Title: Modeling Error in Equilibrium-Restricted Factor Analysis

Nathanael Kazmierczak and Douglas A. Vander Griend, Ph.D., Department of Chemistry & Biochemistry, Calvin College

**Introduction**
- Factor analysis: a data mining technique that expresses a matrix as the product of two factor matrices.
- Application: UV-vis spectrophotometric titrations of equilibrium metal-ligand and host-guest complexation systems.
- Multiple error sources with non-normal distributions.
- Error propagation poorly understood.
- Technique often treated as a “black box.”
- Beer’s law: \( A = e \times \lambda \times c \)
- Factor matrices obtained through constrained nonlinear least-squares optimization:

\[
\begin{align*}
\text{Absorbance (p x m)} & = \frac{1}{(n x m)} \times \text{Equilibrium Concentrations (m x p)} + \text{Residual (n x p)} \\
\end{align*}
\]

**Titration Anatomy**
- Composition (starting chemicals)
- Equilibrates
- Spectrometer measures absorbance data
- Analyzed by

**Composition Error**
- \([H] = 0.001, \text{no transmittance error}\)
- \([H] = 0.001, \text{no compositional error}\)
- Error in calculated \( \Delta G \)

**Transmittance Error**
- \([H] = 0.001, \text{no composition error}\)

**Stoichiometry Error**
- RMS Residual for 1:1 stoichiometry complexation, error-free artificial data

**1:1 Exergonic Limits**

<table>
<thead>
<tr>
<th>([H]_0)</th>
<th>(\Delta G^o) lower bound (kJ/mol) at 298 K</th>
<th>(K)</th>
<th>(K[H]_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1 \times 10^{-6})</td>
<td>-84</td>
<td>(5 \times 10^{14})</td>
<td>(5 \times 10^{16})</td>
</tr>
<tr>
<td>(1 \times 10^{-5})</td>
<td>-79</td>
<td>(7 \times 10^{13})</td>
<td>(7 \times 10^{15})</td>
</tr>
<tr>
<td>(1 \times 10^{-4})</td>
<td>-72</td>
<td>(4 \times 10^{12})</td>
<td>(4 \times 10^{14})</td>
</tr>
<tr>
<td>(1 \times 10^{-3})</td>
<td>-68</td>
<td>(8 \times 10^{11})</td>
<td>(8 \times 10^{13})</td>
</tr>
<tr>
<td>(1 \times 10^{-2})</td>
<td>-62</td>
<td>(7 \times 10^{10})</td>
<td>(7 \times 10^{12})</td>
</tr>
<tr>
<td>(1 \times 10^{-1})</td>
<td>-56</td>
<td>(6 \times 10^{9})</td>
<td>(6 \times 10^{11})</td>
</tr>
<tr>
<td>1</td>
<td>-51</td>
<td>(8 \times 10^{8})</td>
<td>(8 \times 10^{10})</td>
</tr>
</tbody>
</table>

- 0.0003 transmittance error added (best-case spectrometer error)

**Conclusions**
- Monte Carlo adaptations of Sivvu™ software allowed high-volume error studies.
- Each type of error has a distinct signature.
- Each type of error can be diagnosed separately.
- Composition and transmittance error effects are amplified for highly exergonic reactions owing to a loss of model sensitivity to the \(\Delta G^o\) value.
- There exists a smooth RMS minimum at the correct model stoichiometry, potentially allowing optimization to find the correct model.
- For 1:1 reactions, there exists a fundamental lower bound to the \(\Delta G^o\) value that can be reliably calculated in the presence of spectrometer transmittance error.

**Acknowledgements**
- Calvin College
- Arnold and Mabel Beckman Foundation
- National Science Foundation