A novel method of analysis of Experimental Data

CASE STUDY:
Multipole Extraction from Nucleon Resonances

C.N. Papanicolas (with E. Styliaris)

Jefferson Lab, May 4, 2007
I. Introduction
Typical problem: To determine experimentally certain parameters of a theory or model

• The experimental results which are characterized by a statistical and systematic uncertainty are employed and the parameters are determined by minimization of the $\chi^2$.

• For complex physical problems the $\chi^2$ is a hypersurface with complex topology. The experimentally surveyed surface is known with limited precision. Identifying the minima is a difficult task. Associating uncertainties to the extracted parameters is even more difficult.
Case Study: Extraction of multipole amplitudes from the electroexcitation of the $\Delta^+(1232)$

- For the discussion of the method the underlying physics is not really important
- It is a typical difficult problem:
  - We have a good understanding of the problem and a reliable parameterization of the process
  - The connection of the parameters of interest to the observables is indirect and convoluted.
  - Constraints are imposed (unitarity in our case) which interconnect the degrees of freedom
- Issue extensively studied and extensively published.
The signal for deformation in the $N \Rightarrow \Delta$ transition

$p(qqq)$

$I = \frac{1}{2}, J = \frac{1}{2}$

$\Delta(qqq)$

$I = \frac{3}{2}, J = \frac{3}{2}$

938 MeV

1232 MeV

Spherical $\Rightarrow$ M1

Deformed $\Rightarrow$ M1, E2, C2

Deformation signal

$$CMR = \text{Re} \left( \frac{S_{1+}^{3/2}}{M_{1+}^{3/2}} \right)$$

$$EMR = \text{Re} \left( \frac{E_{1+}^{3/2}}{M_{1+}^{3/2}} \right)$$
Using the precision of the electromagnetic probe

\[ \sigma = J_\Omega \Gamma_v \frac{p_{cm}}{k_{cm}} \left( R_T + \epsilon_L R_L + \epsilon R_{LT} \cos 2\phi_{X\gamma} \right. \\
\left. - \nu_{LT} R_{LT} \cos \phi_{X\gamma} - h \nu'_{LT} R'_{LT} \sin \phi_{X\gamma} \right) \]
Extract the relevant information from the Interference Responses

\[ R_{LT} = -\sin \theta_{pq}^* \text{Re} \left \{ S_{0+}^* [M_{1-} - M_{1+} + 3E_{1+}] - \left [ 2S_{1+}^* - S_{1-}^* \right ] E_{0+} \right \} \\
-6 \cos \theta_{pq}^* \left [ S_{1+}^* (M_{1-} - M_{1+} + E_{1+}) + S_{1-}^* E_{1+} \right ] \}

\[ R'_{LT} = -\sin \theta_{pq}^* \text{Im} \left \{ S_{0+}^* [M_{1-} - M_{1+} + 3E_{1+}] - \left [ 2S_{1+}^* - S_{1-}^* \right ] E_{0+} \right \} \\
-6 \cos \theta_{pq}^* \left [ S_{1+}^* (M_{1-} - M_{1+} + E_{1+}) + S_{1-}^* E_{1+} \right ] \}

Primarily: Sensitive to C2

\[ R_{TT} = 3\sin^2 \theta_{pq}^* \left [ \frac{3}{2} |E_{1+}|^2 - \frac{1}{2} |M_{1+}|^2 \right ] \\
-\text{Re} \left \{ E_{1+}^* [M_{1+} - M_{1-}] + M_{1+}^* M_{1-} \right \} \]

Primarily: Sensitive to E2
Caveats and words of caution

- Multipoles are not observables; cross sections are.
- Need to obtain the information needed by fitting the multipoles (while comparing the cross section).
- Multipoles do not provide an orthogonal basis, they are correlated.
\( E_{L+}, E_{L-}, M_{L+}, M_{L-}, L_{L+}, L_{L-} \quad 0 \leq L \leq L_{\text{cut}} \)

\( F_1, F_2, F_3, F_4, F_5, F_6 \quad \text{(CGLN)} \)

Response Functions: \( R_T, R_L, R_{TT}, R_{LT}, \ldots \)

OBSERVABLES: Cross sections, asymmetries, etc.
Extracting Multipoles

Usual procedure followed:
1. Extract the multipoles of interest (typically only 3 to 4, “dominant”) by $\chi^2$ minimization fitting. However the data base is not rich enough to fit all multipoles.
2. Assume an $\text{L}_{\text{max}}$ and keep multipoles below that multipolarity.
3. “Not fitted” multipoles are fixed by a model.
4. One could, although not often practiced, estimate the model error of the extracted multipoles by using all available “reasonable” models.

- Extracted amplitudes and their ratios (EMR, CMR) are characterized by statistical, systematic and model error. The separation of the three is not clean.
- Model error often dominates.
- Model error is a guestimate, and therefore it does not provide a quantitative criterion for determining the acceptance or rejection of a theory/model.
II. The AMIAS Method
Randomly Vary ALL Amplitudes $A_i$ (uniformly $\pm N\sigma$) Subject to imposed constraints

AMIAS Flowchart

Make explicit the assumptions of the model and the parameters that are to be determined \{Ai\}

Experimental Data

Calculate $\chi^2$

Use Model to Calculate observables

Construct ensemble of solutions: $[\{Aj\}, \chi^2]$

Analyze the properties of the ensemble.

RESULTS $A_i$ and $\delta A_i$
Random Variation of ALL Amplitudes $A_i$ (uniformly $\pm 1\sigma$, $\pm 2\sigma$, ...)

Unitarization

$L = 0...5$

Total = (36-5) complex Multipoles

Experimental Data

Calculation of Cross Sections

Calculation of $\chi^2$

Will of course result in solutions with varying $\chi^2$

PaStyl Flowchart
\( \chi^2 \) -Distribution

Variation of ALL Amplitudes

Wider range in the variation yields different distributions

After a sufficiently wide range in the variation a CONVERGENCE in \( \chi^2 \) is reached.
Sensitivity on Multipole $A_i$

For each solution we can project the dependence of a given amplitude on $\chi^2$

$A_1, \ldots A_i \ldots A_{31}, \chi^2$

$A_i$ is uniformly distributed (varied)

$\chi^2$ versus $A_i$
Applying $\chi^2$ Cut on SENSITIVE Amplitude $A_i$

$A_i$ Distribution

Microcanonical ensemble

- $\chi^2 < 200$
- $\chi^2 < 120$
- $\chi^2 < 80$
- $\chi^2 < 40$
Uncertainty depends on the $\chi^2$ cut.

Central value remains stable.
Correlations: Standard fitting

\[ \chi^2_{\text{min}} = 1.40257, \sigma = 0.0879771 \]
Applying $\chi^2$ Cut on NON SENSITIVE Amplitude $A_i$

$A_i$ Distribution

ALL VALUES
- $\chi^2 < 200$
- $\chi^2 < 120$
- $\chi^2 < 80$
- $\chi^2 < 40$

PROJECTION
Correlations

Amplitude Correlations are automatically included through randomization in the ensemble and can be easily investigated.
Explore the allowed solutions in a given observable or derived quantity.
**Canonical Ensemble**: Use all solutions, weighted by the probability that they represent “reality”.

\[
\text{erfc} \left( \frac{\chi^2 - \chi_{\text{min}}^2}{\chi_{\text{min}}^2} \right)
\]

**Sensitive Amplitude L1+**

What is the probability that a given solution characterized by \(\chi^2\) represents “reality”??
**Canonical Ensemble:** Use all solutions, weighted by the probability that they represent “reality”

Non Sensitive Amplitude

E3+

$$\text{erfc} \left( \frac{\chi^2 - \chi_{\text{min}}^2}{\gamma^2} \right)$$
INSTEAD OF PROJECTING Out the “best solutions” (Microcanonical ensemble)

Weigh the significance of each solution by its likelihood to be correct (canonical ensemble)
Microcanonical and canonical Ensemble of Solutions

- **Microcanonical**: easier to understand, pedagogical
- **Canonical**: rigorous, far more efficient computationally
- They yield equivalent results

Use Canonical ensemble of solutions
III. Validation of AMIAS
How Reliable is the extraction of Multipoles?

Test it with pseudodata.

Algorithm:

– Step 1: Generate pseudodata
– Step 2: Analyze them
– Step 3: Demonstrate that you get input values
The extraction of Multipole values is reliable!

<table>
<thead>
<tr>
<th>Multipole</th>
<th>Generator</th>
<th>Set A</th>
<th>Set B</th>
<th>Set C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{1+}$</td>
<td>27.248</td>
<td>27.23 ± 0.13</td>
<td>27.229 ± 0.013</td>
<td>27.249 ± 0.001</td>
</tr>
<tr>
<td>$L_{0+}$</td>
<td>3.500</td>
<td>3.70 ± 0.23</td>
<td>3.515 ± 0.022</td>
<td>3.502 ± 0.002</td>
</tr>
<tr>
<td>$L_{1+}$</td>
<td>1.048</td>
<td>1.03 ± 0.08</td>
<td>1.047 ± 0.008</td>
<td>1.048 ± 0.001</td>
</tr>
<tr>
<td>$E_{1+}$</td>
<td>1.481</td>
<td>1.49 ± 0.18</td>
<td>1.489 ± 0.017</td>
<td>1.482 ± 0.002</td>
</tr>
<tr>
<td>$E_{0+}$</td>
<td>4.225</td>
<td>3.68 ± 1.02</td>
<td>4.278 ± 0.135</td>
<td>4.239 ± 0.013</td>
</tr>
<tr>
<td>$M_{1-}$</td>
<td>4.119</td>
<td>4.47 ± 1.31</td>
<td>4.161 ± 0.126</td>
<td>4.124 ± 0.013</td>
</tr>
<tr>
<td>$L_{1-}$</td>
<td>1.205</td>
<td>1.05 ± 0.43</td>
<td>1.170 ± 0.080</td>
<td>1.203 ± 0.008</td>
</tr>
<tr>
<td>$E_{2-}$</td>
<td>1.024</td>
<td>1.07 ± 0.45</td>
<td>1.053 ± 0.061</td>
<td>1.027 ± 0.006</td>
</tr>
<tr>
<td>$L_{2+}$</td>
<td>0.007</td>
<td>0.02 ± 0.01</td>
<td>0.008 ± 0.001</td>
<td>0.008 ± 0.001</td>
</tr>
<tr>
<td>$E_{2+}$</td>
<td>0.006</td>
<td>0.01 ± 0.01</td>
<td>0.009 ± 0.001</td>
<td>0.007 ± 0.001</td>
</tr>
</tbody>
</table>

$Q^2 = 0.127 \ (GeV/c)^2 \ , \ W = 1232 \ MeV)$, kinematics

Norms. The three sets are identical except for the statistical accuracy
Pseudo-Data Set

\[ Q^2 = 0.127 \text{ (GeV/c)}^2, \quad W = 1232 \text{ MeV}, \text{ kinematics} \]
**Reproduction of Values & identification of sensitivities**

Analysis of the Pseudo Data (amias_ps4_400d)
Both channels n° p (220 points) & n° n (180 points)
\[ Q^2 = 0.127 \text{ (GeV/c)}^2 \quad W = 1232 \text{ MeV} \]

<table>
<thead>
<tr>
<th>Multipole</th>
<th>MAID</th>
<th>( \sigma_0 / A_0 )</th>
<th>Ratio</th>
<th>Sens</th>
<th>MAID</th>
<th>( \sigma_0 / A_0 )</th>
<th>Ratio</th>
<th>Sens</th>
</tr>
</thead>
<tbody>
<tr>
<td>E0+</td>
<td>6.731</td>
<td>0.520 / 6.716</td>
<td>8 %</td>
<td>12.9</td>
<td>8.697</td>
<td>1.880 / 10.39</td>
<td>18 %</td>
<td>5.5</td>
</tr>
<tr>
<td>E1+</td>
<td>1.137</td>
<td>0.190 / 1.205</td>
<td>16 %</td>
<td>6.3</td>
<td>0.869</td>
<td>0.150 / 0.941</td>
<td>16 %</td>
<td>6.3</td>
</tr>
<tr>
<td>E2+</td>
<td>0.282</td>
<td>0.070 / 0.299</td>
<td>23 %</td>
<td>4.3</td>
<td>0.426</td>
<td>--</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>M1+</td>
<td>0.981</td>
<td>0.500 / 1.690</td>
<td>30 %</td>
<td>3.4</td>
<td>41.115</td>
<td>0.290 / 40.79</td>
<td>0.7 %</td>
<td>140.7</td>
</tr>
<tr>
<td>M1-</td>
<td>1.304</td>
<td>--</td>
<td>-----</td>
<td>-----</td>
<td>4.237</td>
<td>1.200 / 2.480</td>
<td>48 %</td>
<td>2.1</td>
</tr>
<tr>
<td>L0+</td>
<td>1.477</td>
<td>0.180 / 1.620</td>
<td>11 %</td>
<td>9.0</td>
<td>5.395</td>
<td>0.300 / 5.512</td>
<td>5.4 %</td>
<td>18.4</td>
</tr>
<tr>
<td>L1+</td>
<td>0.410</td>
<td>0.048 / 0.472</td>
<td>10 %</td>
<td>9.8</td>
<td>1.345</td>
<td>0.140 / 1.394</td>
<td>10 %</td>
<td>10.0</td>
</tr>
<tr>
<td>L1-</td>
<td>3.253</td>
<td>0.520 / 3.085</td>
<td>17 %</td>
<td>5.9</td>
<td>3.143</td>
<td>1.140 / 2.530</td>
<td>45 %</td>
<td>2.2</td>
</tr>
</tbody>
</table>

MAID: Input MAID-2003 value (multipole norm)
\( \sigma_0 \): Uncertainty found (gauss fit)
A\(_0\): Central value for multipole (gauss fit)
Ratio: \( \sigma_0 / A_0 \)
Sens: Multipole Sensitivity (\( = A_0 / \sigma_0 \))
Sensitivity Analysis with Pseudo-Data
Sensitivity Analysis with Pseudo-Data
How Reliable is the extraction of Uncertainties?

Testing of the reproduction of central values is straightforward.

But how do we test the validation of the extracted uncertainties?

Test it with pseudodata. Algorithm:

- Step 1: Generate ensemble of pseudodata
- Step 2: Generate “master distribution”
- Step 3: Analyze “sample data”
- Step 4: Demonstrate consistency and reliability
Bates-Mainz Data \((Q^2=0.127 \text{ (GeV/c)}^2, W=1232 \text{ MeV})\)

Take the published solution to our data to be “the truth”. Then generate pseudodata. An ensemble of pseudodata sets.
Each point in the distribution corresponds to an “experiment”.

Proceed to analyze it as described earlier and drive from it a “solution”. We can then plot the histogram of the $X^2$ of the derived solutions. This an ensemble of the solutions resulting from the sampling of the master distribution.

The distribution of the derived solutions gives an absolute definition of the uncertainty.
Central Value + Deviation
Global Set (900 runs)
Study of Pseudodata

• We reproduce both the values and the uncertainties to the appropriate statistical significance.

• The extracted uncertainties have precise statistical meaning and the resulting distributions can be interpreted precisely in terms of confidence levels.

• AMIAS allows the extraction of maximum information and identification of sensitivities
Apply the Model Independent Analysis for Multipole Extraction

$L_{\text{cut}} = 5$
## Extracted Values

<table>
<thead>
<tr>
<th>Multipole</th>
<th>Full Data Set</th>
<th>Reduced Data Set</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extracted Value</td>
<td>Extracted Value</td>
<td></td>
</tr>
<tr>
<td>$M_{1+}$</td>
<td>27.36 ± 0.15</td>
<td>27.21 ± 0.17</td>
<td>+0.5 %</td>
</tr>
<tr>
<td>$L_{1+}$</td>
<td>0.93 ± 0.08</td>
<td>0.91 ± 0.08</td>
<td>+2.2 %</td>
</tr>
<tr>
<td>$L_{0+}$</td>
<td>1.61 ± 0.22</td>
<td>2.67 ± 0.41</td>
<td>−66.0 %</td>
</tr>
<tr>
<td>$E_{1+}$</td>
<td>1.37 ± 0.20</td>
<td>1.36 ± 0.18</td>
<td>+1.0 %</td>
</tr>
<tr>
<td>$E_{0+}$</td>
<td>2.95 ± 1.13</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>$L_{1-}$</td>
<td>1.00 ± 0.52</td>
<td>0.47 ± 0.21</td>
<td>+112.0 %</td>
</tr>
<tr>
<td>$E_{2+}$</td>
<td>0.09 ± 0.11</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>$L_{2+}$</td>
<td>0.06 ± 0.09</td>
<td>0.04 ± 0.04</td>
<td>+33.3 %</td>
</tr>
</tbody>
</table>
Derived Uncertainties

Left panel: $\sigma_{LT}$
Right panel: $\sigma_{E2}$

Parameters: $\psi_{pq}$ (deg)
- Ad hoc
- Not quantitative
- Model dependent

N. Sparveris et al PRL 94, 022003 (2005)
Conclusions

• The new method is a model independent analysis for identifying sensitivities and extracting Multipole values from experimental data on Nucleon Resonances.

• The method has been examined extensively with pseudodata and with limited set of experimental data. It is stable and robust.

Remaining Issues (work in progress)

• Self adapting randomization width

• Additional variation of phases with respect to unitarization

• Extend the method to handle \( W \) dependence