Finding Specificity in Multivariate Curve Resolution

Jeremy M. Shaver, Neal B. Gallagher
Eigenvector Research, Inc.
Manson, WA 98155

Shaver@Eigenvector.com

EIGENVECTOR RESEARCH INCORPORATED

Finding Specificity in Multivariate Curve Resolution

- Multivariate Curve Resolution (MCR)
- Simulated Evolution Example
  - Multiplicative Ambiguity
  - Rotational Ambiguity
  - Initial Guesses
- *Simple* Dissolution Example
- Conclusions

EIGENVECTOR RESEARCH INCORPORATED

Multivariate Curve Resolution

- MCR is most often used with spectra,
  - also known as self-modeling curve resolution, self-modeling mixture analysis, “end member extraction”
- Often used when you do NOT have calibration data
- Goal is to recover underlying "factors" which represent physically-identifiable features.

MCR

- Based on the classical least squares (CLS) model, attempt to estimate **C** and **S** given **X**:

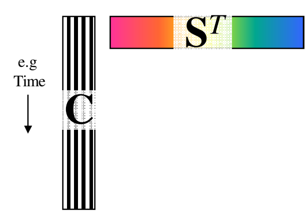
$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E}$$

X *M*x*N* measured responses,

C *M*x*K* pure analyte contributions,

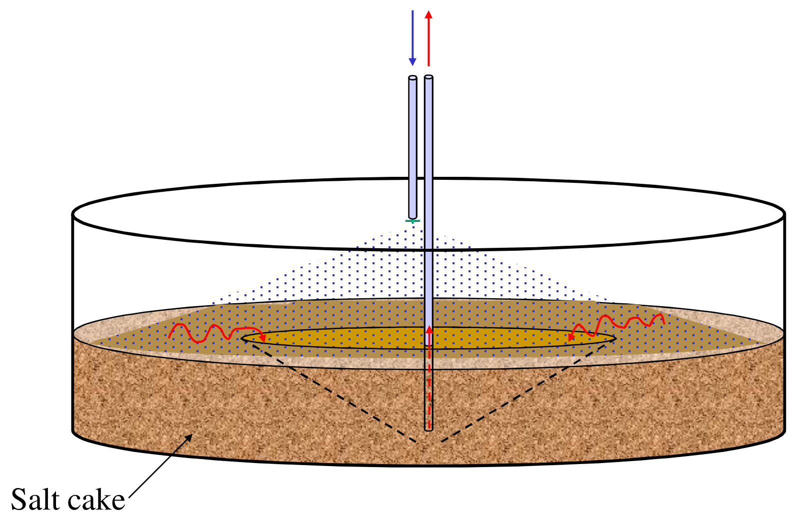
S *N*x*K* pure analyte spectra, and

E *M*x*N* residuals.

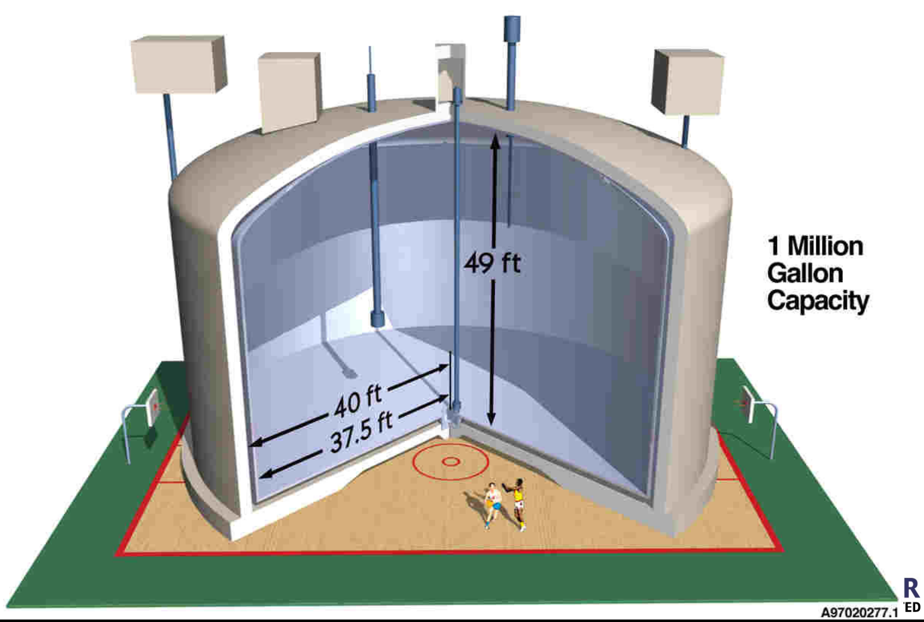


- Non-Negative Alternating Least Squares (NN-ALS)

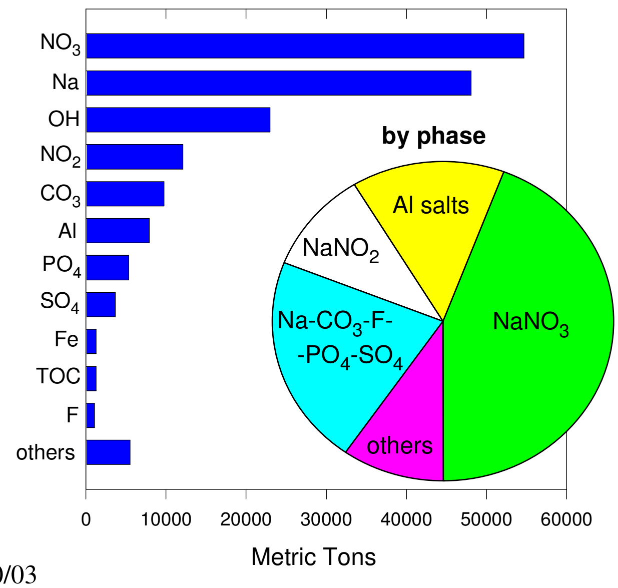
Nuclear Waste Storage Tanks



Hanford Double-Shell Tank

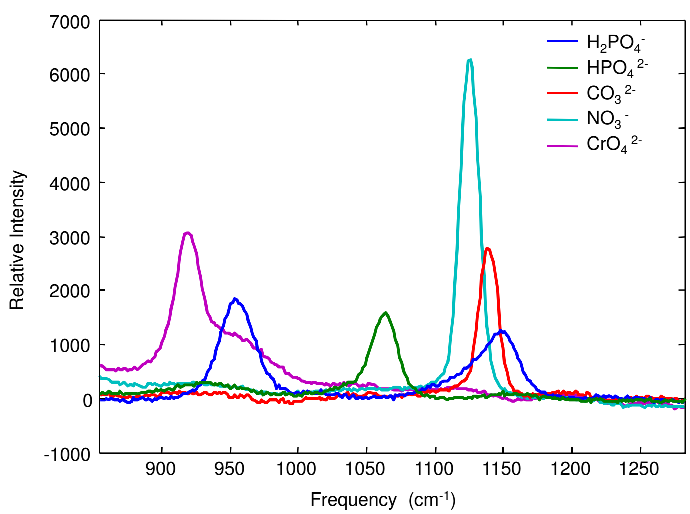


Tank Farm Chemical Inventory



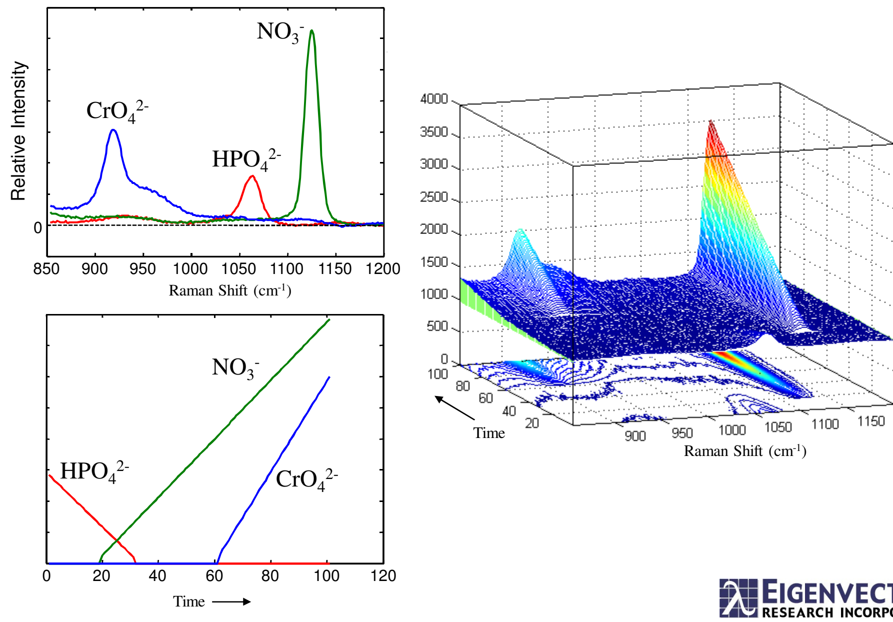
BBI 3/20/03

Raman Spectra of Ionic Species



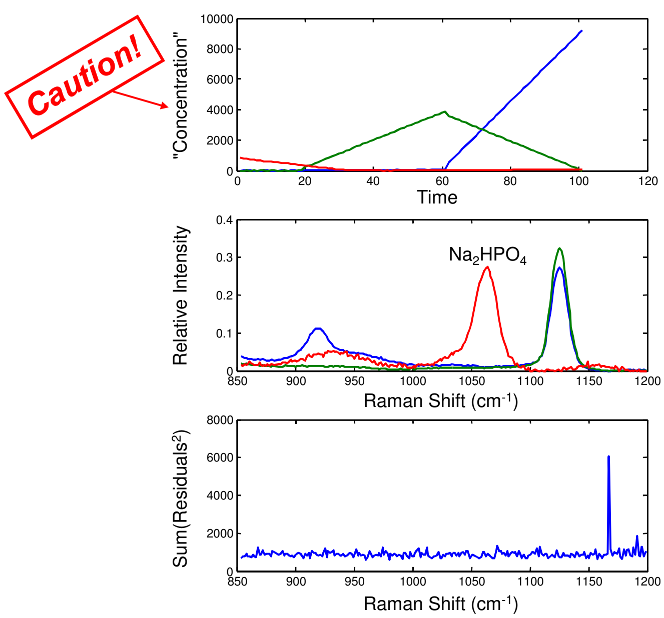
Data courtesy Sam Bryan, Battelle (PNNL)

Simulated Evolution



EIGENVECTOR RESEARCH INCORPORATED

Multiplicative Ambiguity



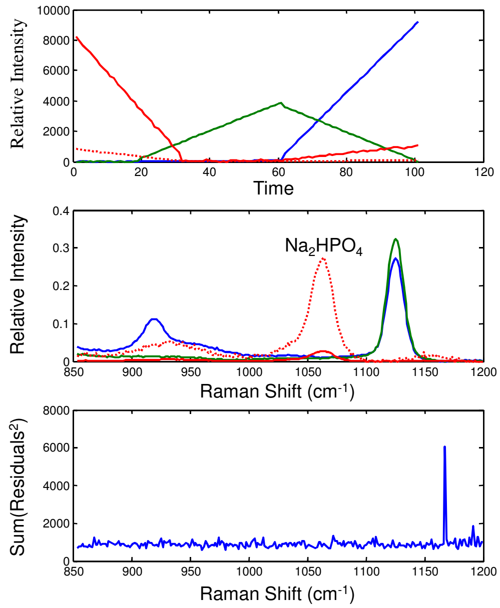
Scores (C)

Factors (S)

Residuals (E)

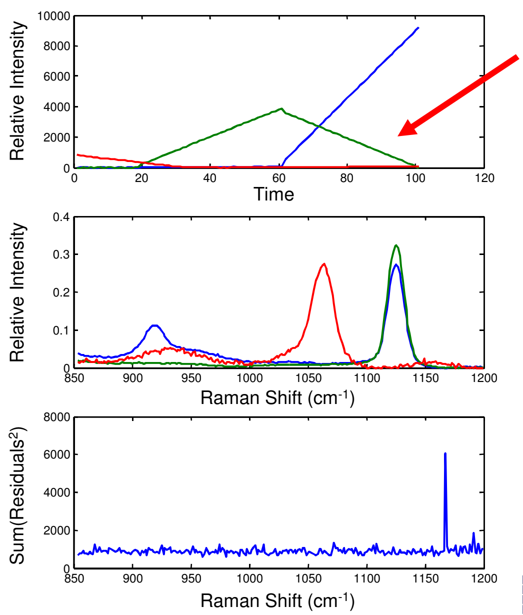
EIGENVECTOR RESEARCH INCORPORATED

Multiplicative Ambiguity



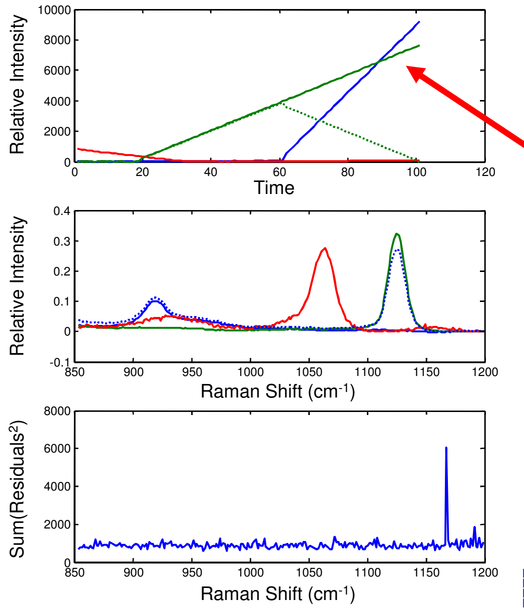
EIGENVECTOR RESEARCH INCORPORATED

Rotational Ambiguity



EIGENVECTOR RESEARCH INCORPORATED

Rotational Ambiguity



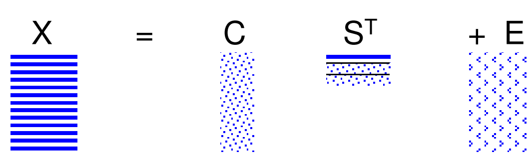
EIGENVECTOR RESEARCH INCORPORATED

Ambiguity Work-Arounds

- **Fix factors** to pure spectra measured at known concentrations.

With all components, this → *CLS!* but must know species *a priori*

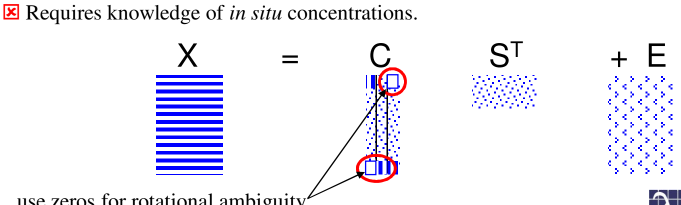
☒ Caution for matrix effects which change spectral profile (e.g. band shifts).



- **Fix contributions** at given times to known concentration values.

☑ Less sensitive to matrix effects

☒ Requires knowledge of *in situ* concentrations.

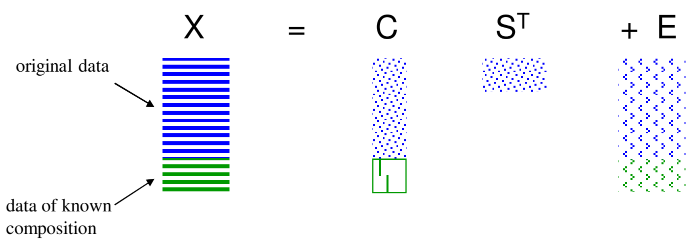


EIGENVECTOR RESEARCH INCORPORATED

Ambiguity Work-Arounds

•**Add spectra** (at known concentrations) to data along with fixed concentrations.

- "Soft" constraint - allows some error "wiggle-room".



☑ Less sensitive to matrix effects on spectral shape
☒ Still sensitive to matrix effects on Scattering efficiency / absorbtivity

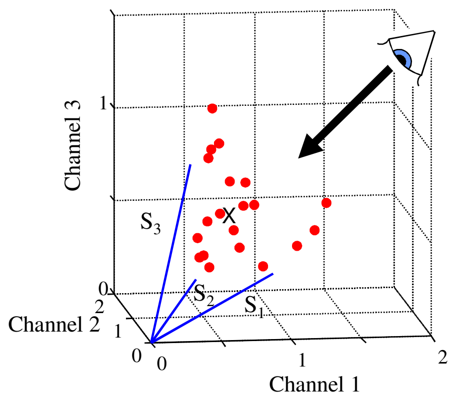
Ambiguity Work-Arounds

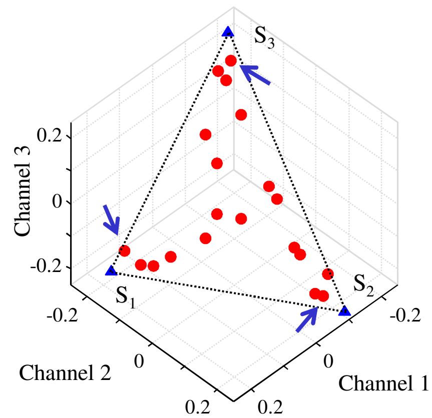
- **Multiplicative:** Use known chemical relationships to get *RELATIVE* concentrations (e.g. identification of species + order of reaction, or Closure)
- **Rotational:** Use expected simplicity of contribution profiles to direct solution towards simple explanation (e.g. unimodality).

☒ **All these methods:**
Unexpected components? *Interpret Carefully!*

Sum Scaled Intensity
Sum Scaled Current
Sum Scaled …

Initial Guess
"Pure[est] Samples"

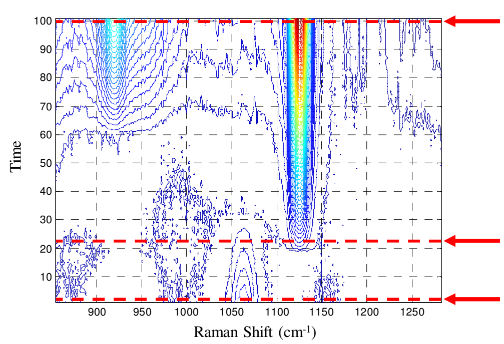


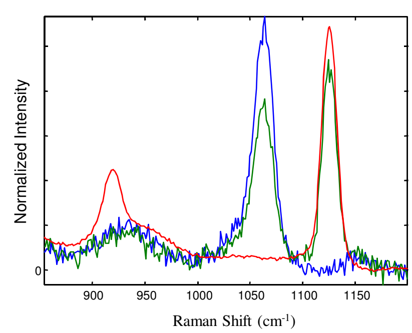


— normalize + mean center →

EIGENVECTOR RESEARCH INCORPORATED

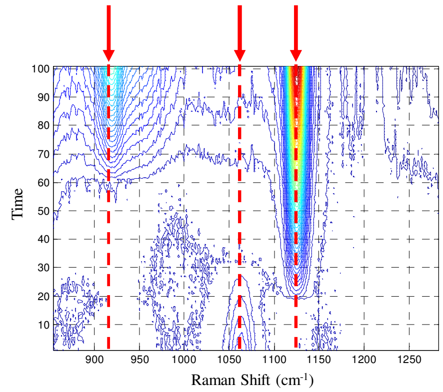
Initial Guess
"Pure[est] Samples"

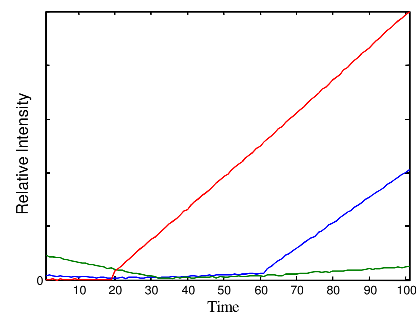




EIGENVECTOR RESEARCH INCORPORATED

Initial Guess
"Pure[est] Variables"

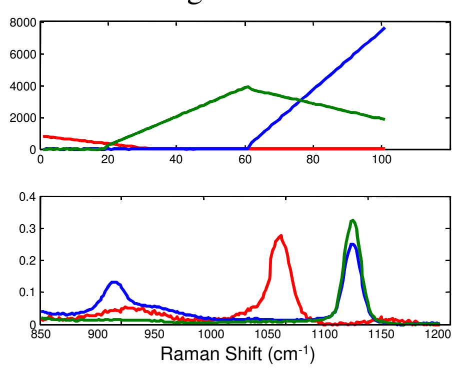




EIGENVECTOR RESEARCH INCORPORATED

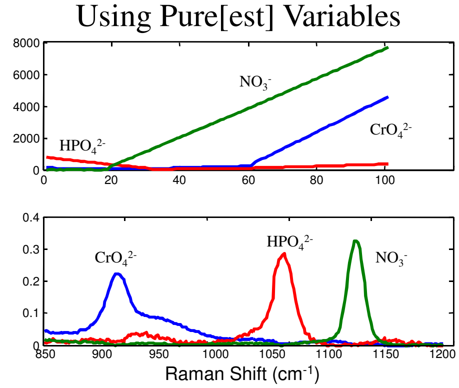
Initial Guess
"Pure[est] Variables"

Using PCA Scores



(similar to results with random starting vectors)

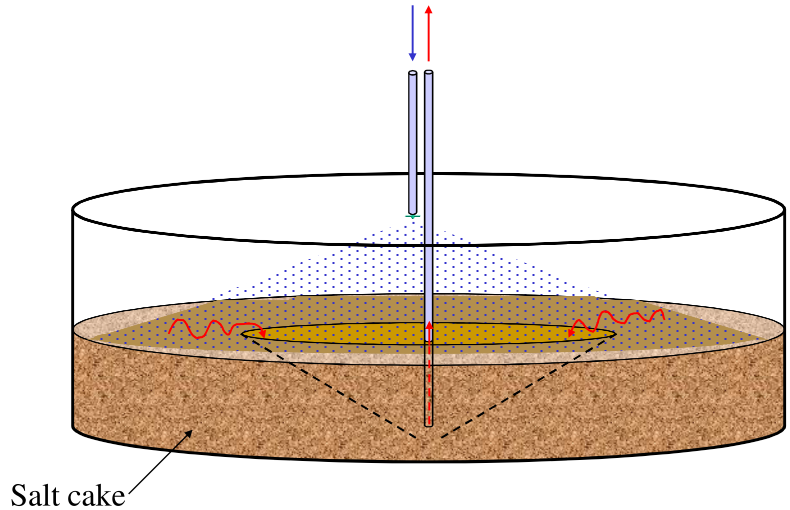
Using Pure[est] Variables



CrO_4^{2-} and HPO_4^{2-} were the spectrally difficult ones!

EIGENVECTOR RESEARCH INCORPORATED

Nuclear Waste Storage Tanks



Simple Dissolution Simulation

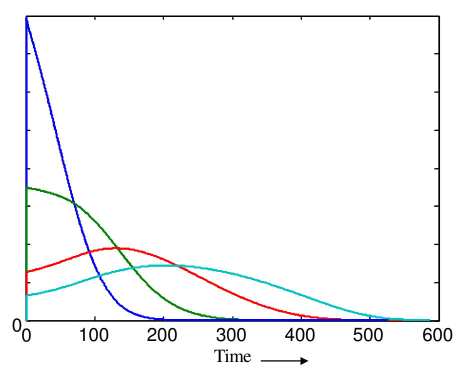
common ion — observed species

$$XA_{(s)} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X^+ + A^- \xrightarrow{k_T} elution$$

$$XB_{(s)} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} X^+ + B^- \xrightarrow{k_T} elution$$

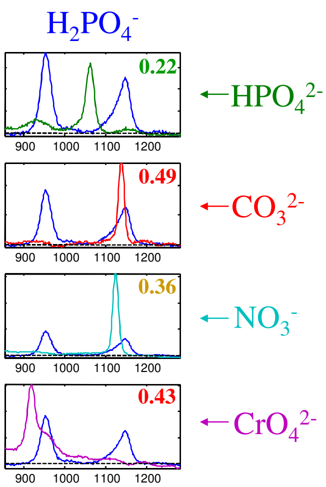
$$XC_{(s)} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} X^+ + C^- \xrightarrow{k_T} elution$$

$$XD_{(s)} \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} X^+ + D^- \xrightarrow{k_T} elution$$



- Chemistry of species assigned to a given evolution profile does NOT necessarily match the equations above (e.g. HPO_4^{2-})
- Super-saturation effects (H-bonding) NOT included in model
- Ionic strength effects NOT included in model

Quantifying Spectral Overlap



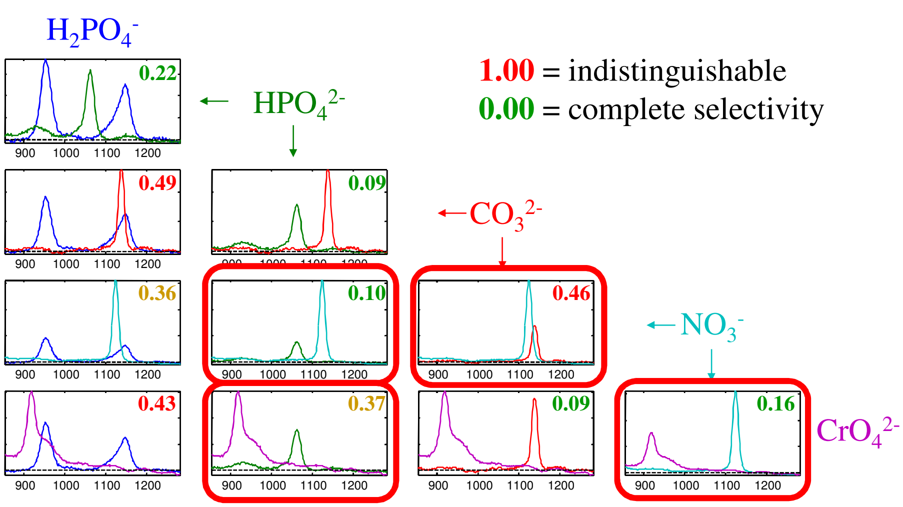
1.00 = indistinguishable
0.00 = complete selectivity

$s_{1,norm} = ||s_1||$
$s_{2,norm} = ||s_2||$
$c = s_{1,norm} \bullet s_{2,norm}^T$

EIGENVECTOR RESEARCH INCORPORATED

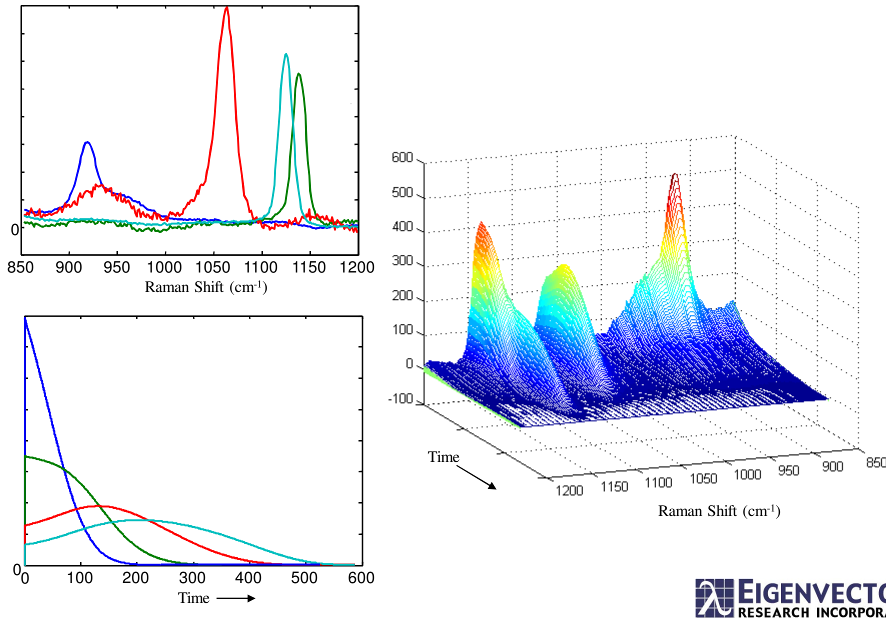
Quantifying Spectral Overlap

1.00 = indistinguishable
0.00 = complete selectivity

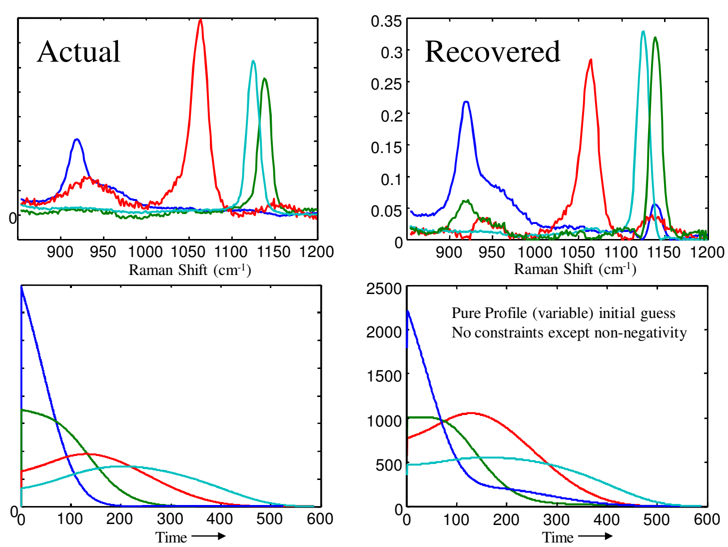


EIGENVECTOR RESEARCH INCORPORATED

Simple Dissolution Simulation



Simple Dissolution Simulation



Conclusions

- A variety of hard and soft constraints exist to reduce effect of multiplicative and rotational ambiguities BUT
- Sometimes, just starting in the right neighborhood goes a long way!

EIGENVECTOR RESEARCH INCORPORATED