Support vector moments of acceptance-corrected angular distributions from multi-particle experimental data

Richard T. Jones*

University of Connecticut

(Dated: May 26, 2023)

Abstract

Measured moments of multidimensional angular distributions are useful as an intermediate step in partial-wave amplitude analysis (PWA) because they they have a well-known quadratic relation to the partial wave amplitudes on the one hand, and on the other hand they are simple model-independent representations of the experimental data, at least in the limit of full experimental acceptance. However, introducing a realistic model for the acceptance can break the simplicity of how the moments are extracted from the measured data. Standard practice is to use a joint multi-parameter fit to experimental data and simulated Monte Carlo events that generally involves a selection of the moments included in the fit to match those required for a favored model interpretation. This approach hampers the clean separation between analysis of experimental data and their interpretation in terms of amplitudes. A new approach to this problem is proposed for extraction of large basis sets of acceptance-corrected moments from experimental data using only linear algebra, enabling a clean separation between analysis and interpretation of experimental data without the use of fits. Extraction of trial-model PWA amplitudes from measured angular distributions is demonstrated using mock experimental data from a simulation of the GlueX experiment, without the use of multi-parameter fits at any stage.

INTRODUCTION

The standard moment h_α is defined as an integral over the multi-body final state angular distribution for a reaction multiplied by a product of modified spherical harmonics,

$$h_\alpha = \int \tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}(\Omega_1)\ldots \frac{dN}{d\Omega_1\, d\Omega_2\ldots}\, d\Omega_1 d\Omega_2\ldots \tag{1}$$

where $\tilde{Y}_{\alpha_i}$ are the real-valued spherical harmonics which may be defined as

$$\tilde{Y}_{\ell,m} = \begin{cases} \sqrt{2}(-1)^m \Im\left[Y_\ell^{|m|}\right] & :\ m<0 \\ Y_\ell^0 & :\ m=0 \\ \sqrt{2}(-1)^m \Re\left[Y_\ell^m\right] & :\ m>0 \end{cases} \tag{2}$$

in terms of the standard complex spherical harmonics Y_ℓ^m. In Eq. 1, pairs of harmonic indices ℓ, m are abbreviated as α_i, and α is a composite index containing the α_i for all angles.

The orthogonality of the spherical harmonics allows Eq. 1 to be inverted to express the angular density as a sum over the moments.

$$\frac{dN}{d\Omega_1\, d\Omega_2\ldots} = \sum_\alpha h_\alpha\, \tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}(\Omega_1)\ldots \tag{3}$$

In the case of perfect experimental acceptance, a good statistical estimator for the moment h_α is given by

$$\hat{h}_\alpha = \frac{(4\pi)^D}{N}\sum_i^N \tilde{Y}_{\alpha_1}(\Omega_{1i})\tilde{Y}_{\alpha_2}(\Omega_{2i})\ldots \tag{4}$$

where the sum is over all measured events i in the real data sample, and Ω_{1i},$\Omega_{2i}\ldots$,Ω_{Di} are the measured angles for pair $1, 2, \ldots, D$ in event i. If the experimental acceptance is perfect then the orthogonality of the spherical harmonics guarantees that the $\hat{h}_\alpha$ are good (unbiased, minimum-variance) estimators for the true moments h_α defined in Eq. 1. Once a realistic experimental acceptance is introduced, however, many of these desirable properties are lost; the $\hat{h}_\alpha$ defined in Eq. 4 are no longer good estimators for the h_α, and the decomposition implied in Eq.3 may not be unique.

The question then arises, what sort of estimator would be suitable to use in the place of the $\hat{h}_\alpha$ in the presence of a realistic acceptance? The equation to be solved is a generalization of Eq. 3 that includes the acceptance $A(\Omega_1, \Omega_2\ldots)$.

$$\frac{dN}{d\Omega_1\, d\Omega_2\ldots} = A(\Omega_1, \Omega_2\ldots)\sum_\alpha h_\alpha\, \tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}(\Omega_1)\ldots \tag{5}$$

In what follows, I adopt a simplified matrix algebraic notation, in which $dN/d\Omega_1\, d\Omega_2 \ldots$ is written as a column vector dN_j, $j = 1 \ldots R$ whose elements are the integral dN over a small bin j in hyper-angle space. It is not the purpose here to introduce a binned analysis of experimental data, but rather to express integrals over angles as Riemann sums in order to exhibit the linear algebra underlying the problem. In this notation, the acceptance becomes a diagonal matrix $A_{j,j'}$ and the product of harmonics $\tilde{Y}_{\alpha_1}(\Omega_{1,j})\tilde{Y}_{\alpha_{2,j}}(\Omega_2)\ldots$ is written as a $R \times Q$ matrix $H_{j,\alpha}$, where Q is the number of moments that are relevant to the analysis, limited by the experimental resolution. In this notation, Eq. 5 becomes Eq. 6.

$$dN = A\ H\ h \tag{6}$$

It is not possible to solve for h in terms of the measurements dN by directly inverting Eq. 6 because the matrix A is, in general, singular. Even if the coverage of A extends to all corners of angular hyper-space, which is rarely the case, there will be regions where the acceptance is poorly understood and changing rapidly with angle. To reduce systematic errors, it is sometimes desirable to apply a fiducial cut to remove such regions from the analysis, thus creating artificial holes in the acceptance that correspond to singularities in A^{-1}.

The following sections introduce an indirect solution to the inversion of Eq. 6 by rearranging it into a form such that it can be pseudo-inverted in a stable fashion, resulting in a set of "support vector" moments related to the $\hat{h}_\alpha$ that recover the desirable properties that $\hat{h}_\alpha$ had in the case of perfect acceptance. They are linear combinations of the experimental moments $\hat{Y}_\alpha$, with linear coefficients computed from phase-space Monte Carlo simulation of the specific reaction topology under study.

These $P \leq Q$ support vector moments form a compact representation of the information encoded in the experimental angular distributions, in a form that is already corrected for acceptance, suitable for release to the broader scientific community, where they can be studied using varied approaches to amplitude analysis. All necessary information related to the experimental acceptance is encoded in the $P \times Q$ matrix of support vector moment coefficients, that would be published together with the experimental moments and their statistical error matrix.

METHOD

Direct solution by inversion of Eq. 6 proceeds by multiplying both sides by the transpose of H,

$$H^T\, dN = H^T\; A\; H\; h = M\; h \tag{7}$$

where the $Q \times Q$ matrix $M = H^T\; A\; H$ is the acceptance matrix transformed into the moments basis. The final step is to multiply both sides by M^{-1}, recovering a solution for h in terms of the measured angular distributions in the presence of realistic acceptance.

$$h = M^{-1} H^T\, dN \tag{8}$$

This is the key result. Application of Eq. 8 to an experimental sample proceeds via two steps. First, the moments $H^T\, dN$ of the sample must be computed by summing over the accepted events multiplied by the angular moments products $\tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}\Omega_2 \ldots$ separately for the full list of moments $\alpha = 1 \ldots Q$. Second, the matrix M must be evaluated via Monte Carlo integration over the experimental acceptance within a limited kinematic region, inverted, and applied the the moments vector summed over all events within that region. The definition of M in integral form is given in Eq. 9.

$$M_{\alpha,\beta} = \int \tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}(\Omega_2)\ldots\; A(\Omega_1, \Omega_2, \ldots)\tilde{Y}_{\beta_1}(\Omega_1)\tilde{Y}_{\beta_2}(\Omega_2)\ldots\; d\Omega_1\, d\Omega_2 \ldots \tag{9}$$

whose estimator via Monte Carlo integration is given by Eq. 10.

$$\hat{M}_{\alpha,\beta} = \frac{(4\pi)^D}{N_{MC}} \sum_i^{N_{MC}} \tilde{Y}_{\alpha_1}(\Omega_{1,i})\tilde{Y}_{\alpha_2}(\Omega_{2,i})\ldots\; \tilde{Y}_{\beta_1}(\Omega_{1,i})\tilde{Y}_{\beta_2}(\Omega_{2,i})\ldots \tag{10}$$

involving a sum over reconstructed events from a simulation of the detector injected with random events from an isotropic event generator for the reaction under study. In the limit of a large Monte Carlo sample, the Monte Carlo sample $\hat{M}_{\alpha,\beta}$ converges to $M_{\alpha,\beta}$. The statistical error on the estimator for a given accepted Monte Carlo sample size N_{MC} is evaluated in the standard way, assuming independence of the Monte Carlo events. To give an accurate model of the acceptance in each kinematic bin, the generator would need to match the distributions in the real data for any kinematic quantities that are not included as decay angles in the definition of the moments. Other than that, there is no dependence between the real data analysis and the computation of the M matrix.

POORLY DETERMINED MOMENTS

While direct application of Eq. 8 is formally correct in all contexts, it is only of practical use in limited situations where the acceptance is essentially complete. The reason for this is that M generally has one or two eigenvalues that are very small or zero, provided the moments basis is large enough to fully describe the measured data. One way around this is to restrict the moments basis to a limited set that avoids the singularities in M. This approach works with toy Monte Carlo exercises in which the number of moments in the Monte Carlo generator respects the restrictions employed in the analysis. However, with real experimental data, that approach introduces uncontrolled systematic errors and model-dependence to a data analysis. The correct and model-independent approach is to enlarge the moments basis to include as many moments as are needed to fully describe the experimental data. When this is done, the spectrum of the full M matrix for a typical multi-particle final state will generally have zeros in some regions of the multi-dimensional kinematic space. In the case of Monte Carlo estimates $\hat{M}$ a zero eigenvalue is any eigenvalue whose statistical error makes it consistent with zero.

These zeros in M arise naturally due to angular regions in A where the acceptance is either zero (eg. cuts to remove background) or very small and poorly determined. These singularities in A do not go away just because A is rewritten in the moments basis as M. The way around this issue in Eq. 8 is to do a partial inversion of M by expressing it in its singular value decomposition,

$$M = \sum_{k}^{P} e_k \, w_k v_k^T = W \; E \; V^T \tag{11}$$

where E is a diagonal $P \times P$ matrix with the non-zero eigenvalues e_k of M down the diagonal, in decreasing order, P is the number of eigenvalues of M which are non-zero, V is the $Q \times P$ matrix of column right-eigenvectors of M corresponding to the P non-zero eigenvalues, and W is the corresponding matrix of left-eigenvectors. The fact that the original acceptance matrix A has all of its eigenvalues between 0 and 1 implies the same thing for M. If degeneracies exist between the eigenvalues e_α, any choice for the ordering of the eigenvectors leads to the same final result.

The eigenvector matrices V and W are orthogonal in the sense $V^T V = W^T W = I$. Substituting the expression for M from Eq. 11 into Eq. 7 and multiplying both sides from

the left by row vector v_k^T leads to P independent equations of the form Eq. 12 for $k = 1 \ldots P$.

$$\frac{1}{e_k}\, w_k^T\, H^T\, dN = v_k^T\, h = s_k \tag{12}$$

On the left is an integral over the experimental data that is represented by a weighted sum over measured events as shown below, and on the right is a linear sum of the standard moments h_α of the acceptance-corrected angular distribution with coefficients that are derived from the experimental acceptance. The role of the real linear coefficients $v_{k,\alpha}$ is to pick out specific linear combinations of the h_α moments which are supported by the experimental acceptance, and dropping those that are rendered ambiguous by the imperfections of the acceptance. This meaning suggests the designation of *support vector moments* for the s_k.

The statistical estimators for the s_k on a sample of N reconstructed experimental events are

$$\begin{aligned} \hat{s}_k &= \frac{1}{e_k} \sum_i^N y_{k,i} \\ y_{k,i} &= \sum_\alpha^Q v_{k,\alpha} \tilde{Y}_{\alpha_1}(\Omega_{1,i}) \tilde{Y}_{\alpha_2}(\Omega_{2,i}) \ldots \end{aligned} \tag{13}$$

The estimated statistical error on $\hat{s}_k$ for the measured sample is

$$\sigma^2(\hat{s}_k) = \frac{1}{e_k^2} \sum_i^N y_{k,i}^2 \tag{14}$$

and the off-diagonal statistical error matrix elements are

$$C_{k,k'} = \frac{1}{e_k e_{k'}} \sum_i^N y_{k,i}\, y_{k',i} \tag{15}$$

DETECTOR RESOLUTION

In the description so far, the detector acceptance was treated as a simple function of angular coordinates whose value between 0 and 1 gives the probability that an interaction whose final state is described by the angular coordinates would produce a hit pattern in the detector leading to a reconstructed topology matching the true one, and passing all of the cuts imposed by the analysis. Indeed, the way that experimental acceptance corrections are often applied seems to suggest that this is the right way to view it. However, the effects of experimental measurement errors and incompleteness cannot be simply described by an

angle-dependent efficiency, because the imperfect experimental resolution means that the reconstructed angles $\Omega'_{1,i}, \Omega'_{2,i}, \ldots$ are not the same as the true angles $\Omega_{1,i}, \Omega_{2,i}, \ldots$ for event i. The experimental design would ideally ensure that the measurement resolution in the reconstructed angles might be sufficiently small to be neglected in a moments analysis, but this cannot always be assumed.

The method described in the previous section can be readily adapted to take into account the experimental resolution, and treat it on the same footing as the detection efficiency. To see how this can be done, reconsider Eq. 5 where the acceptance function was first introduced. One can incorporate finite reconstruction errors in this equation by replacing the acceptance function $A(\Omega_1, \Omega_2, \ldots)$ with an *acceptance kernel* $A(\Omega_1, \Omega_2, \ldots; \Omega'_1, \Omega'_2 \ldots)$ which applies a convolution to the ideal angular distribution given by the moments expansion on the right-hand side in terms of the $\Omega_1, \Omega_2, \ldots$ and yields the measured angular distribution on the left-hand side in terms of the measured angles $\Omega'_1, \Omega'_2, \ldots$.

$$M_{\alpha,\beta} = \int\int \tilde{Y}_{\alpha_1}(\Omega_1)\tilde{Y}_{\alpha_2}(\Omega_2)\ldots\ A(\Omega_1, \Omega_2, \ldots; \Omega'_1, \Omega'_2, \ldots) \tilde{Y}_{\beta_1}(\Omega'_1)\tilde{Y}_{\beta_2}(\Omega'_2)\ldots\ d\Omega_1\, d\Omega_2 \ldots d\Omega'_1\, d\Omega'_2 \ldots \tag{16}$$

$$\hat{M}_{\alpha,\beta} = \frac{(4\pi)^D}{N_{MC}} \sum_i^{N_{MC}} \tilde{Y}_{\alpha_1}(\Omega_{1,i})\tilde{Y}_{\alpha_2}(\Omega_{2,i})\ldots\ \tilde{Y}_{\beta_1}(\Omega'_{1,i})\tilde{Y}_{\beta_2}(\Omega'_{2,i})\ldots \tag{17}$$

These expressions are already captured in exactly this form by Eq. 6. The only difference now is that the matrix A has off-diagonal matrix elements which are non-zero, instead of simply having acceptance factors down the diagonal. The fact that A was already represented as a matrix rather than a column vector in Eq. 6 suggests that it plays the role of a convolution kernel in this formalism. For a well-designed detector, the A matrix should be sharply peaked around the diagonal, but it need not be exactly diagonal as if the resolution were perfect, nor does the spread around the diagonal need to be uniform as a function of angle.

With this more general interpretation of Eq. 6, all of the results presented in Sect. Method cover the situation with imperfect experimental resolution with only minor modification. Because the acceptance kernel A now depends on both the ideal and the measured sets of angles, the first set of spherical harmonics to the left of A in Eq. 9 should be functions of the ideal angles $\Omega_1, \Omega_2, \ldots$ while those to the right of A should be functions of the measured

angles $\Omega'_1, \Omega'_2, \ldots$, and the integral should be over both sets of angles. This follows from the definition of M as the moments-transform of the A matrix, now that A is no longer diagonal.

PRACTICAL CONSIDERATIONS

Although the outlined procedure distinguishes only between zero and nonzero eigenvalues of M, it is rare that one would actually find an eigenvalue that is exactly zero. In a practical context, a zero eigenvalue would be one for which the statistical error on the eigenvalue arising from the Monte Carlo sampling error in the computation of M makes it consistent with zero within a reasonable confidence interval. The Monte Carlo sampling error on the eigenvalues of M can be estimated by repeating the singular value decomposition of M on $\hat{M}$ from different subsets of the isotropic event sample, and applying a bootstrap analysis. Systematic errors on the SVD results can similarly be investigated by generating different isotropic Monte Carlo samples under slightly different assumptions for the detector configuration and analysis criteria.

The definition of the support vector moments results in the $\hat{s}_k$ having zero statistical correlation for the uniform case where the angular distributions in real data are described by isotropic angular distributions before acceptance is applied. If the physical angular distributions are not uniform then the measured support vector moments will have non-zero statistical correlations, but they are expected to be small compared with what the standard h_α moments would show for the same data sample because acceptance-induced correlations are zero within the former set, but are present in the latter.

TRIAL EXERCISE

A trial with Monte Carlo data was carried out to test the effectiveness of the support vector method for extracting acceptance-corrected moments from experimental data. The reaction chosen for this study was $\gamma p \to \eta\pi^0 p$ at 9 GeV photon beam energy and fixed target (hydrogen). The η and π^0 mesons are reconstructed in their 2γ decay modes.

TABLE I: Properties of the generated $\eta\pi^0 p$ sample used in the trial.

parameter	value
total events generated	10^8
incident photon beam energy	9 GeV
total decaying to $4\gamma p$	3.94×10^7
azimuthal dependence of cross section	uniform
t-dependence of cross section	$e^{-\|t\|}$
$m(\eta\pi^0)$ dependence of cross section	broad Breit-Wigner centroid = 1.5 GeV FWHM = 3.0 GeV
meson *cm* decay angular distributions	uniform

TABLE II: Selection cuts used to form the reconstructed event sample used in the trial.

cut parameter	condition
number of charged tracks	$1 \leq n_t \leq 2$
number of neutral showers	$n_\gamma == 4$
kinematic fit confidence level	$C.L. > 0.01$
number of combinations accepted per event	1

the sample

The trial sample was based on 100 million simulated events from the reaction $\gamma p \to \eta\pi^0 p$ produced by 9 GeV photons incident on a liquid hydrogen target in the GlueX spectrometer. The kinematic properties of the generated sample are given in Table I.

The $\eta\pi^0 p$ final state was simulated and reconstructed in the decay channels $\eta \to \gamma\gamma$, $\pi^0 \to \gamma\gamma$ using the standard GlueX tools for detector simulation and final state event reconstruction. The cuts applied in final sample selection are listed in Table II. A 4C kinematic fit to the $\eta\pi^0 p$ hypothesis was applied to each reconstructed combination of $4\gamma p$, and the combination with the greatest kinematic fit confidence level was chosen from each event. Out of the original 10^8 events simulated, there were 27.1 million events that passed all of these selection cuts.

The data used for this trial are stored on a public xrootd server, and made available to anyone who would like to explore the results presented in this section on their own. Instructions are provided in the Appendix for how to download and build the software

used to generate these results on a local user platform. Using those instructions, all of the plots included below can be reproduced, and further tests carried out on this sample. Furthermore, a jupyter notebook [1] is provided containing code that generates all of the figures from these public xrootd data. A Jefferson Lab user account with Jupyterlab access is required to run this notebook. Simply download the notebook file from the link provided, start a user session on jupyterhub.jlab.org and upload, open the notebook. The notebook assumes the "ROOT (PyROOT, ROOT C++)" jupyter image has been selected for the session, and runs best on 16 cores with 4GB/core selected.

Figs. 1-2 show the kinematic distributions of the generated and reconstructed events. As expected for pure signal Monte Carlo, the amount of background in Figs. 2 is negligible.

the model

The goal of this trial is a convincing demonstration of the feasibility and power of the support vector method presented in this article, not a realistic modeling of the physics of $\eta\pi^0$ photoproduction. For this purpose, a model was created using an unconstrained combination of moments with arbitrary analytic dependence of the decay angular distribution of the $\eta\pi^0$ system on the invariant mass of $\eta\pi^0$ and Mandelstam t. The quantity R in Eq. 18 multiplies the internal phase space factor of the Monte Carlo generator to form the simulated production/decay cross section for this trial. The photon beam is unpolarized, so the only structure in the angular distributions is in the center of mass angles of the $\eta\pi^0$ system. For these, the Gottfried-Jackson frame was used.

$$R(\theta, \phi, m_X, t) = \left| \sum_{L=0}^{3} \sum_{M=-L}^{L} c_{LM}(m_X, t) Y_M^L(\theta, \phi) \right|^2 \tag{18}$$

where m_X refers to the invariant mass of the $\eta\pi^0$ system. The 16 functions $c_{LM}(m_X, t)$ are arbitrarily chosen complex analytic functions formed from products of gaussian functions of m_X multiplied by polynomials in t. For more details on the model, see the `model1_density` method of `DSelector_etapi0_moments.C` referenced in the Appendix.

Fig. 3 shows the first four moments in the real spherical harmonics basis (Eq. 2) of the trial model as functions of m_X for three different choices of Mandelstam t. The remaining 45 non-zero moments are plotted on a single graph in Fig. 4. The point highlighted by this figure is that the different moments span a wide range in magnitude, not just as functions of m_X

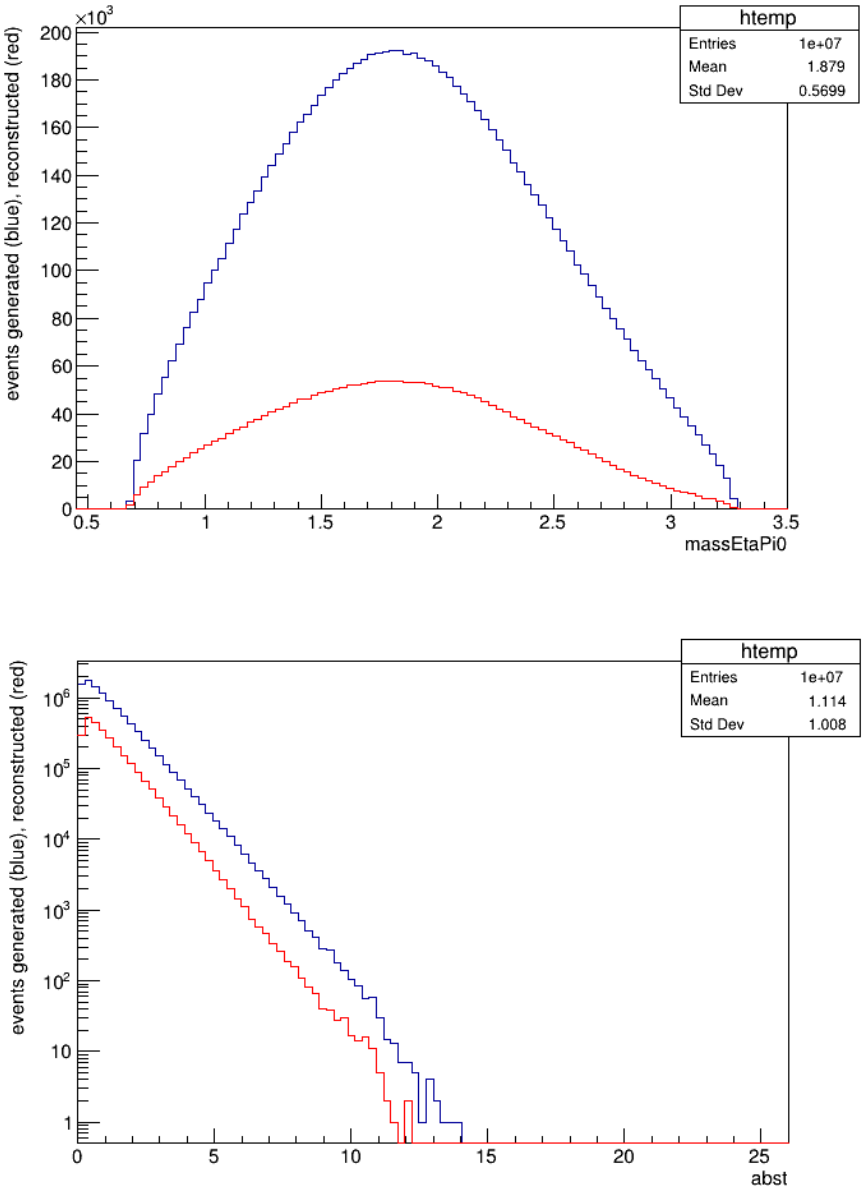

FIG. 1: Invariant mass (left) and Mandelstam t (right) of the $\eta\pi^0$ system for generated (navy) and reconstructed (red) reconstructed events in the trial Monte Carlo sample, based on 10% of the sample.

and t, but also in comparison with each other at fixed m_X, t. The challenge in this exercise is to see whether the small moments are accurately recovered by the acceptance correction procedure, or whether they are biased by feed-through from more dominant components.

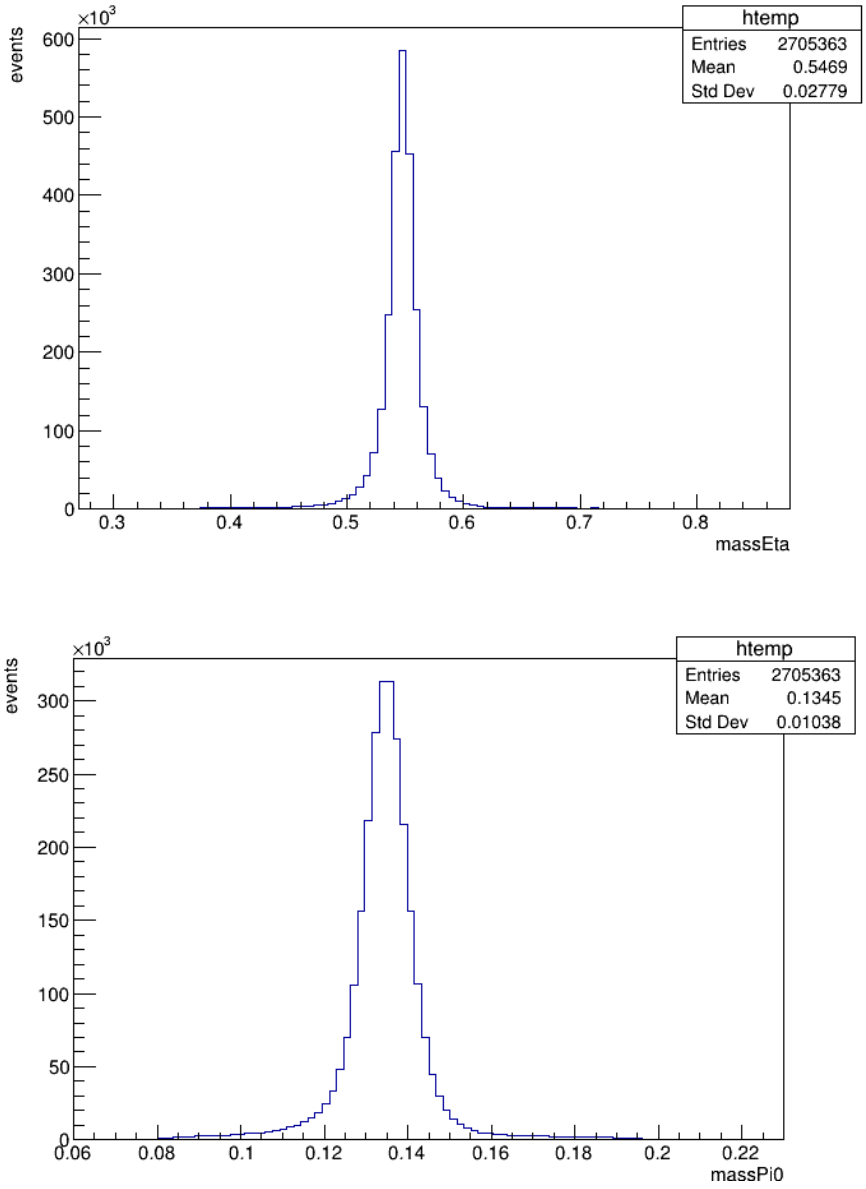

FIG. 2: Measured invariant mass distribution of the η meson (left) and π^0 (right) in the reconstructed trial sample, based on 10% of the sample.

the trial sample

The full set of 100 million generated $\eta\pi^0$ events was divided into two equal parts of 50 million events each. The first half of the events were used to produce a mock data sample by the procedure described in this section, while the second half was reserved for acceptance estimation, as described in a following section. The first 50 million simulated $\eta\pi^0$ events were partitioned into a bins laid out on a regular grid on the m_X, t plane spanning the range

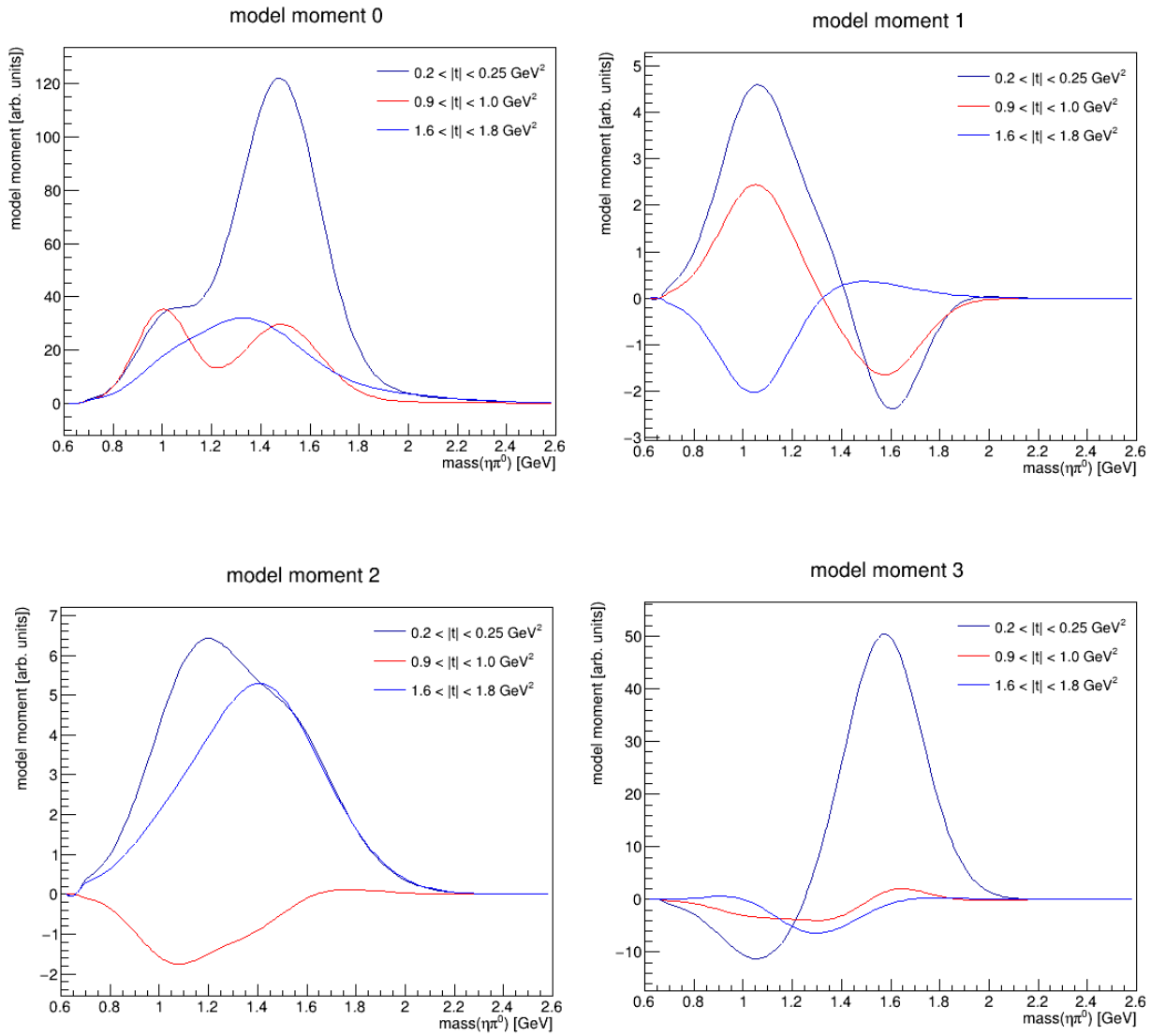

FIG. 3: Angular moments of the $\eta\pi^0$ decay distribution used in the trial, where moment 0 refers to $L = 0, M = 0$, moment 1 to $L = 1, M = -1$, moment 2 to $L = 1, M = 0$, and moment 3 to $L = 1, M = 1$.

(0.6,2.4) GeV in m_X and (0,2.5) GeV^2 in $|t|$. Each event was assigned a weight R evaluated according to Eq. 18 using the generated (not the reconstructed) angles θ, ϕ in the Gottfried-Jackson frame. Events were then sorted in bins in m_X, t, and in each bin a maximum event weight was found for the generated events. This maximum R value, different for each bin, was then used as the cut-off in a random accept-reject procedure to select events belonging to the mock experimental data sample for that bin.

The reader may wonder why this procedure was carried out bin-by-bin, rather than using

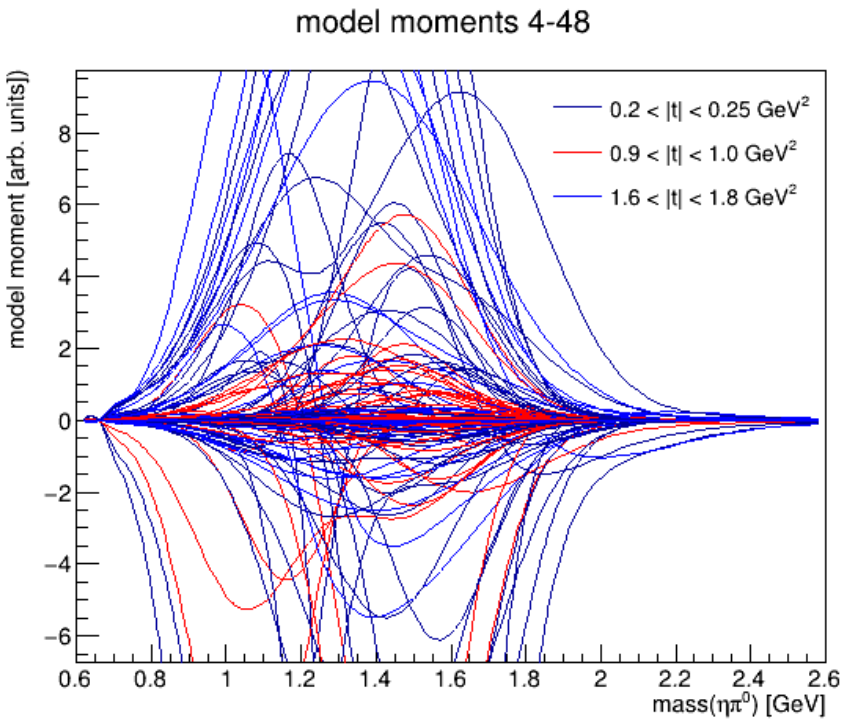

FIG. 4: Angular moments of the $\eta\pi^0$ decay distribution used in the trial all on a common scale, including all moments with $L = 2, 3, 4, 5, 6$.

a single maximum weight for the accept-reject cut-off for all bins. The reason for this choice was to allow the trial sample to explore all regions in the m_X, t plane, including those where the value of R is a few percent of its maximum value that appears around $m_X = 1.5, |t| = 0.3$. This way the trial is sensitive to behaviors of the moments far from the center of the kinematic plot, where deviations might be suspected to be most likely to occur.

The procedure described above would be expected to produce discontinuities in plots of sample moments vs m_X and t. To avoid this, all plots of sample moments vs m_X and t shown below are produced by forming the plots individually in each bin and them summing them up, each multiplied by the maximum weight that was used for the sample cut-off in that bin. Examples of this are shown in Fig. 5 where the combined event distribution is shown for reconstructed mock experimental data in their m_X and $|t|$ projections.

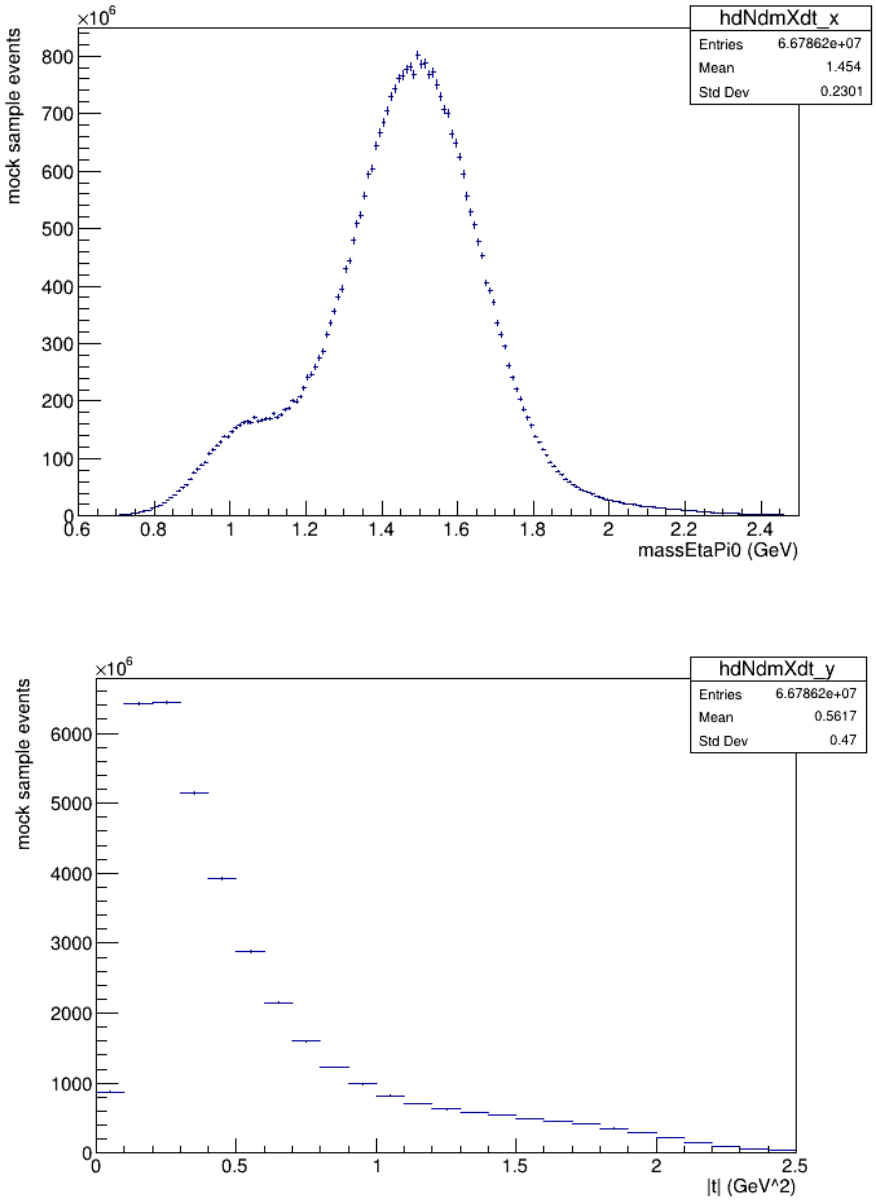

FIG. 5: Distributions of the reconstructed events from the mock experimental data sample projected onto m_X (top panel) and Mandelstam $|t|$ (bottom panel).

choice of a moments basis

In order to be computed numerically, the sum in Eq. 3 must be truncated. This truncation defines the basis of moments in which the experimental acceptance will be evaluated for the reaction of interest, so that acceptance-corrected moments can be extracted from a data sample. In the case of the trial sample, this basis must include all moments up to $L = 6$ because of how the trial model was constructed (see Eq. 18. But there are good reasons

to extend the basis used for the acceptance correction well past the upper limit where one expects significant components to be present in the signal. One reason is that backgrounds that may be present together with the signal in the data sample are more likely to be spread out over many moments. Measuring a large number of moments where signal is expected to be zero is a means for estimating the amount of background present in the sample. A second reason, particularly relevant in the specific case of the $\eta\pi^0$ reaction, is that the signal itself may contain contributions from processes that preferentially contribute to high-L moments, such as multi-peripheral production, multi-Regge exchange contributions, etc.

In the case of $\eta\pi^0 p$, evidence of such processes is seen in the Gottfried-Jackson angular distributions in the region $m_X > 1.8$ GeV. Furthermore, the fact that it is dominant in the higher-mass region suggests that it is likely to be present in the intermediate mass region $1.4 - 1.8$ GeV as well, even though it is less dominant there. Including many higher-L moments allows these components to be extracted from the experimental data in a model-independent way, whereas truncating the sum introduces an uncontrolled bias in the extraction of the lower moments by the unwarranted assumption that the higher moments are not present in the sample. In the case of ideal acceptance, the moments are independent and unbiased, but in the case of realistic acceptance this is not the case, as will be shown in the following sections. Only moments that are included in the basis used for the acceptance correction

On the other hand, one may worry that inflating the basis much beyond the region of interest in the physics analysis is that it will result in larger statistical errors on the moments that one is most interested in measuring. This is true in principle, although the size of this effect has to be determined empirically. For this study, all moments up to and including $L = 12$ were included and acceptance-corrected moments were extracted across the full kinematic region of the study. This basis has a total of 169 moments, which means that the acceptance M matrix has dimensions 169×169. This should be considered small, and imposes no computational constraints whatsoever on the moments extraction process, from considerations of memory allocation, cpu time, rounding error, etc.

The option of including higher moments in the individual decay angular distributions for the η and π^0 was also explored. If there were non-signal or non-resonant backgrounds present in either of these subsystems, that would make sense to do. In a preliminary investigation, a basis including all moments up to $L = 6$ in all three angle pairs: $\eta\pi^0$, η, and π^0 was adopted.

TABLE III: Kinematic binning schemes used in the analysis of the trial sample.

scheme	bin count	$\|t\|$ ranges (GeV^2)	$m(\eta, \pi^0)$ ranges (GeV)
coarse	24	0.0-0.3	0.6-0.9
		0.3-0.6	0.9-1.2
		0.6-1.2	1.2-1.5
		1.2-2.4	1.5-1.8
			1.8-2.1
			2.1-2.5
intermediate	48	0.0-0.3	0.6-0.75
		0.3-0.6	0.75-0.9
		0.6-1.2	0.9-1.05
		1.2-2.4	1.05-1.2
			1.2-1.35
			1.35-1.5
			1.5-1.65
			1.65-1.8
			1.8-1.95
			1.95-2.1
			2.1-2.25
			2.35-2.4
fine	144	0.0-0.2	0.6-0.7
		0.2-0.3	0.7-0.8
		0.3-0.45	0.8-0.9
		0.45-0.6	0.9-1.0
		0.6-0.9	1.0-1.1
		0.9-1.2	1.1-1.2
		1.2-1.6	1.2-1.3
		1.6-2.4	1.3-1.4
			1.4-1.5
			1.5-1.6
			1.6-1.7
			1.7-1.8
			1.8-1.9
			1.9-2.0
			2.0-2.1
			2.1-2.2
			2.2-2.3
			2.3-2.4

Because of the symmetry in the η and π^0 decays, only event L values can contribute for those subsystems, leading to a total of $28 \times 28 \times 49 = 38416$ moments in hyper-angle space. The software used for the trial sample study was originally developed and tested on this larger moments basis, and the results were consistent with what is shown here. However,

seeing that there is no significant background under the η and π^0 peaks in the trial sample, spending so much time extracting $38416 - 49 = 38367$ moments that are all consistent with zero did not seem worthwhile, and it was decided that extending the basis in the $\eta\pi^0$ system beyond $L = 6$ and truncating the η and π^0 bases at $L = 0$, the only moment present in the trial signal, would be more meaningful.

acceptance estimation

Everything that is relevant for the extraction of acceptance-corrected moments from the trial sample is contained in the matrix M defined in Eq. 17 whose Monte Carlo estimator is given in Eq. 17. Instances of the M matrix were evaluated individually together with its statistical error for each kinematic bin in the m_X, t plane, 24 bins in the coarse-binned scheme, 48 bins in the intermediate scheme, and 144 bins in the fine-binned scheme (see Table III). One way to visualize the results is to plot the eigenvalue spectrum. Fig. 6 shows the eigenvalues of M ordered from largest to smallest for a selection of kinematic bins.

These results show that, while M is non-singular, it has a few eigenvalues that are very small compared to the others. This indicates that a few linear combinations of the moments are poorly determined given the limitations imposed by the detector resolution, reconstruction efficiency, and analysis cuts. In the case of a real experimental result, these moments would also be taken with a grain of salt because they are sensitive to the detector performance and simulation fidelity in edge regions where the acceptance is small. In this trial, however, these systematic errors do not enter because the simulation is the common source of both the experimental data sample and the source of information about the acceptance. Hence the trial results are shown for all moments, including contributions from moments that have very small acceptance, which means that the acceptance-corrected moments can be shown in the standard spherical basis. In reviewing these results, it must be kept in mind that the support vectors are the basic output from the method, and their un-mixing into the standard spherical basis following the acceptance correction was done to facilitate their comparison with the model inputs.

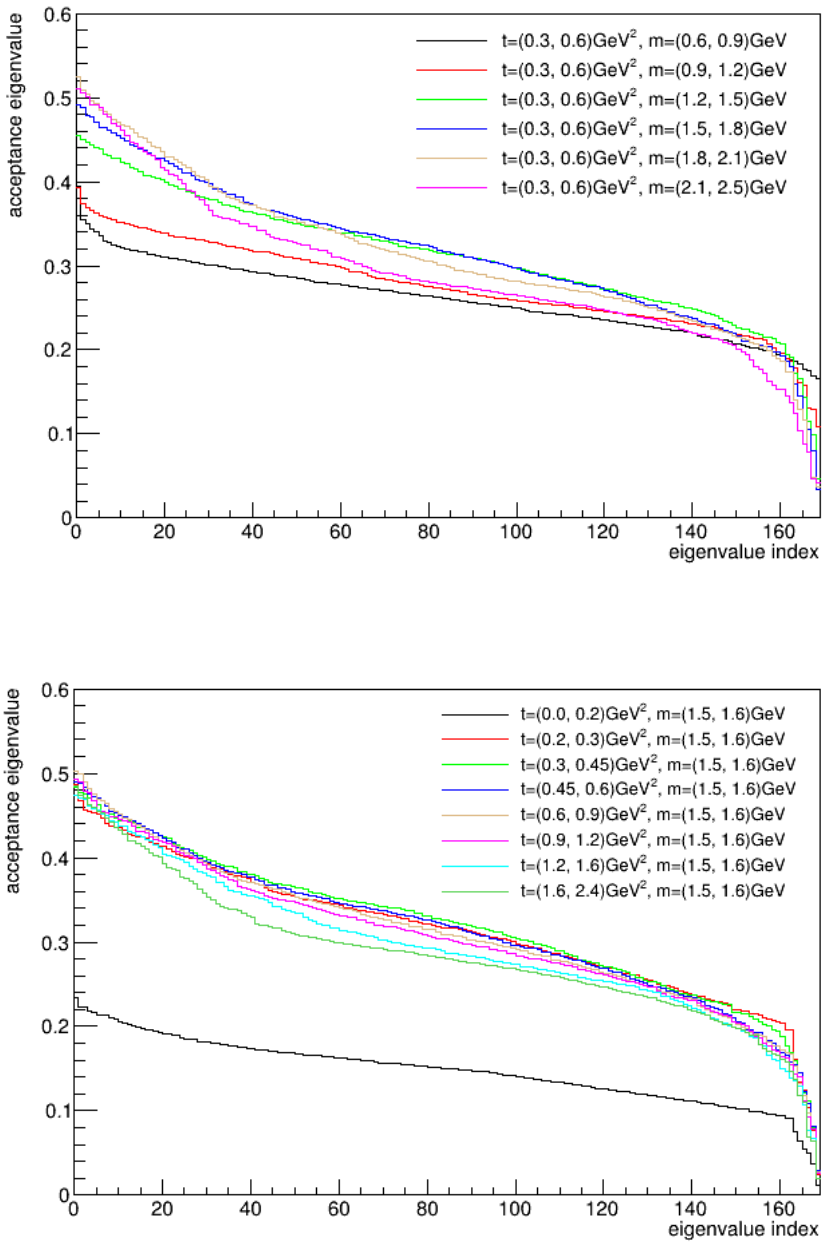

FIG. 6: Eigenvalues of the acceptance matrix M for $\eta\pi^0$ ordered from largest to smallest. The top panel shows the variation with m_X for fixed $|t|$ in the range (0.3,0.6) GeV^2, while the bottom panel shows the variation with $|t|$ for fixed m_X in the range (1.5,1.6) GeV.

acceptance-corrected moments

The acceptance correction is applied by the following procedure. First the m_X, t plane in divided into a regular array of kinematic regions. In each region, the mock data sample and the acceptance matrix M are extracted from their respective Monte Carlo datasets

and recorded. The M matrix in each region is diagonalized, and the eigenvectors (support vectors) used as a basis to go back to the original Monte Carlo acceptance sample and compute the acceptance as a scalar function of (m_X, t) with finer binning within the region, one for each support vector, and normalized to the number of generated events binned the same way on the same region. These acceptances are saved as 2D histograms with errors given by their Monte Carlo integration statistics.

Turning next to the trial data sample, the support vector moments defined for each region above are computed on the mock data events and the results binned in m_X, t using the same binning scheme as was used for the acceptance eigenvalues. The real data support vector moments are then divided by the acceptance, and the results are the acceptance-corrected moments in the support vector basis. The last step is to transform those support vector moments back into the standard spherical moments basis to facilitate comparison with the trial model moments.

Results for a selection of the 169 moments from the trial data sample are shown in Fig. 7. The red data points are the input model moments, plotted here with error bars because they are averages over the bins in the (m_X, t) plane that are obtained from the input model functions by Monte Carlo integration within each bin. The light blue points are the moments before acceptance correction, rescaled by a constant factor for ready comparison with the model inputs, while the dark blue points are after the acceptance correction. These particular moments were selected for display in this figure to highlight the fact that the acceptance correction is far from simply rescaling the uncorrected moments by a smoothly varying efficiency factor; in most cases it completely changes the shape, and sometimes even the sign. Notice also that the errors on the acceptance-corrected moments are significantly larger than the uncorrected ones. This is mostly a result of the few support moments that have extremely small acceptance eigenvalues. The same results can be expressed equivalently either by reporting moments in the standard spherical basis with a covariance matrix that carries these strong correlations between the moments, or in the support moments basis where the correlations are much weaker and the statistical errors on the measured moments are correspondingly smaller, except for one or two of them.

The comparisons in Fig. 7 give a visual impression of the agreement between the acceptance-corrected and model moments. A more comprehensive, quantitative comparison is shown in Fig. 8 where the χ^2 difference between the blue and red points shown in

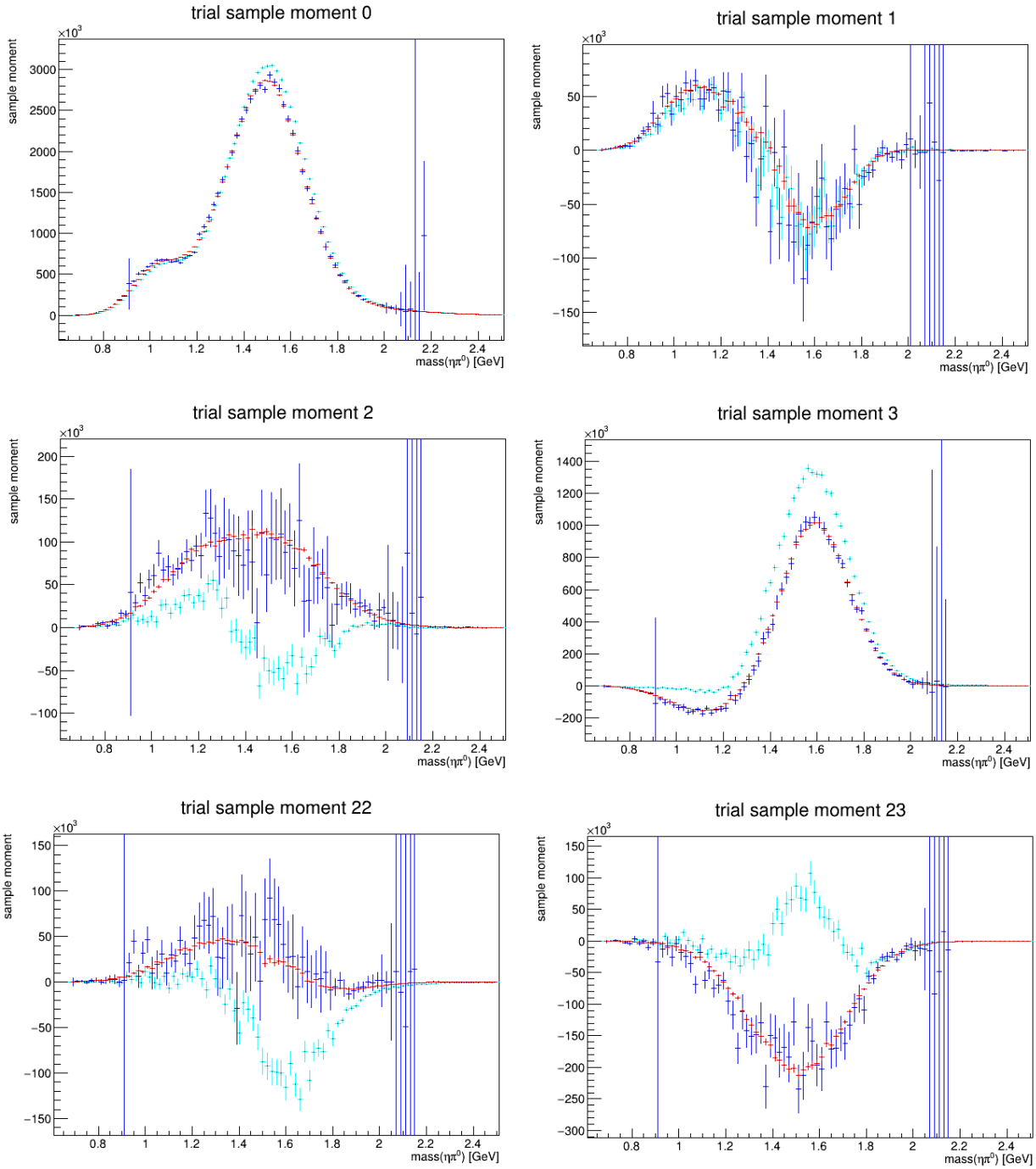

FIG. 7: Spherical moments of the trial data sample as a function of m_X for $0.3 < |t| < 0.6$ GeV^2. Moments are shown before acceptance correction (light blue), after acceptance correction (dark blue), and input model (red). Uncorrected moments have been rescaled by a global arbitrary constant common to all moments. No adjustable parameters are used in the comparison between the acceptance-corrected moments and the model.

Fig. 7 for all 169 moments are formed for all 24 bins in (m_X, t) in the coarse-binned scheme, and for 142 of the 144 bins in in the fine-binned scheme. Two of the bins in the fine-binned scheme had to be dropped for lack of statistics. The sample χ^2 means in Fig. 8 are somewhat less than the expected value of 95 because the error propagation in the conversion from standard moments to support vector basis does not take into account the correlations between the sample moments. These correlations are dominated by statistical noise at the level of individual bins at the resolution shown in Fig. 7, but can be estimated by averaging over the local region in (m_X, t), as described in the next section.

systematic error

While Eqs. 8 and following are exact solutions to a linear system, their application to an experimental dataset has systematic errors arising from the fact that the acceptance is estimated by Monte Carlo integration. In the case of this mock data trial, two separate effects contribute to this systematic uncertainty, one associated with the limited statistics of the Monte Carlo dataset, and the second arising because the acceptance is not a constant but varies somewhat across the span of the kinematic bins in m_X, t over which the Monte Carlo sum is performed.

To study these effects in the trial sample, a more precise means of quantifying the statistical errors in the comparison between the acceptance-corrected moments and the input model is needed than the method used to generate Figs.r̃effig:correctcompare-8. For that to be possible, a more coarse-grained binning must be performed on the (m_X, t) plane so that a full covariance matrix in the sample moments can be reliably extracted from the mock experimental dataset for each kinematic bin. These errors can then be propagated through the multiplication by M^-1 within each bin (see Eq. 8) to find the full covariance matrix for the acceptance-corrected moments in that same bin, assuming that the M matrix is exact. Systematic errors in the estimation of M will show up as excess χ^2 in Eq. 19.

$$\chi^2 = h^T \, C^{-1}(h) \, h \tag{19}$$

where h is the column vector of acceptance-corrected sample moments from a solution to Eq. 8 and C is the covariance matrix based on the Poisson statistics of the original event sample propagated by standard error propagation through Eq. 8 into the acceptance-corrected

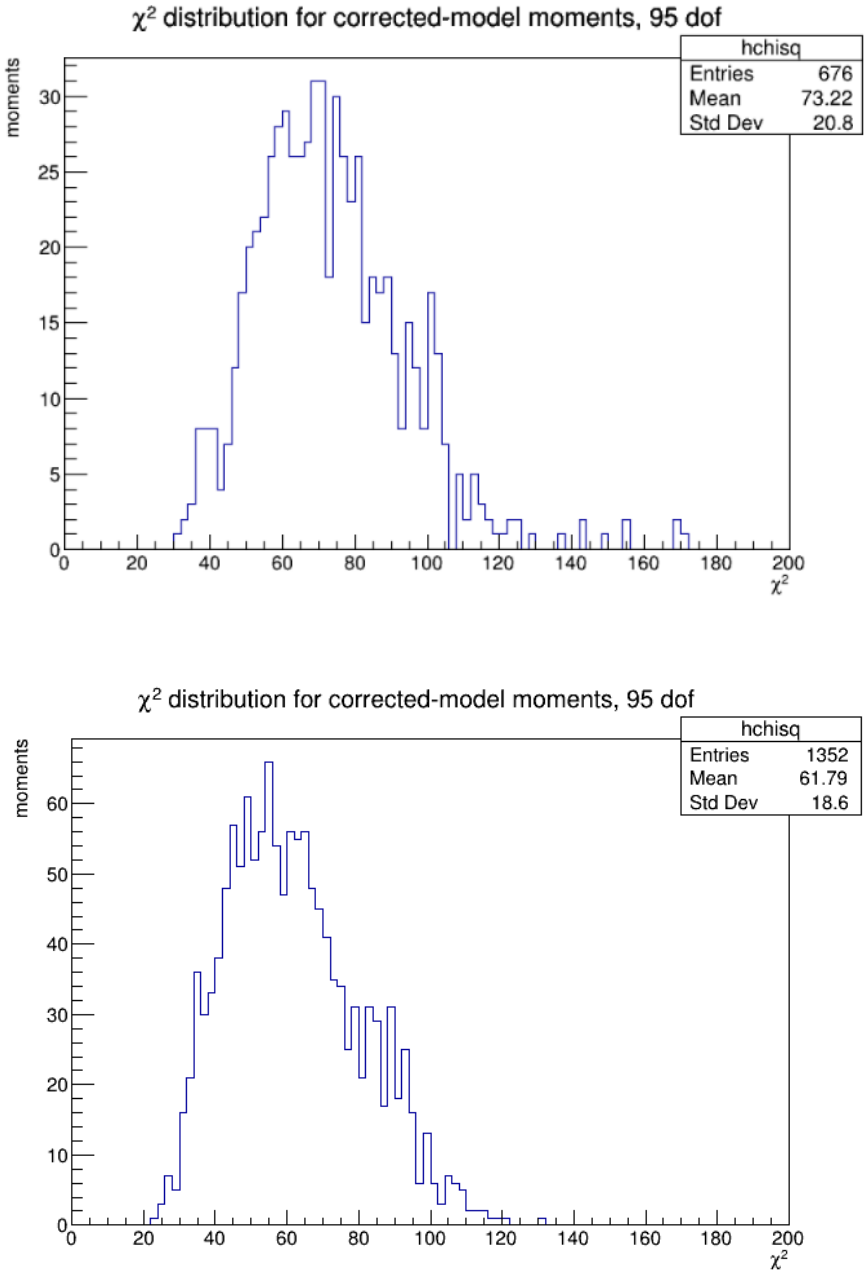

FIG. 8: Quantitative comparison between acceptance-corrected and model moments for the coarse-bin subdivision of the (m_X, t) plane (top panel) and the fine-bin subdivision (bottom panel). The χ^2 value for each moment is formed with 95 degrees of freedom.

moments basis. Fig. 9 shows the resulting χ^2 distribution based on Eq. 19 for a basis with 169 moments. If there were no systematic error present, the average χ^2 value in Fig. 9 should be 169.

Variation of the exact M within a single kinematic bin argues for making the bins smaller.

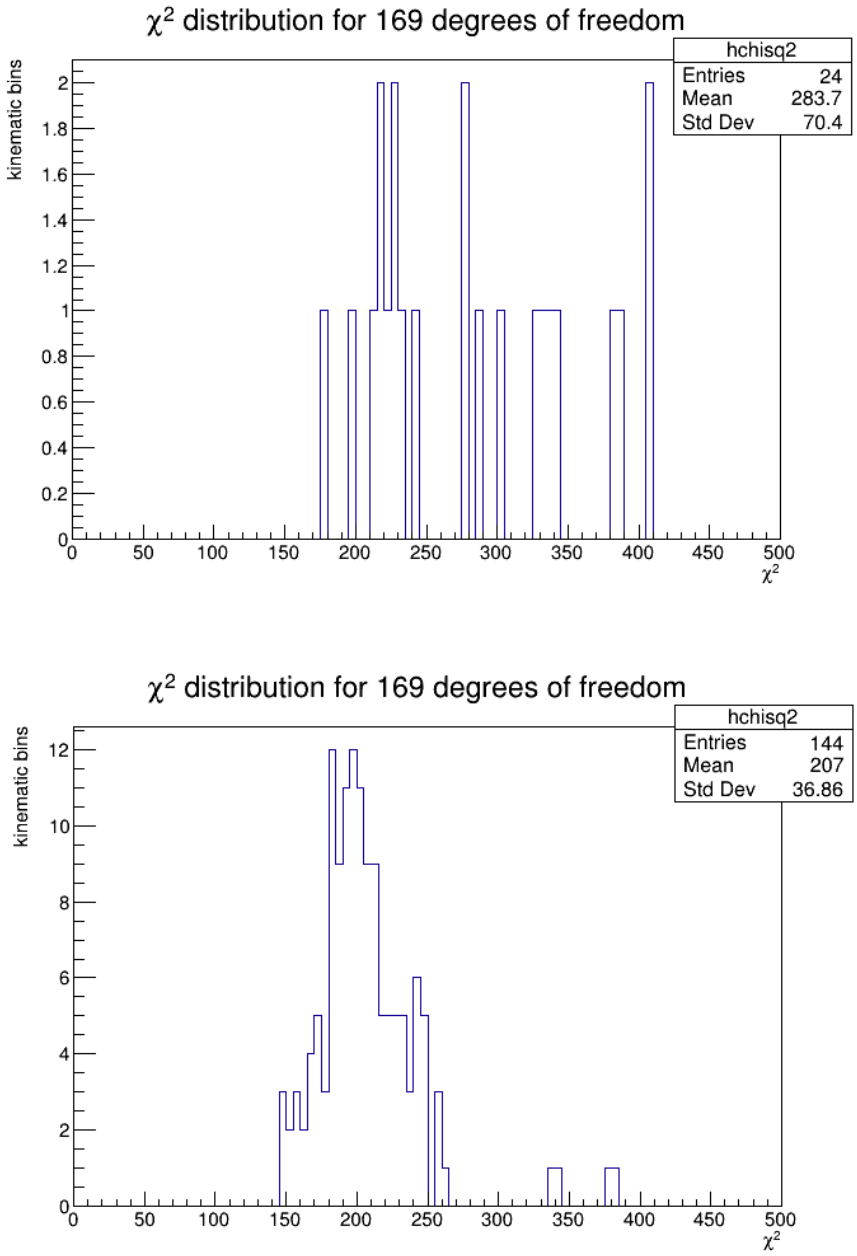

FIG. 9: χ^2 comparison between acceptance-corrected and model moments as a function of the inverse of the Monte Carlo acceptance sample size, for a coarse-binned subdivision of the (m_X, t) plane into 24 bins (upper panel), and a finer-binned scheme with 144 bins (lower panel).

On the other hand, the Monte Carlo statistics available for evaluating M in each bin argues for making them larger. To investigate the effects of Monte Carlo statistics at constant bin size, the Monte Carlo acceptance sample of 50M events was split into a subset of 10M events plus the other 40M, and the χ^2 plots shown in Fig. 9 were compared for M matrices

computed on acceptance samples with with N_{gen}=10M, 40M, and 50M. In the limit of large N_{gen}, the contribution to the χ^2 on the acceptance-corrected moments for the same mock data sample coming from the systematic error on M should scale like $1/N_{gen}$. This is shown in Fig. 10 for three different kinematic binning schemes. These data show that, in the limit of large N_{gen} there is still some excess systematic error at fixed kinematic binning, except possibly for the finest binning scheme shown in the figure.

A better way to study this systematic error might be to vary N_{gen} and the kinematic bin size at the same time, keeping the number of events per bin constant. The same data shown in Fig. 10 are re-plotted in the second panel of Fig. 10 where the two points have been chosen from among the 9 points in the upper panel to have approximately equal numbers of acceptance sample events per bin. In this second plot, the linear extrapolation to infinite Monte Carlo statistics comes to 175, which is within 5% of 169. In practice, one might be well satisfied with an acceptance correction whose systematic error contributes only 20% excess variance to the acceptance-corrected moments. Other sources of systematic error in M coming from imprecision in aspects of the simulation might very well enter at a similar level. The most important result from this study is that acceptance estimation based on a Monte Carlo sample with a factor 4 more statistics than the experimental dataset might be expected to add of order 20% to the variance on the extracted moments for a near-optimum choice of kinematic binning, and that this systematic scales with the acceptance Monte Carlo dataset size as $1/N_{gen}$.

A third potential source of systematic error is that coming from truncation of the moments expansion in Eq. 3 and following. This does not contribute to the trial results shown so far because the number of moments Q=169 used in the analysis was deliberately chosen to be much greater than 49, the number of non-zero moments assigned non-zero values in the model. The M matrix elements estimated using Eq. 10 are independent of the truncation, but its inverse M^{-1} is not. Thus it is expected that any truncation of the moments series such that moments actually present in the data are dropped from the expansion will result not merely in those excluded moments being dropped but also the lower moments that are extracted having a systematic bias associated with the truncation.

To evaluate this systematic error, four different truncations were tested on the trial sample. In the first test, the series was truncated from 169 to 49 moments, which should have no systematic error because none of the dropped moments are non-zero in the input model.

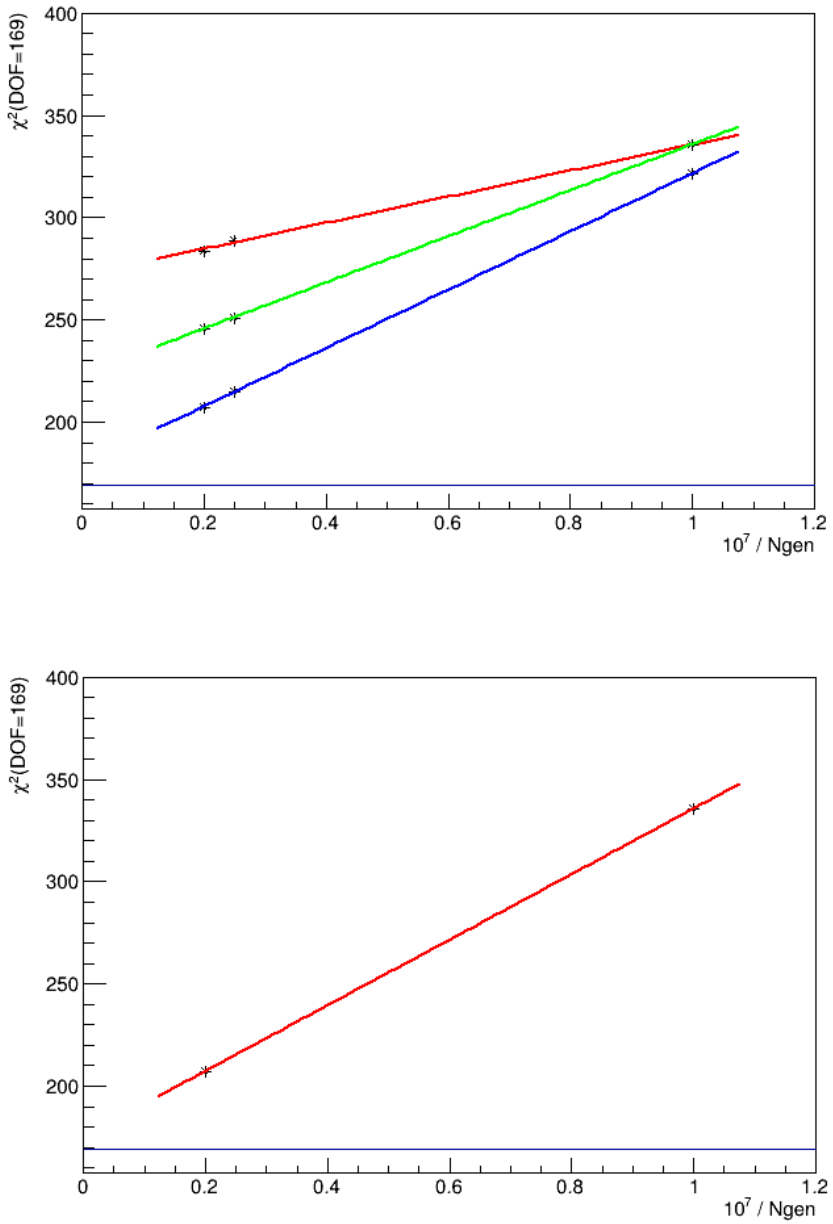

FIG. 10: χ^2 comparison between acceptance-corrected and model moments as a function of the inverse of the Monte Carlo acceptance sample size. The top panel shows the full comparison results for the coarse-bin subdivision of the (m_X, t) plane (red line) and the intermediate binning (green line) and the fine-bin subdivision (blue line). The line are fitted to the data assuming the linear behavior in $1/N_{gen}$ expected in the limit of large N. The lower figure shows the same data for a pair of points from the upper panel have approximately the same number of events per kinematic bin, showing the extrapolation to large N_{gen} if the events per bin is held fixed, while the number of bins is increased without bound.

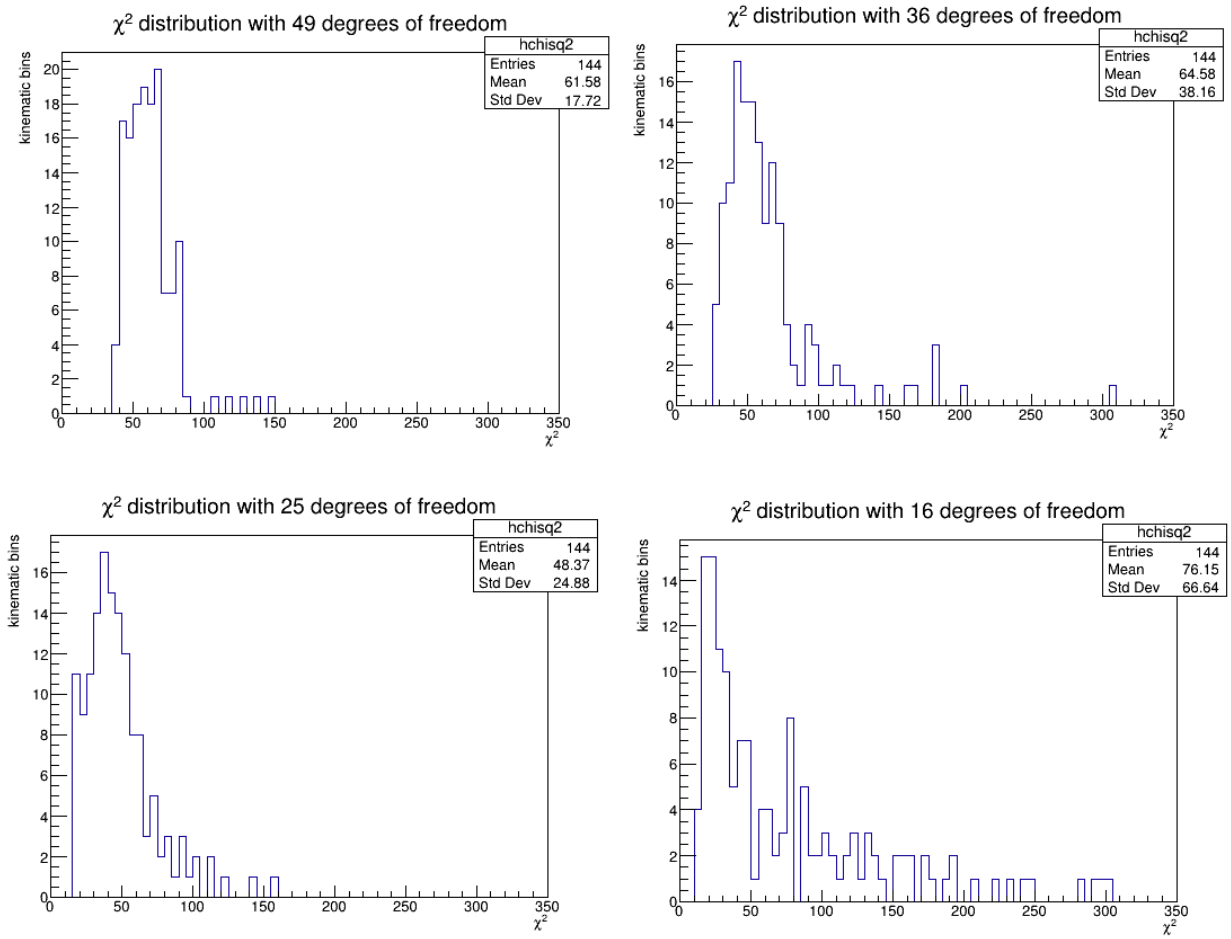

FIG. 11: χ^2 comparison between acceptance-corrected and model moments in an analysis where the acceptance correction was applied only to the first 49 moments (panel 1), the first 36 (panel 2), the first 25 (panel 3), and the first 16 (panel 4).

In the following tests, the series was truncated by dropping all moments with $L \geq 6$, $L \geq 5$ and $L \geq 4$. The results show significant and increasing systematic error in the results for the lower moments as soon as higher moments that are present in the data are dropped from the analysis.

As a final check of systematics in the trial sample, the acceptance M matrix was computed assuming that the experimental resolution on the reconstructed angles is perfect, while the mock experimental data still reflected the simulated experimental resolution on those angles. The results of this systematic error are shown in Fig. 12. The impact of neglecting the experimental resolution in reconstructed angles is clearly visible in the trial sample, although the size of the effect is relatively modest for this reaction.

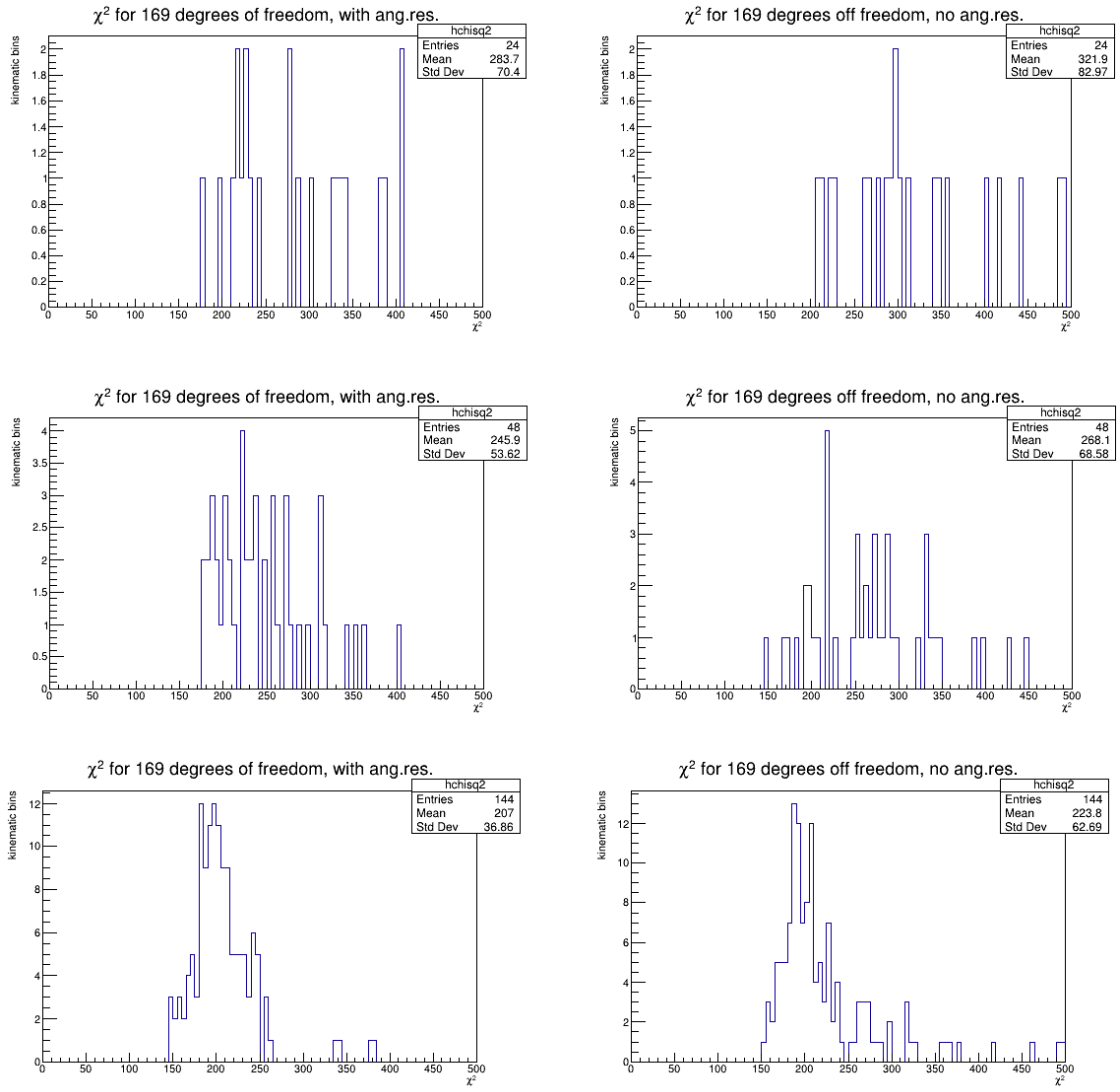

FIG. 12: χ^2 comparison between acceptance-corrected and model moments in an analysis where the effects of detector resolution in the reconstructed angles was taken into account in the acceptance correction (left plots) and neglected (right plots). The first row shows the results for the coarse-grained kinematic binning scheme, while the second and third rows show the results for the intermediate and fine binning schemes.

adding constraints

Truncation of the moments basis introduces a model-dependence to the analysis of experimental data. In the previous section it was shown that significant systematic error was introduced in the acceptance-corrected low-order moments obtained from the analysis of

the trial experimental sample once the series was truncated to exclude a set of higher-order moments that were present in the model. A model-independent approach to this issue is exemplified in the trial sample analysis presented in this study, where the analysis is performed using a moments basis that is much larger than what is seen in the uncorrected experimental moments. The full covariance matrix of the corrected moments in this large basis can then be used to control the truncation to the minimal basis needed to describe the data. This is readily accomplished through the application of linear constraints, expressed through the constraint matrix S defined in Eq. 20,

$$S\,h = 0 \tag{20}$$

where S is a $Z \times Q$ matrix representing Z linear constraints between the acceptance-corrected moments. For example, truncating the moments basis in the trial sample analysis from $L \leq 12$ to $L \leq 6$ would be represented by a 120×169 matrix S with zeros everywhere except for the 120 elements $S_{i,i+49} = 1$ for i=1 . . . 120.

Representation of these simple simple constraints in matrix form may seem excessively complicated, but adopting the form of Eq. 20 for the constraint condition allows a much more general class of constraints coming from theoretical models to be introduced within a unified analysis toolkit. Application of constraints follows the standard procedure of Lagrange multipliers by introducing a generalized χ^2 function X,

$$X(h) = \chi_0^2 + (h - h_0)^T\, C^{-1}\, (h - h_0) + 2\lambda^T\, S\, h \tag{21}$$

where h_0 are the acceptance-corrected moments prior to application of the constraints, and column vector λ contains the Z Lagrange multipliers associated with the constraints embodied in S. Extremizing X with respect to h and λ leads to updated moments h_0' and their updated covariance matrix C'.

$$\begin{aligned} h_0' &= (1 - D)\, h_0 \\ D &= C\, S^T (S\, C\, S^T)^{-1}\, S \end{aligned} \tag{22}$$

$$C' = C - DC - CD^T + DCD^T \tag{23}$$

The updated experimental moments h_0' automatically respect the constraints in Eq. 20 exactly, while their statistical errors in the updated covariance matrix C' are reduced by

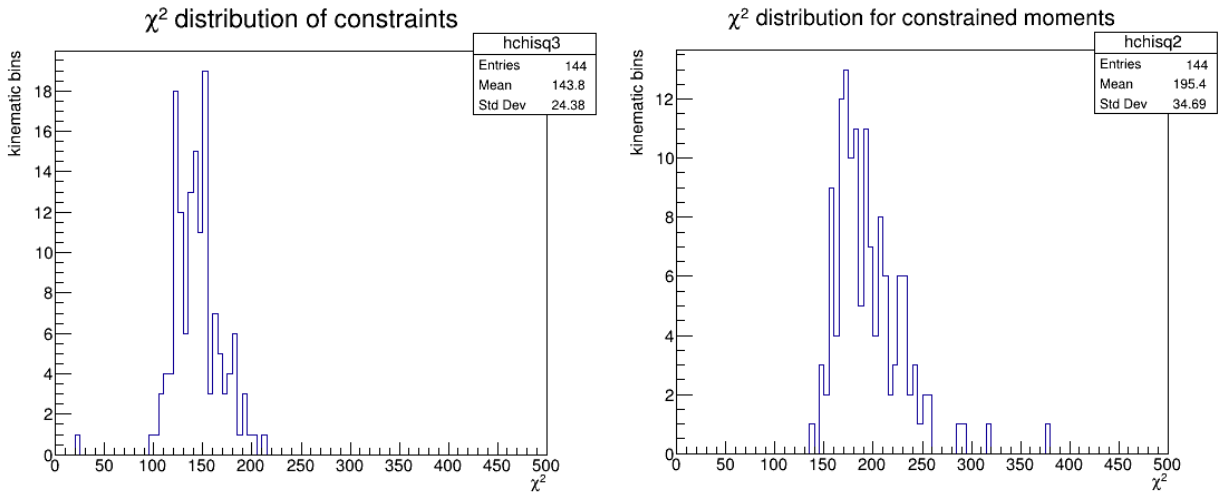

FIG. 13: Results of a test of 120 constraints applied to the set of 169 acceptance-corrected moments derived from the trial sample, segmented according to the fine binning scheme. The left plot shows the χ^2 statistic with 120 degrees of freedom from each kinematic bin, while the right plot shows the χ^2 comparison between the 169 constrained moments and the trial model inputs for the same kinematic bins.

the introduction of additional information through the constraints. Furthermore the value of X at $h = h'_0$ is distributed as χ^2 with Z degrees of freedom. A standard χ^2 test can be applied to this sample statistic to verify the consistency of the experimental data with the proposed constraints. This allows constraints between the angular moments to be applied in a systematic manner, and avoid uncontrolled systematic errors from model-dependent assumptions slipping into the analysis of experimental data.

As an example, 120 constraints setting the moments with $L > 6$ to zero were applied to the trial sample results extracted in the 169 moments basis using the approach described above. The results are shown in Fig. 13, where both the χ^2 test statistic of the 120 constraints is shown in the left plot, and the χ^2 comparison of the constrained acceptance-corrected sample moments with the models inputs is shown in the right plot. Acceptable agreement with expectations are seen in both cases.

the density matrix

Reduction of the moments from the original basis of 169 to 49 does not exhaust the full set of constraints that are present in the input model for this mock experiment. Eq. 18 shows that there is a total of 16 independent complex amplitudes, or 32 real functions that define the angular distributions. Of those, one is an invisible global phase, leading to a maximum of 31 constants that together define the 49 non-zero moments measured by the experiment. This means that there must be 18 additional constraints between the measured moments that have not been uncovered in the discussion so far.

A convenient way to expose these remaining constraints is to introduce the quantum mechanical density matrix ρ. Together with the cross section $d\sigma/dt\, dm_X$, the density matrix contains everything that can be known about the $\eta\pi^0$ system in this reaction. Eq. 18 describes a rank one density matrix. It corresponds to artificial beam conditions where the incident photon is in a pure helicity state, eg. pure left-circular polarized, the initial proton is polarized, and the final state proton detector selects a pure helicity state. Data taken under more realistic beam and target conditions would have multiple density matrices, one for each of the initial beam polarization states, each of which would be rank 2 coming from the spins of the proton that are unmeasured (4 reduced to 2 by parity). The simplest possibility, rank one, was chosen for this trial example because it exemplifies all of the features of the method without the unessential complications of working with multiple density matrices with higher-rank. Everything presented here can be readily generalized to density matrices of higher rank.

The density matrix, which contains a complete description of a quantum state, can be expressed in a basis-independent form as the density operator. The density operator underlying the model PDF in Eq. 18 is

$$\hat{\rho} = \frac{1}{|R|} \sum_{L,L'=1}^{3} \sum_{M=-L}^{L} \sum_{M'=-L'}^{L'} c_{LM}(t, m_X)\, c^*_{L'M'}(t, m_X)\ |L', M'\rangle\, \langle L, M| \tag{24}$$

where $|R|$ represents the norm of R defined such that the density operator obeys the standard normalization condition $Tr\,\{\hat{\rho}\} = 1$. The quantum expectation value of the spherical moments operators are given in terms of the density operator as

$$\left\langle \hat{h}_\alpha \right\rangle = |R|\, Tr\left\{ \hat{\rho}\, \hat{h}_\alpha \right\} \tag{25}$$

where $\hat{h}_\alpha$ are the operators whose eigenvalues are the real spherical harmonic functions when acting on states of good Ω.

$$\hat{h}_\alpha = |\Omega\rangle\, \tilde{Y}_\alpha(\Omega)\, \langle\Omega| \tag{26}$$

Setting $\left\langle \hat{h}_\alpha \right\rangle$ equal to the moments h_α measured in an experiment leads to a set of *linear* relations between them and the matrix elements of $\hat{\rho}$ in the $|L, M\rangle$ basis,

$$\begin{aligned} h_\alpha \;=\; |R| \sum_{L_1,L_2=0}^{3} \sum_{M_1=-L_1}^{L_1} \sum_{M_2=-L_2}^{L_2} \sqrt{\frac{(2L_1+1)\,(2L_2+1)\,(2L_\alpha+1)}{4\pi}} \times \\ \times \begin{pmatrix} L_1 & L_2 & L_\alpha \\ M_1 & -M_2 & -M_\alpha \end{pmatrix} \begin{pmatrix} L_1 & L_2 & L_\alpha \\ 0 & 0 & 0 \end{pmatrix} g(\rho_{L_1 M_1, L_2 M_2}, M_\alpha) \end{aligned} \tag{27}$$

in terms of Wigner's 3j symbols, where

$$g(z, M) = \begin{cases} z & :\; M = 0 \\ \sqrt{2}\,\Re\{z\} & :\; M > 0 \\ -(-1)^{L_1+L_2+L_\alpha}\,\sqrt{2}\,\Im\{z\} & :\; M < 0 \end{cases} \tag{28}$$

There are 49 non-zero moments on the left side of Eq. 27 but 256 independent real functions of (t, m_X) on the right hand side, including the norm of R plus 255 real values that determine the Hermetian ρ matrix subject to the condition $Tr\{\rho\} = 1$. Further information, for example knowledge of the rank of the density matrix, is needed in order to invert Eq. 27 and infer the density matrix elements from the measured angular moments. If the 16×16 density matrix in the L, M basis up to $L = 3$ is rank 1 then there must be 225 constraints between the 255 degrees of freedom in the density matrix to reduce 255 down to the 30 needed to completely define a 16×16 Hermetian unit-trace matrix of rank 1. Those 225 constraints can be expressed as 15 real relations of the form of Eq. 29, plus 105 complex relations of the form of Eq. 30.

$$\left(\rho_{re,LM}\right)^2 - \rho_{re,re}\,\rho_{LM,LM} = 0 \tag{29}$$

$$\rho_{re,LM}\;\rho_{LM,L'M'} - \rho_{re,L'M'}\;\rho_{LM,LM} = 0 \tag{30}$$

where re labels whichever partial wave LM is chosen to have a pure real amplitude in order to fix the global phase. Once the rank-1 density matrix elements are known, the amplitudes $c_{LM}(t, m_X)$ can be simply read out using Eq. 24 as

$$c_{LM}(t, m_X) = |R| \frac{\rho_{LM,re}}{\sqrt{\rho_{re,re}}} \tag{31}$$

The task here is to use Eq. 27 to translate those constraints on the density matrix into constraints on the moments h_α, and verify that no additional assumptions are needed in order to successfully invert Eq. 27 and infer complete information about the quantum final state from the measured moments.

The constraints in Eqs. 29-30 are quadratic in the density matrix elements and hence in the moments, in contrast to the form of Eq. 20 whose linearity in the moments enabled the direct solution to the constrained χ^2 minimization problem through the use of Lagrange multipliers. Using Lagrange multipliers to apply constraints that are quadratic in the moments leads to nonlinear equations for the constrained solution. The conventional approach is to adopt the complex c_{LM} coefficients themselves as variational parameters, and conduct a nonlinear least-squares fit to the measured moments, letting the rank-1 constraints be imposed implicitly through Eqs. 24,25. A different approach based on group theory is followed here, which should be equivalent to the conventional approach in its outcome, but might be interesting to the reader.

The density operator in the LM basis is a 16×16 positive-definite Hermetian matrix, which can without restriction be expanded as a sum over the generators of SU(N), $N = 16, \sigma_i$, $i = 1 \cdots 255$.

$$\rho = \frac{1}{N} + \alpha_j \sigma_j \tag{32}$$

in terms of 255 real parameters α_j. The orthogonality relation of the SU(16) generators,

$$Tr\left\{\sigma_j \sigma_k\right\} = \frac{1}{2}\delta_{jk} \tag{33}$$

allows Eq. 34 to be inverted to obtain Eq. 34.

$$\alpha_j = 2\, Tr\left\{\rho \sigma_j\right\} \tag{34}$$

Eq. 33 leaves a great deal of freedom in how the set of traceless matrices are chosen and organized into the generator sequence $\{\sigma_j\}$, which can be used to advantage by identifying the first $j = 1 \cdots 48$ with the measured moments h_α. Thus without any effort, the values of $\alpha_j = h_j$ needed to determine the density matrix in Eq. 32 are known. This is simply a restatement of Eq. 27, or alternatively, Eq. 27 gives an explicit representation for the first 48 generators σ_j by identifying α_j with h_α for $j = \alpha = 1 \cdots 48$. The remaining 207 generators $\sigma_j, j = 49 \cdots 255$ are assigned in an arbitrarily order from the remaining 207 traceless, Hermetian 16×16 matrices that obey Eq. 33.

The rank-1 constraint is most compactly expressed through the idempotent condition, $\rho = \rho^2$, written in terms of the α as follows.

$$\frac{1}{2}\,\alpha_i\,d_{ijk}\,\alpha_j = \left(1 - \frac{2}{N}\right)\alpha_k \tag{35}$$

where the $dijk$ are the symmetric (anticommutation) structure constants of SU(N), $N = 16$. Directly solving for the roots of these 255 simultaneous quadratic equations subject to 48 of the α being known would be a daunting task. Instead, it is convenient to start out with a random α_j that satisfies Eq. 35 and apply a general set of rotations in the SO(255) space of the $\vec{\alpha}$ that preserves the rank of ρ. To find such a general set of rotations, consider the following.

Any two rank-1 density matrices ρ and ρ' are related to one another through a change of basis represented by the 16×16 unitary operator U.

$$\rho' = U\,\rho\,U^{-1}\ ,\quad U = e^{i\omega_j\sigma_j} \tag{36}$$

for some real rotation vector ω_j in 255 dimensions. Translating U acting on ρ in SU(16) into the equivalent expression in terms of vector $\vec{\alpha}$ in SO(255) leads to

$$\alpha' = D\,\alpha\ ,\quad D = e^{-\omega_k F_k}\ ,\quad [F_k]_{ij} = f_{ijk} \tag{37}$$

where f_{ijk} are the antisymmetric (commutation) structure constants of SU(N), $N = 16$. Viewed another way, the F_i are the real generators of SU(N) in the adjoint representation. Following this approach, the solution for ρ is obtained by the following algorithm.

1. Generate a random starting rank-1 ρ matrix using Eqs. 29-30.

2. Convert ρ to α using Eq. 34.

3. Compute the distance $|\vec{h} - \vec{\alpha_j}|$ between the current $\vec{\alpha}$ and the measured moments, using the moments inverse covariance matrix as a metric.

4. Find the rotation vector $\vec{\omega}$ that minimizes the distance, and apply that rotation matrix to solve for α at the minimum.

5. Repeat until all minima are found.

The distance referenced in step 3 is defined as follows.

$$|\vec{h} - \vec{\alpha}| = (\alpha_j - h_j)\, C^{-1}_{jk}\, (\alpha_k - h_k) \tag{38}$$

where C is the statistical covariance matrix among the measured moments $j = 1 \cdots 48$. Thus defined, the distance at each minimum found by the above algorithm is a χ^2 test statistic with 18 degrees of freedom that can be used to discard any solutions that are unable to match the measured moments with an acceptable confidence level. Given more than one solution within an acceptable distance margin, the data are not able to distinguish between them and all such solutions must be reported.

Appendix: Programmer's Guide

This section serves as a guide to the software that was developed to carry out the trial of the support moments method, as described in Sec. Trial Exercise.

---

* richard.t.jones@uconn.edu; https://zeus.phys.uconn.edu; GlueX Collaboration

[1] R.T.Jones, *Jupyter notebook: Support moments trial 1* (2023), URL https://github.com/rjones30/gluex_moments_analysis/support_moments_trial_1.ipynb.