate uniformly distributed points in parameter space in our Monte Carlo searching procedure, but rather find the distribution of local minima reached from such a set of points by the minimizer. The general tendency of the minimizer is to move toward the center of a cluster (more precisely, toward the point of absolute minimum X^2 which lies near the center), so that (C2) will tend to underestimate the covariance matrix even when used with the corrected weight functions of Eq. (C5). The size of the underestimate depends on how successful the variable-metric minimizer usually is in moving across the cluster in a search for the absolute minimum, before getting trapped in a local minimum. Observation of the results of the Monte Carlo search procedure indicates that the corrected minimizer estimates U^n of the local covariance matrices provide a reasonably consistent measure of this effect. The fact that this is comparable to the cluster size itself is not unexpected because most of the volume in a multidimensional region is concentrated near the surface, and since the search regions are considerably larger than a single cluster we know that the minimizer must be able to move through distances comparable to several cluster "radii" in order for the clusters to be formed. To correct for this effect, we add the average of U^n to the intracluster dispersion as expressed by Eq. (C2).[60]

Finally, we increase the cluster size for clusters of low weight $W \simeq 1$ to guard against statistical fluctuations. Replacing W by $W - \frac{1}{2}$ in the denominator of the covariance matrix estimate increases the estimate by a factor of 2 for clusters of low weight, and has negligible effect if $W \gg 1$.

To sum up, we use the weight functions of Eq. (C6) in calculating cluster averages. These are used in Eq. (C1) to calculate the central point of a cluster. The algorithm used to calculate the covariance matrix of a cluster is

$$V_{pJs,p'J's'} = \frac{W}{W-\frac{1}{2}}\left(T_{pJs,p'J's'} + \overline{U}_{pJs,p'J's'}\right), \tag{C7}$$

where

$$T_{pJs,p'J's'} = \frac{1}{W}\sum_n w_n (f^n_{pJs} - \bar{f}_{pJs})(f^n_{p'J's'} - \bar{f}_{p'J's'}),$$
$$\overline{U}_{pJs,p'J's'} = \frac{1}{W}\sum_n w_n U^n_{pJs,p'J's'}. \tag{C8}$$

This completes the description of the simplified procedure in which points are first scanned and assigned to clusters, and then combined. In the procedure actually used, we increase the efficiency of the process by building up the clusters during the scan and by doing the scan in two passes in which the central parts of the clusters are formed first, and then the tails. Suppose we have scanned $n-1$ local minima on the first pass; some have been combined into clusters and some have not so that we have a collection of $k \leq n-1$ points (clusters and local minima) with which to compare the nth local minimum. On this pass the comparison with the mth point, say, is only carried out if $|X_n^2 - X_m^2| < \delta X^2$. This greatly reduces the amount of computation necessary during initial cluster formation. If none of the k previously scanned points satisfies the conditions for combination with n, then n is added to this set of points and the scan proceeds to the $(n+1)$th local minimum. If a local minimum m satisfies the conditions for combination with n, then the two minima are combined and this pair of minima is replaced by the resulting cluster. Although they will be added later as in Eq. (C7) it is necessary at this stage to keep track separately of the T and $\overline{U}$ contributions to the covariance matrix of the new cluster as well as the total weight W. The values of X^2, χ^2, s, etc., for the local minimum with the smaller value of X^2 are also recorded as the values associated with the new cluster.

Now consider what happens when we compare n to a previously formed cluster m. In the X^2 comparison X_m^2 is the previously found minimum value within the cluster. In the comparison distance calculation we replace the matrices $2u^m_{Js}$ by the corresponding submatrices of V^m. (Note that these matrices coincide for a cluster consisting of a single solution.) If the criteria for combination

of n and m are then satisfied we combine them and obtain new cluster weight, mean, and covariance matrix estimates. The new estimates are obtained using Eqs. (C9)–(C12) below for the special case in which one cluster consists of a single solution and has unit weight.

In the second pass we drop the condition $|X_n^2 - X_m^2| < \delta X^2$, compare points in sequence as in the first pass, and combine all pairs of points which satisfy (2.55). The distance is now defined with $2u^m$ and $2u^n$ replaced by the appropriate submatrices of V^m and V^n, and with cluster parameters, such as X_m^2, s_m, etc., being defined as the values obtained for the minimum-X^2 solution in each cluster. Suppose that two clusters with weights W_m and W_n satisfy the conditions for combination. Then it is easy to show from the defining equations (C1), (C4), (C6)–(C8) that the new combined cluster has weight,

$$W = W'_m + W'_n \tag{C9}$$

where

$$W'_l = W_l \exp\left\{\frac{-1}{2\delta X^2}\left[X_l^2 f(s_l) - X^2_{\min} f(s_{\min})\right]\right\}, \quad l = m \text{ or } n \tag{C10}$$

and

$$X^2_{\min} = \min(X_m^2, X_n^2), \tag{C11}$$

and that the new mean and covariance estimates are

$$\begin{aligned}
\bar{f}_{pJs} &= \frac{1}{W}(W'_m \bar{f}^m_{pJs} + W'_n \bar{f}^n_{pJs}), \\
\bar{U}_{pJs,p'J's'} &= \frac{1}{W}(W'_m \bar{U}^m_{pJs,p'J's'} + W'_n \bar{U}^n_{pJs,p'J's'}), \\
T_{pJs,p'J's'} &= \frac{1}{W}(W'_m T^m_{pJs,p'J's'} + W'_n T^n_{pJs,p'J's'}), \\
&\quad + \frac{W'_m W'_n}{W}(\bar{f}^m_{pJs} - \bar{f}^n_{pJs})(\bar{f}^m_{p'J's'} - \bar{f}^n_{p'J's'}).
\end{aligned} \tag{C12}$$

The foregoing is an essentially complete description of the cluster formation process. Certain modifications are sometimes made in special cases, mainly having to do with further subdivision of the comparison process into several different passes because of computer memory size limitations and problems with missed connections between close clusters. It is apparent that because of the sequential nature of this procedure, the results depend to some extent on the order in which the initial local minima are processed. We have verified through numerous tests that this dependence is quite weak, and that in the great majority of cases the clusters are well-defined and well-separated objects in parameter space.

---

*Now at Physics Dept., St. Bonaventure Univ., St. Bonaventure, N. Y. 14778.
†Permanent address: Physics Dept., Univ. of Bristol, Bristol, England.
‡Now at Dept. of Nuclear Science and Engineering, Carnegie-Mellon University.
§Now at Jet Propulsion Laboratory, Pasadena, California.
‖Deceased.

[1]R. E. Cutkosky *et al.*, Phys. Rev. Lett. 37, 645 (1976); R. E. Cutkosky *et al.*, in *Proceedings of the Topical Conference on Baryon Resonances*, edited by R. T. Ross and D. H. Saxon (Oxford Univ. Press, London, 1976), p. 48.
[2]R. L. Kelly and R. E. Cutkosky, preceding paper, Phys. Rev. D 20, 2782 (1979).
[3]R. E. Cutkosky *et al.*, following paper, Phys. Rev. D 20, 2839 (1979).
[4]G. C. Fox *et al.*, LBL and Caltech Report No. LBL-92/CALT-68-719 (in preparation).
[5]C. Lovelace *et al.*, LBL Report No. LBL-63, 1973 (unpublished).
[6]D. J. Bardsley *et al.*, in *Proceedings of the Topical Conference on Baryon Resonances*, edited by R. T. Ross and D. H. Saxon (Oxford Univ. Press, London, 1976), p. 66.
[7]R. J. Tapper, private communication.
[8]C. M. Hughes, Bristol University thesis, 1972 (unpublished).
[9]K. Abe *et al.*, Phys. Rev. D 10, 3556 (1974).
[10]J. F. Martin *et al.*, Nucl. Phys. B89, 253 (1975).
[11]M. G. Albrow *et al.*, Nucl. Phys. B37, 594 (1972).
[12]C. R. Cox *et al.*, Phys. Rev. 184, 1453 (1969).
[13]D. Hill *et al.*, Phys. Rev. Lett. 27, 1241 (1971).
[14]R. M. Brown *et al.*, Nucl. Phys. B117, 12 (1976); B137, 542 (1978).
[15]R. M. Brown *et al.*, Nucl. Phys. B144, 287 (1978).
[16]J. E. Nelson *et al.*, Phys. Lett. 47B, 281 (1973).
[17]S. R. Shannon *et al.*, Phys. Rev. Lett. 33, 237 (1974).
[18]R. E. Cutkosky and B. B. Deo, Phys. Rev. 174, 1859 (1968).
[19]J. W. Alcock and W. N. Cottingham, Nucl. Phys. B41, 141 (1972) and private communication.
[20]The need for high partial waves in the Born term has been exhaustively studied in a similar analysis of K^+p elastic scattering. See M. L. Griss and G. C. Fox, Phys. Rev. D 7, 74 (1973) and R. E. Cutkosky and J. C. Sandusky, *ibid.* 9, 2189 (1974).
[21]B. Tromborg *et al.*, Phys. Rev. D 15, 725 (1977).
[22]L. D. Roper *et al.*, Phys. Rev. 118, B190 (1965).
[23]C. J. Bebek *et al.*, Phys. Rev. D 13, 25 (1976).
[24]It is a common practice in practical calculations to remove a factor $e^{2i\sigma_0}$ from f_{em} and g_{em} and to simultaneously replace Σ_{Js} by $\Sigma_{Js} - \sigma_0$ in f_N and g_N. The calculation of observables is then slightly simplified because g_{em} becomes purely real. This notation is

used in Ref. 21, but is not used here. The reader can easily supply the changes in notation necessary to make Eqs. (2.41)–(2.43) consistent with this modification.

[25]A. A. Carter *et al.*, Phys. Rev. 168, 1457 (1968); A. A. Carter *et al.*, Nucl. Phys. B26, 445 (1971).

[26]T. R. Engelmann and R. E. Hendrick, Phys. Rev. D 16, 2891 (1977) and private communication.

[27]H. N. Sarma, Imperial College thesis, 1976 (unpublished); D. M. Binnie, private communication.

[28]R. Bhandari and Y.-A. Chao, Phys. Rev. D 15, 192 (1977).

[29]N. C. Debenham *et al.*, Phys. Rev. D 12, 2545 (1975).

[30]R. E. Cutkosky, Ann. Phys. (N.Y.) 54, 110 (1969); R. E. Cutkosky, in *Lectures in Theoretical Physics*, edited by A. O. Barut and W. E. Brittin (Colorado Associated University Press, Boulder, Colorado, 1973), Vol. 14A, p. 423.

[31]S. Almehed and C. Lovelace, Nucl. Phys. B40, 157 (1972).

[32]R. Ayed, University of Paris-Sud thesis, Report No. CEA-N-1921, 1976 (unpublished); private communication to Particle Data Group.

[33]F. James and M. Roos, MINUIT, CERN Computer Center Program Library No. D506 D516 (unpublished).

[34]Particle Data Group, Phys. Lett. 75B, 1 (1978).

[35]Y.-A. Chao *et al.*, Phys. Lett. 57B, 150 (1975).

[36]R. C. E. Devenish *et al.*, Nucl. Phys. B59, 237 (1973); R. G. Moorhouse *et al.*, Phys. Rev. D 9, 1 (1974); G. Knies *et al.*, *ibid.* 9, 2680 (1974); I. M. Barbour and R. L. Crawford, Nucl. Phys. B111, 358 (1976).

[37]H. Wessel, Univ. of Bonn thesis, Report No. Bonn-IR-75-17, 1975 (unpublished).

[38]P. Noelle, Prog. Theor. Phys. 60, 778 (1978); Nagoya Univ. Report No. DPNU-16-77, 1977 (unpublished).

[39]E. Pietarinen, Nucl. Phys. B59, 273 (1975).

[40]G. Höhler *et al.*, *Handbook of Pion-Nucleon Scattering*, (Fachinformationszentrum Energie, Physik, Mathematik, Karlsruhe, 1979), Physik Daten Vol. 12-1.

[41]D. V. Bugg, private communication.

[42]D. V. Bugg, A. A. Carter, and J. R. Carter, Phys. Lett. 44B, 278 (1973); W. S. Woolcock, Nucl. Phys. B75, 455 (1974).

[43]D. M. Binnie *et al.*, Phys. Rev. D 8, 2789 (1973).

[44]R. E. Cutkosky, J. Math. Phys. 14, 1231 (1973).

[45]J. W. Alcock *et al.*, Nucl. Phys. B84, 503 (1975).

[46]This remark is based on the following property of a linear least-squares fit using ν free parameters to fit data with covariance matrix V. Let U be the output covariance matrix, i.e., the covariance matrix of the values of the fitted function evaluated at the data points. Then $\mathrm{Tr}\, V^{-1}U = \nu$. This formula can be easily derived by the techniques used in Sec. V of the preceding paper (Ref. 2), where similar results are presented.

[47]The programs for single-energy fitting run on a CDC 7600 at LBL. The HDR programs run on a DEC VAX-11/780 at CMU.

[48]E. Pietarinen, Univ. of Helsinki Report No. HU-TFT-78-23, 1978 (unpublished); see also Ref. 40.

[49]E. Pietarinen, Univ. of Helsinki Report No. HU-TFT-78-13, 1978 (unpublished).

[50]Some reduction in these cross sections in the low-energy region could remove the wriggle in $\mathrm{Re}\, C(\pi^+)$ between 1550 and 1600 MeV (see Fig. 4), with negligible changes in the amplitudes above 1620 MeV (or below 1320 MeV). This modification would make the moduli of the elastic amplitudes somewhat closer to Pietarinen's in the region 450–700 MeV/c, but the phases, paradoxically, would be somewhat more different.

[51]R. E. Rothschild *et al.*, Phys. Rev. D 5, 449 (1972).

[52]P. J. Richards *et al.*, Phys. Rev. D 10, 45 (1974).

[53]J. Vavra *et al.*, Phys. Rev. D 16, 2687 (1977).

[54]R. J. Ott *et al.*, Phys. Rev. D 16, 2699 (1977).

[55]R. L. Kelly *et al.*, Phys. Rev. D 10, 2309 (1974).

[56]W. R. Frazer, *Elementary Particles* (Prentice Hall, Englewood Cliffs, N.J., 1966), p. 181.

[57]M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) 7, 404 (1959).

[58]R. E. Cutkosky *et al.*, Nucl. Phys. B102, 139 (1976).

[59]Y.-A. Chao and E. Pietarinen, Phys. Rev. Lett. 26, 1060 (1971).

[60]It is perhaps worth noting that the minimizer occasionally fails to converge, and that, though we retain points found in such searches, we expect the associated estimates U^n to be particularly unreliable. In calculating the average of U^n these points are not omitted but are given a reduced weight of $(0.3)w_n$. We ignore this technical point in the following and give only the algorithms appropriate to the case in which all searches converged.