Speciation of Mixtures

Overview

- Least Squares Analysis
- Example: Least-Squares Fitting
- Analysis of Sulfur K-edge Spectra of Living Mammalian Cells
- Principal component analysis
- Example: A photo-chemical reaction
- Choice of Model Spectra
- Summary

Near-Edge Spectra as Tools

- Deduce local information:
  - Electronic, geometric, etc
  - Empirically or using calculations
- Identify chemical forms present in "unknown" sample
  - "Speciation"
  - Spectra used as fingerprint
  - Compare with library of standards
X-ray Absorption Near Edge spectra

Near-edge spectra provide a “fingerprint” of chemical type.

Need a comprehensive library of “model compound” spectra.

More sensitive than EXAFS (low concentrations are OK)

RSO3

RSO2

Fe4S4

SO4

2-

RS=O

R3S+

RS-Me

RS-H

RS-SR

S8

SO3

2-
Near-Edge Spectra in Speciation

- XAS can analyze all of the element in the sample, no matter how it is present:
  - Aqueous solution
  - Crystalline or amorphous solid
  - Species adsorbed on a surface
  - Trapped gas
  - …

- X-ray absorption spectroscopy (XAS) can analyze a material essentially without pretreatment
  - Although:
    - Usually freeze the sample
    - May grind it to give a more representative sample

Analysis of mixtures

$$A + B = \text{mixture}$$

Near-edge Spectra of Mixtures

- The spectrum of a mixture of more than one chemical forms of an element appears as the sum of the spectra of the individual components

- The height of the edge jump of each component is proportional to the quantity of the element in that form
Speciation of Mixtures

Near-edge Spectra of Mixtures

- Generally spectra are normalized so that their edge jump is unity

- Model spectra are normalized to the same amount of the element
- The height of the edge jump of each component is proportional to the fraction of the element present in that form

Fitting can yield Quantitative Analysis

Model spectrum of mixture with a sum of model compounds, and use least-squares fitting to quantitatively analyze how much of each is present.

Fitting can yield Quantitative Analysis

Composition = X% blue + Y% red
Least Squares Fitting

- In least-squares fitting we minimize a function \( F \):
  \[
  F = \frac{1}{N} \sum_{j=1}^{N} (y_{j,\text{obs}} - y_{j,\text{calc}})^2
  \]
- Where the calculated intensity is given by:
  \[
  y_{i,j,\text{calc}} = \sum_{i=1}^{m} x_i I_i(x_j)
  \]

Methods of Deconvolution – Least-Squares Fitting.

Can be applied to limited data sets - a single spectrum can be analyzed (that's often all we have).

- Pro - can yield quantitative estimate of chemical composition.
- Con - difficult to judge whether small contributions are significant.
- Con - cannot detect small fractions if something is there in excess.
- Con - absolutely require model spectra.

Fitting of Derivatives

- Derivatives of the spectra can also be fit
- Near-edge and derivative fits should agree

Example: Least-Squares Fitting Analysis of Sulfur K-edge Spectra of Living Mammalian Cells

Graham N. George
Canada Research Chair in X-ray Absorption Spectroscopy

Sulfur K-edge X-ray Absorption Spectra

The spectra are very sensitive to the chemical form of sulfur.

Curve-fitting example - Sulfur forms in whole blood

Whole blood can be easily separated into plasma and erythrocytes by centrifugation.

Use sulfur K-edge XAS to investigate the chemical forms of sulfur in horse erythrocytes and plasma.

Erythrocytes show a predominance of reduced sulfur forms (i.e. thiols) while plasma shows oxidized sulfur forms (i.e. disulfides).

This confirms the well known idea that the inside of cells is much more reducing than the outside.
Grow cells on Transwell™ plates - on a 0.2μm thick polycarbonate film.

MDCK cells are considered to be prototypical polarized epithelial cells.

Cells are mounted in the X-ray beam (with minimal possible disturbance, but rotated through 90°) still growing on their polycarbonate substrate.

Cells cultures are kept in an incubator at beamline until loading.
Methods of Deconvolution - Principal Component Analysis

This is a technique that can be applied to extended data sets - m spectra of a mixture of < m components.

1. Principal Component Analysis - can (in principal) count the number of components.
2. Target Transformation - can identify components of mixtures.

Principal Component Analysis

Based on a mathematical technique called Single Value Decomposition

\[ A = U \cdot V \cdot W^T \]

\( m \) Spectra, \( n \) data points each

\[ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nm} \end{bmatrix} \cdot \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{m1} \\ v_{12} & v_{22} & \cdots & v_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{mn} \end{bmatrix} \cdot \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix} \]

Eigenvectors (principal components)
Eigenvalues
Weights

Principal Component Analysis

\[ A = U \cdot V \cdot W^T \]

Spectra A
\[ A \]
PCA
\[ U \cdot V \]

Principal components are NOT the component spectra
In this example 3 components could be used to reconstitute the data
**Target Transformation**

Essentially asks the question "is this model spectrum a member of our set of spectra?". Target transformation can thus be used to test whether candidate component spectra are present.

\[ T^* = U \cdot T \]

- \( U \) can be restricted to include only the significant principal components.
- If \( T \) and \( T^* \) are (essentially) identical then candidate spectrum is part of the set and the candidate compound is a part of the mixture.

**Target Transformation**

Transformed spectrum \( T^* \) and Candidate spectrum for aqueous selenite – \( T \) and \( T^* \) are essentially identical so selenite is present in the mixture.

**Target Transformation**

Transformed spectrum \( T^* \) and Candidate spectrum for \( \alpha \)-selenium – \( T \) and \( T^* \) are different so \( \alpha \)-Se is not present in the mixture.
Principal Component Analysis

Principal Component Analysis will not work when you have fewer spectra than components in the mixture.

All spectra must have common abscissae.

Target transformation absolutely requires accurate model spectra.

Note that a "component" could itself be a complex mixture of species that has invariant composition amongst the spectra in the set.

Quantitative analysis

The conclusions of principal component analysis and target transformation can be used with least-squares fitting to obtain a quantitative analysis.

Results of a three-component least-squares fit
Quantitative Analysis

Set of related spectra → number and identities of components from Principal Component Analysis and Target Transform → Least Squares fitting of each spectrum in set → Quantitative analysis of whole set

Quantitative Analysis

Least Squares fitting of each spectrum in set

Two alternative approaches are available in using the least-squares analysis of the data to obtain a quantitative analysis of the set:

1. Use model compound spectra to do least squares fitting
2. Use target transforms of model compound spectra to do least squares fitting.

The second method is actually rather like fitting the data with itself - we consider fitting with target transforms to be undesirable, although numerically correct.

Example: A photo-chemical reaction
Example - A photo-chemical reaction

R-S'● is an elusive free-radical chemical intermediate of biochemical importance.

From considerations of the expected electronic structure we expect R-S'● to have a very low-lying transition in the sulfur K near-edge spectrum.

Observation - prolonged exposure to beam of a stable sulfenic acid (RSOH) produces a transition in exactly the region expected for R-S'●

RSOH → R-S'● + OH●

Could this be R-S'●?

How many components does the reaction produce?
Use Principal Component Analysis to find out.

Spectral Time Course

Spectra taken with increasing time show increasing levels of low-energy peaks.

Principal Component Analysis

Twenty spectra at different times; only two components are indicated.
**Example – A Chemical Reaction**

\[ \text{RSOH} \xrightarrow{h\nu} \text{RS}^+ \cdot \text{OH}^- \]

Principal component analysis indicates only two components in all spectra.

Use kinetic analysis with the spectrum of the starting material, and difference spectra to compute spectra of product.

---

**Spectrum of RS·**

\[ \text{RSOH} \xrightarrow{h\nu} \text{RS}^+ \cdot \text{OH}^- \]

- **LUMO**
- **LUMO+1**

Computed energy difference ~4eV

The 2465 eV peak can be used as a probe of RS· in biological systems.

---

**Choice of Model Spectra**

---
The Choice of models is critical to a good analysis

Spectra depend on the chemical and physical environment.

It is critical that model spectra are collected under appropriate conditions.

Sulfur K-edge XAS spectra of substituted thiophenes.

Remote structure can have significant affects...

Physical form can be important
Solids vs. Solutions...

In general, peaks of solution spectra are more pronounced than with solids. The reason for this is that crystal packing forces in the solid distort the molecule lifting orbital degeneracy and cause bound state transitions to be more spread out in energy.

pH can be important

Spectra of an aqueous solution of cysteine taken at different pH values.

Sulfite has a pKₐ of 6.2
Summary

One spectrum – use Least-squares fitting.

Set of spectra – use PCA, Target & Least-squares.

PCA ··················· Gives number of components.
Target ··················· Gives identities of components.
Least-Squares Fitting ···· Gives quantitative estimation.

Acknowledgements

Stanford Synchrotron Radiation Laboratory, U.S. DOE and NIH
Canadian Light Source and its many sponsors www.lightsource.ca
Canada Research Chairs Program
University of Saskatchewan
Province of Saskatchewan
CFI, NSERC, NIH, CIHR

Canadian Light Source and its many sponsors www.lightsource.ca

Stanford Synchrotron Radiation Laboratory, U.S. DOE and NIH
ACKNOWLEDGEMENTS