PHYSICAL REVIEW D VOLUME 12, NUMBER 1 1 JULY 1975

Estimates of cluster size in high-energy hadron collisions

T. Ludlam*
Physics Department, Yale University, New Haven, Connecticut 06520

R. Slansky†
Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87545
(Received 31 March 1975)

We discuss the parameterization of the strength of clustering effects in hadronic final states in terms of an average cluster size which is inferred from predictions of a cluster emission model. Particular attention is paid to the general question of sensitivity and the importance of assumptions regarding the kinematic constraints and statistical independence. Evidence that the cluster sizes are significantly larger than the decay multiplicities of the familiar low-mass resonances is found in a large range of experimental data.

I. INTRODUCTION

Attempts to understand the mechanisms of high-energy hadron production—or at least to analyze in a manageable way the particle-production data which are now available at Fermilab and the CERN ISR—have recently been focused on clustering phenomena in high-energy multiparticle final states. Clusters of particles are seen directly in the quasielastic component among events of small final-state multiplicity. However, for the major fraction of the inelastic cross section, the usual projections of the data do not reveal such patterns in an obvious way, so that more sophisticated and indirect techniques are required in order to assess the degree of such structure in the population density of the multidimensional phase space.[1,2] Many of the available techniques place considerable importance on the interplay between data analysis and models. For example, in several recent analyses the strength of clustering effects in multiparticle final states is estimated by comparing the measure of final-state correlations, calculated from the data, with the predictions of a cluster emission model.[3-5] The resultant estimates of cluster size need have no bearing on the actual production of clusters; rather, they provide a parameterization of the cumulative effects of structure in phase space in terms of an average cluster size inferred from the predictions of a simple model. The hope is that the results may provide an accurate guide to more general classes of models. Thus, small cluster sizes suggest multiperipheral or statistical behavior, while large clusters leave open other possibilities. Unfortunately, the range of estimates currently to be found in the literature covers both situations, although there is general agreement that some clustering effects must be present.[2-6]

An independent emission model is appropriate for this kind of analysis in that by varying a single parameter—the average cluster size—the population density in phase space can be made to span the spectrum from a smooth to a highly structured behavior. Many of the predictions of cluster models may be obtained analytically *if* it is assumed that the emitted clusters (or particles) in each event behave as a statistically independent sample. This assumption is, of course, invalidated by the kinematic constraints on each event. It has been argued that if one examines only a subsample of the final-state particles in each event, then the effects of energy and momentum conservation can be ignored and the observed sample treated as if each particle (or cluster of particles) were drawn independently from some parent distribution. For example, if one has measured the rapidities of observed (charged) particles, one might choose to ignore in each event those particles with the largest and smallest rapidities and then proceed as if the kinematic constraints may be ignored in the remaining sample. One of the primary issues discussed here is the effect of energy-momentum conservation on estimates of the cluster size.

Interpreting measures which are constructed from the full inelastic sample (such as the two-particle correlation function) in terms of a cluster emission model presents a second difficulty. Oftentimes a measure will be more sensitive to the multiplicity distribution (e.g., f_2) than to structure in phase space of the exclusive channels composing the data sample. Of course, if one assumes a particular model is literally correct, then one must accept the connection between the cluster size and the multiplicity distribution. However, if these models are to be used as a phenomenological guide for assessing the phase-space structure, this assumption could be regarded as too strong. We discuss this further elsewhere.[7] In this paper we avoid the problem by examining only exclusive and semi-inclusive data.

Our approach to the effects of the constraints on estimates of cluster size is to examine models for cluster emission in which the kinematic constraints are fully imposed. We employ a fluctuation measure which we have previously used to study clustering in multiparticle production.[1] This measure provides a rigorous test of the hypothesis of statistical independence, and is also quite sensitive to clustering effects. We discuss the general question of sensitivity to clustering effects and show that analyses which do not take full account of the kinematic constraints tend to underestimate the true cluster size implied by the model. We compare a number of previously reported experimental analyses and emphasize their interpretation. Our principal conclusion is that the average cluster size is significantly larger than the decay multiplicities of the commonly observed resonances.

II. A MEASURE OF CLUSTERING

In this paper we measure the strengths of clustering effects with a fluctuation parameter which we have employed in previous analyses of multiparticle data.[8] In terms of the left-right fluctuation density this parameter is defined as

$$\kappa_n = \frac{3}{2n}\int dy\, p(y)\left\{\langle[n_L(y)-n_R(y)]^2\rangle - \langle n_L(y)-n_R(y)\rangle^2\right\} . \tag{1}$$

Here n is the number of observed (measured) particles in each event in the data sample. The analysis is to be carried out on samples for which this number is the same for each event—e.g., n prongs in a semi-inclusive data set or n bodies in an exclusive channel. The variable y may be any measurable kinematic quantity. For high-energy semi-inclusive data samples we have found rapidity to be a useful choice. For fully specified, exclusive final states we use center-of-mass longitudinal momentum. The functions $n_L(y)$, $n_R(y)$ are equal to, respectively, the number of particles to the left of right of y in a given event, so that $n_L(y) + n_R(y) = n$. The function $p(y)$ is the differential cross section $d\sigma_n/dy$ normalized to unit area.

As defined in Eq. (1) κ_n has the feature that if all n particles in each event are selected in a statistically independent manner from the distribution $p(y)$, then $\kappa_n = 1$, regardless of the shape of $d\sigma_n/dy$. That is, under these circumstances κ_n is a distribution-free statistic, and the value $\kappa_n = 1$ represents purely statistical fluctuations.[9] If clustering effects are present, the single-particle distribution $d\sigma_n/dy$ is not representative of individual events, but is just an average; the left-right asymmetry in each event is larger than would be obtained from statistical fluctuations alone. In this case, if we ignore the effects of energy-momentum conservation, the expected value of κ_n is greater than 1.

Thus κ_n is at least in principle sensitive to clustering effects, and we show in Secs. III and IV that it is indeed quite a sensitive measure. In an independent emission model, κ_n is primarily determined by the average cluster size. But, it has been our experience that measures which are sensitive to clustering are also sensitive to the kinematic constraints. The advantage of κ_n is that, although the conservation laws imply a large correction, the correction varies little from model to model. Even in the constrained case κ_n is approximately distribution-free.

The sensitivity to the constraints follows from the fact that the requirement of energy and momentum conservation in each event reduces the over-all statistical fluctuations, thereby decreasing κ_n. The value of κ_n calculated from a given data sample is a composite of effects of constraints, which tend to suppress the left-right fluctuations, and dynamical features of the production mechanism which may further suppress the fluctuations (such as in models with strong ordering in momentum transfers) or, in the case of independently emitted clusters, enhance them. If the magnitude of the fluctuations is primarily determined by the constraints, then the value obtained for κ_n will be less than 1. We illustrate this with a simple calculation[1]: Consider sets of n values of y for which $p(y)$ is a Gaussian distribution of unit variance and zero mean, but for which the sum in each set is constrained to be zero. In this case κ_n is less than 0.5 for all $n > 2$. If the distribution of the sums is not identically zero, but a Gaussian with standard deviation given by $\sigma^2 = n/2$, then $\kappa_n \simeq 0.6$. For statistically independent samples the distribution of the sums has $\sigma^2 = n$, and $\kappa_n = 1$. Thus as the constraint is relaxed and the n values of y become statistically independent, κ_n slowly approaches its value of unity.

As is seen in the next section, the constraints do play an important role. The values obtained for κ_n in more realistic models are in general less than unity, and even in cases where strong clustering is present, the value of κ_n may be near to or even less than unity. Note that $\kappa_n \simeq 1$ would imply very small clusters if the effects of the constraints were ignored. This behavior is not significantly changed by the attempt to remove the constraint effects by excluding from the analysis those particles in each event which carry off most of the momentum. In this sense, independent emission does not imply statistical independence of the secondaries in each event, and statistical independence cannot be assumed in estimating cluster sizes.

III. CLUSTERING IN HIGH-ENERGY SEMI-INCLUSIVE DATA

The experimental values obtained for κ_n from 200 and 300 GeV proton-proton collisions lie in the range 0.8–1.0, and are nearly independent of the charged-prong multiplicity, n.[3] In Ref. 3 these results were interpreted in terms of an independent emission model in which clusters are produced according to k_T-damped phase space with cluster decay multiplicities which vary about some average value. We use this same model here to assess the sensitivity of κ_n to the strengths of clustering effects imposed on such a model, and also to test the validity of assumptions regarding statistical independence of central-region particles in high-energy semi-inclusive data samples.

Before proceeding to the conclusions we describe the model. In order to have exact energy-momentum conservation in each event, Monte Carlo techniques are used to generate events which are then analyzed as if they were experimental data. For an n-prong event, a neutral multiplicity $n_{\rm n}$ is taken from a broad distribution of mean value $n/2$. These $n_{\rm n}+n$ particles are then assigned to clusters. The cluster sizes are drawn from a Poisson distribution of mean $\bar{n}_c$. If the first cluster is of size $(n+n_{\rm n}-1)$ or larger, then the event becomes a quasielastic event. Otherwise, clusters are added to the event until all $n+n_{\rm n}$ particles can be assigned. If the sum of sizes exceeds $n+n_{\rm n}$, the last cluster is truncated so that the total number of particles is $n+n_{\rm n}$. A cluster may consist of a single particle, but the quasielastic particles are excluded in computing the average number of particles into which the clusters decay. We then assume that each pion adds 0.6 GeV to the mass of the cluster and generates the four-vectors for the clusters assuming a transverse-momentum-damped phase-space model, and maintaining exact energy-momentum conservation. The clusters are then allowed to decay isotropically by Lorentz-invariant phase space. The mass excitation relation of 0.6 GeV/pion implies an average rms transverse momentum of 450 MeV/c. The final event then exactly conserves energy and momentum, both in the cluster and in the particle generation.

We have computed the values of κ_n for the rapidity fluctuations at 200 GeV for a k_T-damped phase space (all one-body clusters), and for the values $\bar{n}_c=3.5$, 7.0, and 11.0. As a result of the truncation of the last cluster, the true value of the average cluster size in the generated events is not $\bar{n}_c$, but varies with n as follows: For $\bar{n}_c=3.5$, the true cluster size is 3.4 for all $n\geq 6$; for $\bar{n}_c=7.0$, it varies from 5.6 for $n=6$ to 6.2 for $n=12$; and for $\bar{n}_c=11.0$, it varies from 7.2 for $n=6$ to 8.9 for $n=12$. We should emphasize that these averages do include the neutrals.

In Fig. 1 we show the results of the fluctuation analysis for samples of events generated from this model. Figure 1(a) shows κ_n as a function of n for a range of average cluster sizes. Here the analysis includes all n charged particles in each event. The results in Fig. 1(b) are for the case where the end particles—the two charged particles having the largest and smallest rapidities—have been excluded from the analysis, so the analysis is made on $n-2$ values of y per event.

The solid line on each of these plots corresponds to the no-clustering case—here single pions are emitted according to transverse-momentum-cut-

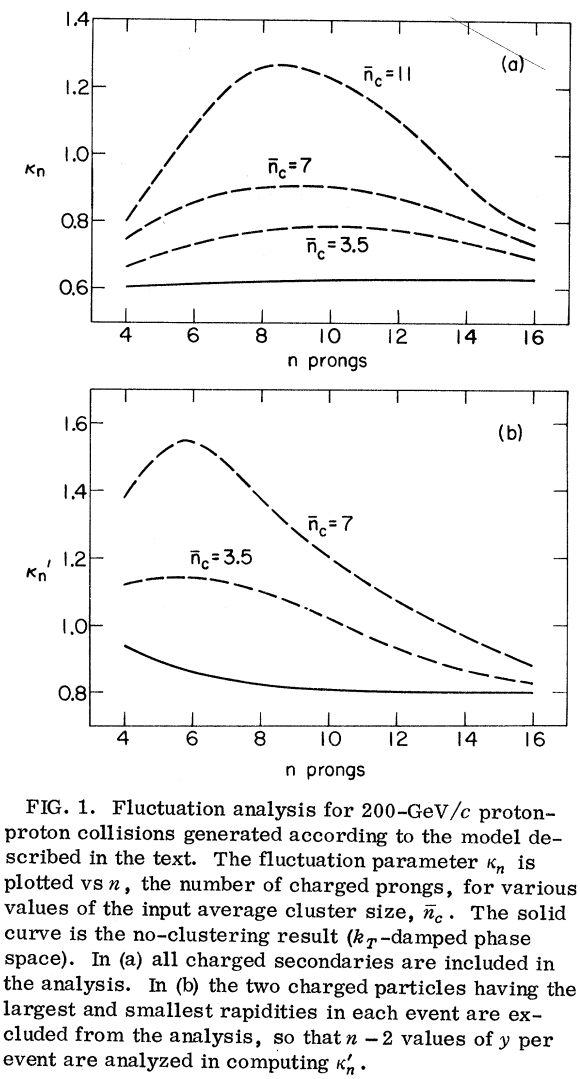

FIG. 1. Fluctuation analysis for 200-GeV/c proton-proton collisions generated according to the model described in the text. The fluctuation parameter κ_n is plotted vs n, the number of charged prongs, for various values of the input average cluster size, $\bar{n}_c$. The solid curve is the no-clustering result (k_T-damped phase space). In (a) all charged secondaries are included in the analysis. In (b) the two charged particles having the largest and smallest rapidities in each event are excluded from the analysis, so that $n-2$ values of y per event are analyzed in computing κ_n'.

off phase space. For this case, as discussed above, the value of κ_n is distribution-free. A value of κ_n which lies above this line for a particular value of n implies some degree of clustering. This is a "model independent" result, as κ_n is quite insensitive to the extent that phase space is occupied as long as the probability density is not strongly structured in this region. The dashed lines corresponding to various cluster sizes follow from the specific model we have chosen. However, two different models having similar degrees of structure in the phase-space population density will give similar values of κ_n. In this sense the specification of κ_n, or the corresponding average cluster size obtained by comparing experimental values for κ_n with the curves in Fig. 1, should be a meaningful parameterization of the strength of the clustering effects in the data.

The effects of the constraints of energy and momentum conservation on the results of the model calculation are evidently quite severe. For the 200-GeV/c pp collisions analyzed in Fig. 1(a), where all charged particles are included in the analysis, only the most extreme clustering effects, with an average of 10 or more particles per cluster, result in left-right fluctuations which are larger than the purely statistical fluctuations which would result if each final-state particle were emitted in a truly independent manner with no constraints. That is to say, only the most extreme clustering gives $\kappa_n > 1$. For the no-clustering case (solid curve), $\kappa_n \simeq 0.6$ for all charged-prong multiplicities.

In Fig. 1(b), where the end particles in each event are ignored, the values of κ_n obtained from the remaining particles still lie well below unity for the no-clustering case. The observed fluctuations are still strongly influenced by the constraints, especially in the higher multiplicities, for which the interpretation of $\kappa_n = 1$ as a no-clustering result would be grossly incorrect.

Thus, in evaluating left-right fluctuations (or other similar measures of correlations among the particles produced in each event) the assumption of statistical independence, even if one excludes from consideration the "leading" charged particles in each event, yields a systematic underestimate of the true strength of clustering effects in the production mechanism. In comparing the experimental results with the predictions of an independent emission model, this may lead to a significant underestimate of the average cluster size.

The values of κ_n obtained in Ref. 3 for semi-inclusive data samples from proton-proton scattering experiments in the Fermilab 30-in. bubble chamber clearly indicate large cluster sizes. On the basis of the analysis discussed above, an average cluster size of the order of 5 to 6 particles per cluster (3 to 4 charged particles per cluster) is inferred from values of κ_n which lie between 0.8 and 1.0. This same cluster size is found for all charged-prong multiplicities obtained for diffractively induced clusters in these experiments. The fluctuations observed in these data are not consistent with a picture in which the clustering effects result solely from independent emission of the low-lying resonances.

IV. ANALYSIS OF EXCLUSIVE CHANNELS

The parameter κ_n becomes more sensitive to specific features of the production mechanism as the kinematic constraints become stronger and can be accounted for more precisely in the model. From the technical point of view, exclusive channels are the ideal data sets for this analysis. They leave no ambiguity as to particle identification or the behavior of unseen neutrals. For these cases κ_n provides a severe test of models as well as a measure of clustering strengths. At very high energies such data sets are available for only a miniscule fraction of the inelastic final states, and are not available at all for the "interesting" final states—those for which the number of final-state particles is at or near the average value for inelastic collisions. However, for proton-proton collisions with incident beam momenta near 30 GeV/c, exclusive data are available for reactions which span the average inelastic multiplicity. In Fig. 2 we show the values of κ_n obtained from 2-, 3-, 4-, and 6-pion production data in pp and pn collisions at 28 GeV/c.[6] The reactions analyzed are

$$pp \to pp\pi^+\pi^- \,, \tag{2}$$

$$pn \to pp\pi^+\pi^-\pi^- \,, \tag{3}$$

$$pp \to pp\pi^+\pi^+\pi^-\pi^- \,, \tag{4}$$

$$pp \to pp\pi^+\pi^+\pi^+\pi^-\pi^-\pi^- \,. \tag{5}$$

For these data the fluctuation analysis is carried out in terms of the center-of-mass longitudinal momentum, rather than rapidity. In Fig. 2 the data are compared with some models relevant to the present discussion.

The lowest-lying solid curve in Fig. 2 is the no-clustering case, calculated for particles emitted according to transverse cutoff phase space. The effects of the constraints on the fluctuations of the observed particles are now maximal, and this is reflected in values of κ_n which are much smaller for this curve than for the analogous ones in Fig. 1. The data points lie well above the no-clustering values.

The open triangles in Fig. 2 give the values of

κ_n calculated from events generated according to a variation of the Reggeized multiperipheral model of Chan, Łoskiewiez, and Allison (CŁA model).[10] Two-body resonance production (ρ and Δ) has been explicitly incorporated into the model calculation in amounts which roughly reproduce the amounts seen in the data. The open boxes are results of the same CŁA calculation, but with 20% production of n-pion resonances added incoherently to the amplitude. The details of the calculation are discussed in Ref. 6. In the case of the 2-pion channel [reaction (2)] the model, which includes Pomeranchuk exchange, reproduces almost exactly the fluctuations seen in the data. Since this reaction is dominated by diffractive excitation of either the beam or target proton, this result is not unexpected. For the higher-multiplicity final states

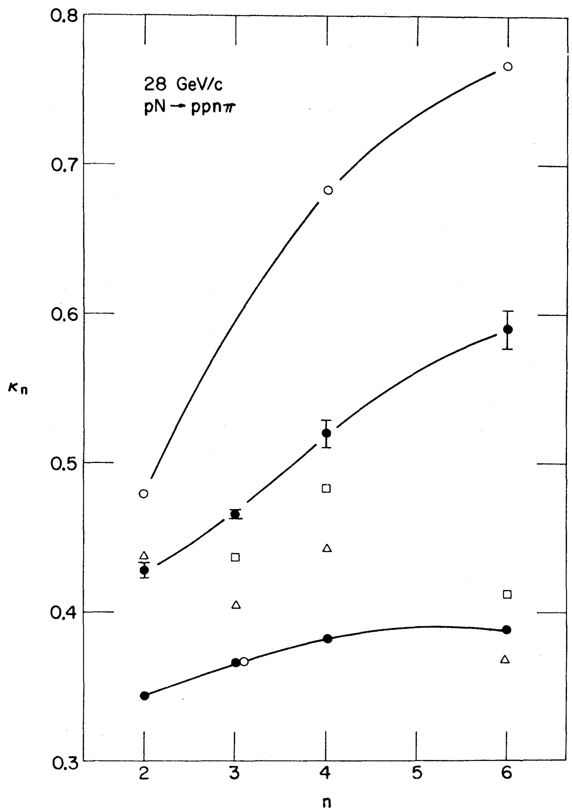

FIG. 2. Fluctuation analysis of the reactions for n pion production, $pp \to pp\pi^+\pi^-$, $pn \to pp\pi^+\pi^-\pi^-$, $pp \to pp2\pi^+2\pi^-$, $pp \to pp3\pi^+3\pi^-$ for 28-GeV/c incident momentum. κ_n is plotted vs n, the number of pions in the final state, but the analysis is done on all n +2 particles in each event. The points with error bars are from the data. The open triangles and boxes are the results of the CŁA model discussed in the text. The lower curve is the k_T-damped phase space result. The upper curve is calculated from a model for single diffractive excitation (see text).

the values of κ_n obtained from the CŁA model fall well below the data, and for the highest multiplicity [reaction (5)] the CŁA values fall *below* the "no-clustering" reference value for κ_n. This is an example of a dynamical mechanism (in this case the enforced small values of momentum-transfer along the multiperipheral chain) acting in concert with the constraints to suppress the fluctuations.

It is worth noting that the CŁA calculations employed here give a very good description of the more traditional projections of the data for reactions (2)–(5), such as invariant-mass and momentum-transfer plots. But the model clearly fails the test of the fluctuation analysis for all but reaction (2). We consider this model a very reasonable phenomenological vehicle for describing the salient features of the data in terms of the production of low-mass resonances along a multiperipheral chain, and its complete failure to reproduce the trend of κ_n in these reactions leads us again to conclude that the clustering effects exhibited by the data result from larger assemblies of final-state particles than can be accounted for by the familiar resonances.[11]

The remaining curve (open circles) in Fig. 2 is intended to set an upper boundary for the scale on which to evaluate the fluctuations observed in the data. It is calculated from a single excitation model in which every event proceeds via diffractive excitation of either the beam or the target.[12] The data are clearly not supportive of such an extreme mechanism for clustering.

V. SUMMARY

We have investigated the effects of clustering phenomena in an independent emission model for high-energy semi-inclusive production upon the left-right rapidity fluctuations of final-state particles. The parameter κ_n, which measures the strength of such fluctuations, is typical of a general class of charge and multiplicity correlation estimators which have recently been employed to assess clustering effects in high-energy data. We have shown that, because of the constraints, an independent emission mechanism for production need not imply statistical independence of observed particles in high-energy semi-inclusive data samples, even if one takes pains to analyze only those produced particles which lie in the central region of the rapidity scale. We have employed a model calculation in which the constraints are rigorously imposed, and find that the cluster size inferred from the observed fluctuations is seriously underestimated if statistical independence is assumed. The parameter κ_n is found to be an accurate estimator of the true cluster size in such models, and, more generally, is a reliable guide to the strength of collective effects in the probability density of multidimensional phase space for high-energy data samples.

In analyses of semi-inclusive data from proton-proton collisions at Fermilab energies we find average cluster sizes of the order of 6 particles per cluster (including neutrals), independent of the charged-prong multiplicity of the final state. We conclude that the clustering effects in these data cannot be accounted for by emission of the familiar low-mass resonances, but rather—if cluster emission is indeed the proper picture—by states comparable in size to the diffractively produced objects seen in the quasielastic component. These results are borne out by an analysis of 4-, 5-, 6-, and 8-body exclusive channels in 28-GeV/c pp collisions, in which the observed fluctuations are much stronger than the prediction of a model in which two- and three-body resonances are produced along a multiperipheral chain, and are otherwise consistent with the qualitative results from the higher-energy data.

---

*Work supported by ERDA under Contract No. AT(11-1) 3075.

†Work supported by ERDA under Contract No. W-7405-ENG-36.

[1]T. Ludlam and R. Slansky, Phys. Rev. D 8, 1408 (1973).

[2]A. Chao and C. Quigg, Phys. Rev. D 9, 2016 (1974).

[3]T. Ludlam *et al.*, Phys. Lett. 48B, 449 (1974).

[4]C. Quigg, P. Pirilä, and G. Thomas, Phys. Rev. Lett. 34, 290 (1975).

[5]T. Kafka *et al.*, Phys. Rev. Lett. 34, 687 (1975).

[6]J. Hanlon *et al.*, Phys. Lett. 46B, 415 (1973).

[7]T. Ludlam and R. Slansky, following paper, Phys. Rev. D 12, 65 (1975).

[8]See Refs. 1, 3, and 6. The parameter κ_n defined here differs from $\langle\omega_n^2\rangle$ of Ref. 1 by a factor of $6n$: $\kappa_n = 6n\langle\omega_n^2\rangle$. Otherwise, although κ_n is displayed in somewhat different form here, the definition is completely equivalent to the one given in these earlier papers. Here we have recast it in an attempt to make its interpretation more transparent.

[9]See, for example, J. D. Gibbons, *Non-Parametric Statistical Inference* (McGraw-Hill, New York, 1971).

[10]Chan, Hong-Mo, J. Łoskiewicz, and W. W. M. Allison, Nuovo Cimento 57A, 93 (1968).

[11]Some additional clustering may result from the symmetrization of identical particles in the amplitudes described by a model of this type. In the construction of the model we have accounted for the Bose statistics only insofar as is possible without a correct prescription for the relative phases of the contributing multiperipheral diagrams. In the spirit of the CŁA model (Ref. 10), the matrix element is an incoherent sum of diagrams.

[12]For the 3-pion production reaction only the target can be excited diffractively. These events cluster in only one region of phase space and thereby constitute a reference calculation. See Refs. 1 or 6 for a more detailed explanation.