Spectral Methods to Extract Useful Information at Very Low Signal-to-Noise

Ralph Levy
CTO
Quant Engineering
Ralph.Levy@QuantEngineering.com

Presented at: **Sensors Expo & Conference 2012, Rosemont, IL**

Work funded by: **Quant Engineering Internal Research and Development (IR&D) project funding**

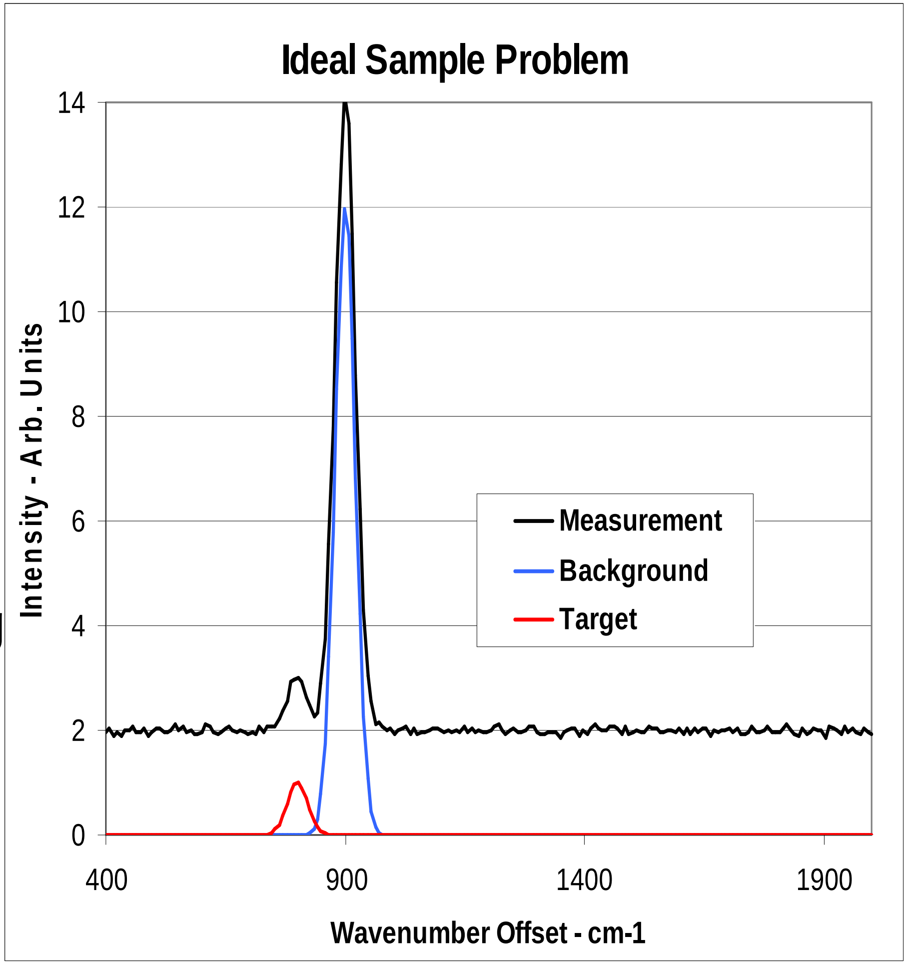
Spectroscopy

- Data
  - One of the good things about spectroscopy is that there is a lot of data
  - One of the problems with spectroscopy is that there is a lot of data

- Data-> Information-> Meaning-> Wisdom

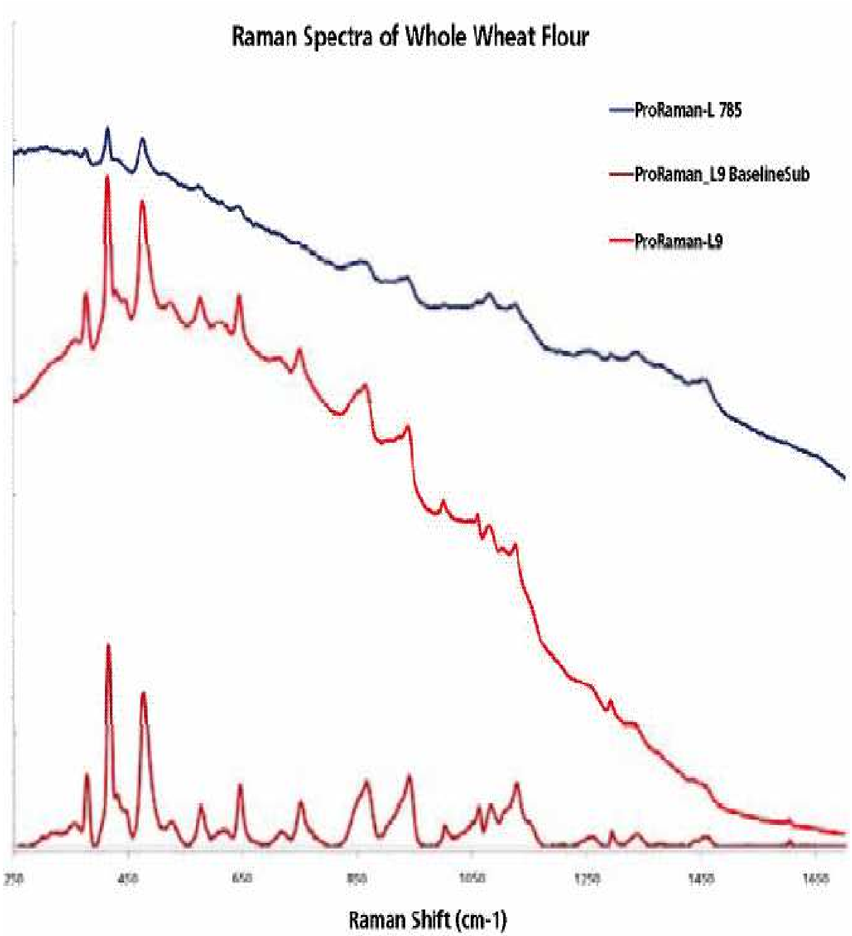
Statement of the Problem – In Principle

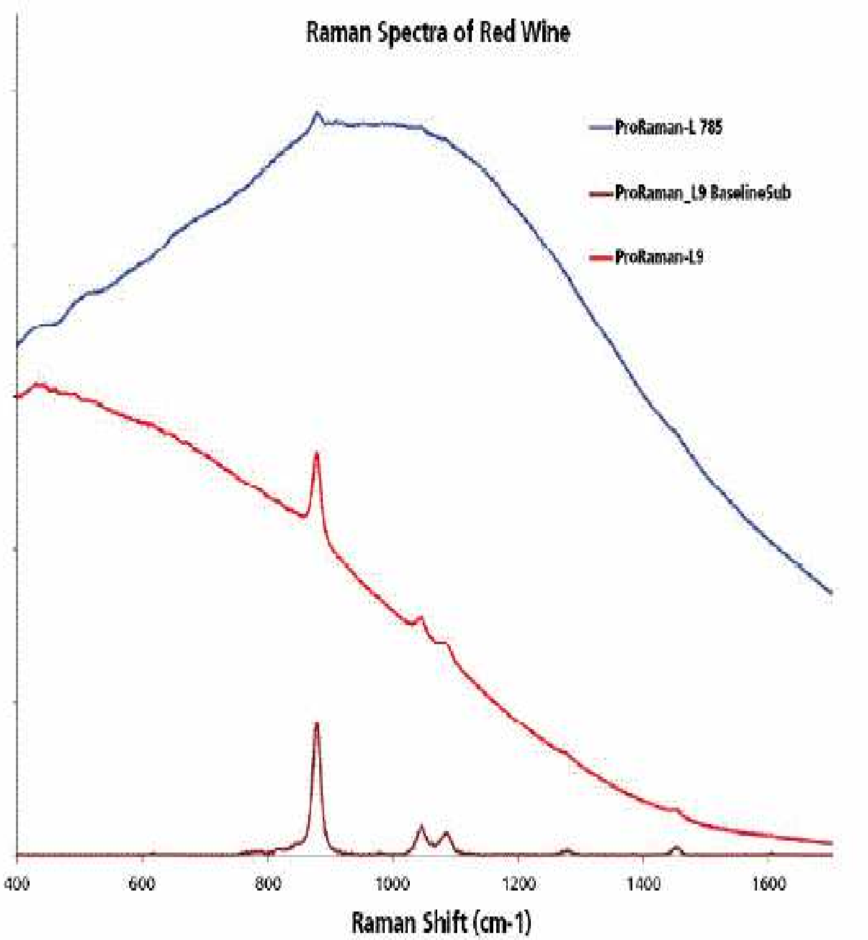
- Raman, FTIR and Fluorescence Spectroscopy (other spectral measurements)
- Known system with Background and Target signals
- **Find the Spectral Abundance of the Target Species – Spectral Unmixing**
- Note:
  - Spectral Abundance may need to be Transformed into Species Abundance (for example Beer's Law)



Examples at Excellent SNR

Enwave Optronics, Inc. www.enwaveopt.com





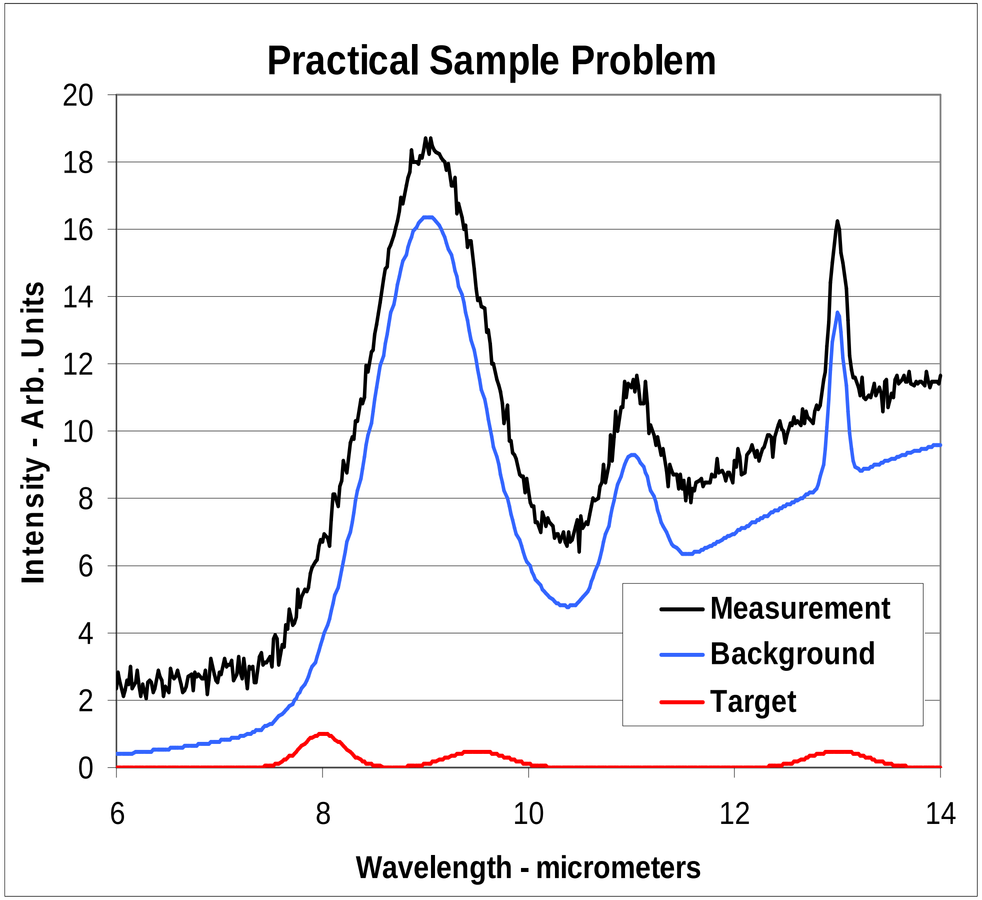
Limitations on Acquisition Signal and Noise

- Acquisition Time
    - Throughput – QA / QC
    - Remote Sensing
    - Transient Phenomena

- Applied Power
    - Hardware Limitations
    - Sample Damage
    - Dynamic Range of the Instrument

Statement of the Problem – In Practice

- Background and Target overlap
- Multiple Background species, each in unknown quantities
- Noise is as big as the Target



Signal: Information, Background and Noise

- Instrument manufacturers normally use Dynamic Range divided by Dark Noise as the listed SNR
  - Appropriate when the application is unknown

- Background signals are often much bigger than what we are trying to measure
  - Contaminants, additives, trace quantities

Signal: Information, Background and Noise

- The critical issue is “Information” not total signal
- Information, our Target Signal, is often 1 to 2 orders of magnitude smaller than background signals
  - Raman
    - Fluorescence, interfering spectra
  - FTIR Remote Sensing
    - Self emission and background species, poorly controlled environment
  - Contaminants as opposed to Principal species
    - Process Monitoring/QC/QA in Chemical, Pharma, Food

Signal: Information, Background and Noise

- In this talk I want us to think of "Target Signal"-to-Noise rather than the term Signal-to-Noise

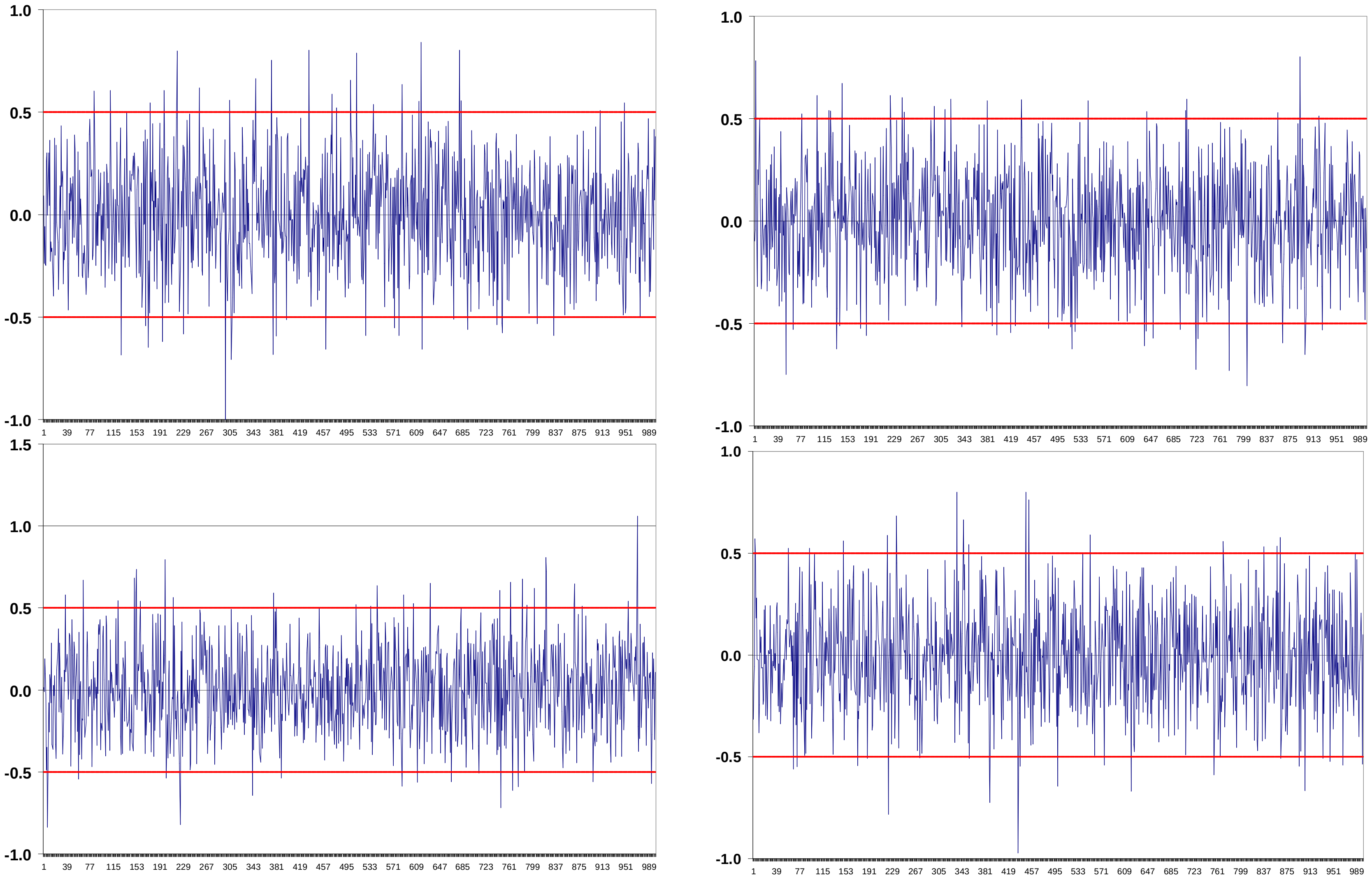
- If we have robust techniques to remove background, then total signal intensity is not the critical parameter in discussing error unless it affects the dynamic range or linear range of the instrument

Noise

- Noise has many sources including thermal noise in the detector, electronics noise and interference
- There are also drift and calibration issues that are outside the scope of this talk
- Noise is modeled here as Gaussian with a mean of Zero and a stated magnitude

Gaussian Noise

Here Magnitude is “defined” as 4 times Standard Deviation



Traditional Approaches

- Single wavelength analysis
  - Sit on the Absorbance/Emission peak and subtract the background
  - Often what is taught at school
  - Deservedly bad reputation

- Integrated peaks
  - A modest improvement in some cases
  - Still need accurate background intensity

Vector Sub-Space Story

- If measuring in 3-space and
  - Information is in X direction and
  - Interfering signal is in Y-Z plane
  - We would project our measurements normal to the Y-Z plane

- If we knew the direction of the Interfering signal we would do all our analysis in a direction orthogonal to the Interfering signal

If Each Spectral Point is a Direction

- If each spectral point (wavelength, frequency, wavenumber) is a direction
- We can project our measurements in a direction orthogonal to all the background spectra

Wait !!!

- With, e.g., 200 or 2000 (or more) spectral wavelengths
- That requires 200 or 2000 (or more) -dimensional Matrix Math !

Wait !!!

- With, e.g., 200 or 2000 (or more) spectral wavelengths
- That requires 200 or 2000 (or more) -dimensional Matrix Math !

- **Oh No ! ! !**

Wait !!!

- With, e.g., 200 or 2000 (or more) spectral wavelengths
- That requires 200 or 2000 (or more) -dimensional Matrix Math !

- **Oh No ! ! !**

- Is there another choice?

What About Least Squares?

- Linear Problem: find the coefficients (α_i) to best fit the measurement:

$$\text{Signal}(\lambda) = \Sigma\ \alpha_i\ S^i(\lambda) + N(\lambda)$$

What About Least Squares?

- Linear Problem: find the coefficients (α_i) to best fit the measurement:

$$\text{Signal}(\lambda) = \Sigma\, \alpha_i\, S^i(\lambda) + N(\lambda)$$

- OK if very few background spectra and low noise

What About Least Squares?

- Linear Problem: find the coefficients (α_i) to best fit the measurement:

$$\text{Signal}(\lambda) = \Sigma\, \alpha_i\, S^i(\lambda) + N(\lambda)$$

- OK if very few background spectra and low noise
- Terrible idea in the real world with multiple background spectra and lots of noise
  - Computes all abundances including background
  - Fits the noise
  - Negative abundances possible (not uncommon) especially for the trace species of interest

OK Back to Matrix Math

- Relax
- It is actually easy
- I will give you the code to go along with *Numerical Recipes in C*.
- Even the matrix inversion is fast and robust using Singular Value Decomposition.

Orthogonal Subspace Projection

Following Chang (2006):

$$\mathbf{Measurement}(\lambda) = \gamma_i * \mathbf{S}^i(\lambda) + \mathbf{N}(\lambda) \qquad (1)$$

Where the γ_i coefficients are the abundances of the $\mathbf{S}^i$ components. Separating the spectra into the Desired (**d**) spectral component and Undesired (**U**i) components:

$$\mathbf{Measurement}(\lambda) = \alpha * \mathbf{d}(\lambda) + \beta_i * \mathbf{U}^i(\lambda) + \mathbf{N}(\lambda) \qquad (2)$$

Create a matrix that will eliminate the $\mathbf{U}^i$ components:

$$\mathbf{OSP = I - UU^{\#}} \quad \textbf{where} \quad \mathbf{U^{\#} = (U^T U)^{-1} U^T} \qquad (3)$$

Multiplying equation (2) by **OSP**:

$$\mathbf{OSP * Measurement}(\lambda) = \alpha * \mathbf{OSP * d}(\lambda) + \mathbf{OSP * N}(\lambda) \qquad (4)$$

To find the abundance α, take the ratio:

$$\alpha = [\ (\mathbf{OSP * Measurement}(\lambda)) \bullet \mathbf{d}(\lambda)\ ]\ /\ [\ (\mathbf{OSP * d}(\lambda)) \bullet \mathbf{d}(\lambda)\ ] \qquad (5)$$

Top-Level View of Code

In parent code:

```
sing = ComputeOSP(MP, NP, Ubackground, OSP, Out); // code below
if (sing) return sing;

//    Transform Sample Measurement (Umeasure) into
//    (Uprojected) to Remove Background
Mproject(OSP, MP, Umeasure, Uprojected);          // code below

//    Abundance is alpha/alphadPd
//    Compute Correlation of Uprojected with Target from Library
alpha = Dot(Uprojected, Utarget, MP);             // dot product

//    Compute Scaling for Target from Library
Mproject(OSP, MP, Utarget, dPd);                  // code below
alphadPd = Dot(dPd, Utarget, MP);                 // dot product

Abundance = alpha/alphadPd;
```

Numerical Recipes uses Arrays that run from 1 to N !

```
void Msubtract(double **a, int m, int n, double **b, double **c)
{
// Input  matrix "a" is mxn
// Input  matrix "b" is mxn
// Output matrix "c" is mxn

//             Note that the indexes go from:
//                       1 to N (as in FORTRAN)
//             rather than from :
//                       0 to N-1 (as in C) !!!

    int         k, l;

    for (k=1; k<=m; k++)
    {
          for (l=1; l<=n; l++)
          {
               c[k][l] = a[k][l]-b[k][l];
          }
    }
}
```

Declarations

```
int ComputeOSP(int MP, int NP, double *Uin, double *OSP, FILE *Out)
{
//  MP    Number of data points per Spectrum
//  NP    Number of Background Spectra
//  Uin   Input Spectral Matrix
//  OSP   Output OSP Matrix
//  Out   File for fprintf output

    double          *w;
    double          **Pu;
    double          **u;
    double          **v;
    double          **UUh;
    double          **UH;
    double          **U;
    double          **Ut;
    double          **UU;
    double          **UUi;
    int   sing = 0;
    int   NN;
    int   i;
    int   j;
```

Allocate Storage and Unload Input Data

```
NN  = max(MP, NP);
w   = dvector(1, NN);
Pu  = dmatrix(1, NN, 1, NN);
UUh = dmatrix(1, NN, 1, NN);
u   = dmatrix(1, NN, 1, NN);
v   = dmatrix(1, NN, 1, NN);
UH  = dmatrix(1, NN, 1, NN);
U   = dmatrix(1, NN, 1, NN);
Ut  = dmatrix(1, NN, 1, NN);
UU  = dmatrix(1, NN, 1, NN);
UUi = dmatrix(1, NN, 1, NN);

//              Unload Input Spectral Matrix
for(i=1; i<=MP; i++)
      for(j=1; j<=NP; j++)
           U[i][j] = Uin[NP*(i-1)+(j-1)];
```

Compute the OSP Matrix

```
//              Ut = Transpose(U)
Mtranspose(U, MP, NP, Ut);
//              UU = Ut * U
Mmult(Ut, NP, MP, NP, U, UU);
//              UUi
sing = Minverse(UUi, UU, NP, NP, w, v);
if (sing)
{
      fprintf(Out,"\n\nSingular %d ! ! ! ! ! ! * * * * * !\n\n", sing);
}
else
{
      //        UH = U# = UUi * Ut
      Mmult(UUi, NP, NP, MP, Ut, UH);
      //        UUh = U * U#
      Mmult(U, MP, NP, MP, UH, UUh);
      //        v =I
      Midentity(v, MP);
      //        Pu = (I-U*U#)
      Msubtract(v, MP, MP, UUh, Pu);

      //        Load Output OSP Matrix
      for(i=1; i<=MP; i++)
           for(j=1; j<=MP; j++)
                OSP[MP*(i-1)+j-1] = (float)Pu[i][j];
}
```

Free Allocations and Return

```
    free_dmatrix(U  , 1, NN, 1, NN);
    free_dmatrix(Ut , 1, NN, 1, NN);
    free_dmatrix(UU , 1, NN, 1, NN);
    free_dmatrix(UUi, 1, NN, 1, NN);
    free_dmatrix(v  , 1, NN, 1, NN);
    free_dmatrix(u  , 1, NN, 1, NN);
    free_dmatrix(Pu , 1, NN, 1, NN);
    free_dmatrix(UUh, 1, NN, 1, NN);
    free_dmatrix(UH , 1, NN, 1, NN);
    free_dvector(w  , 1, NN);

    return sing;
}
```

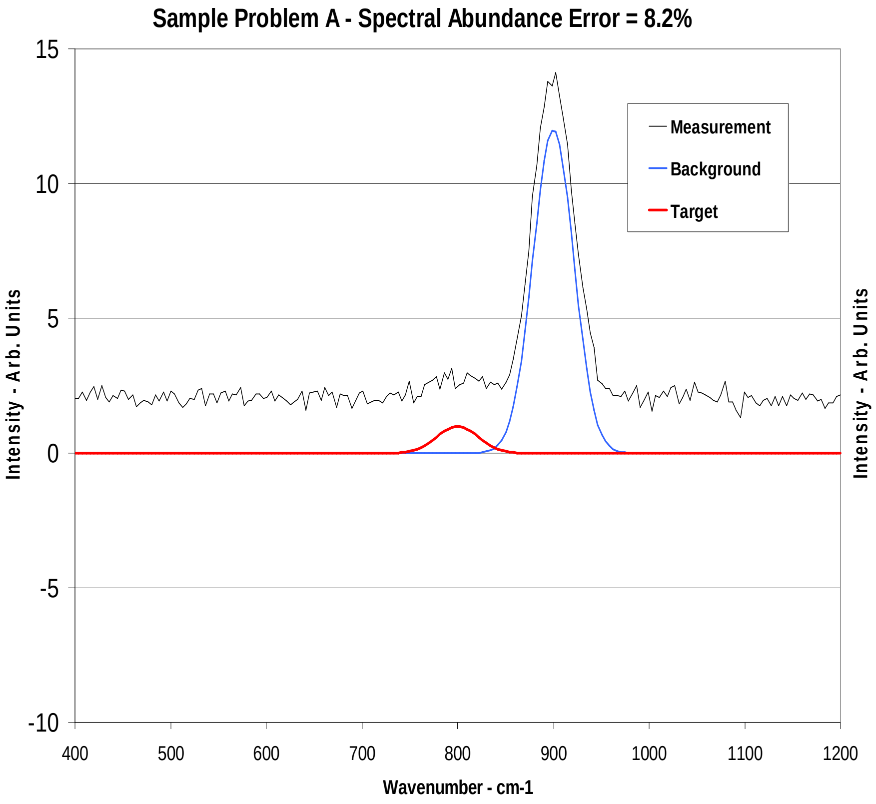
Mproject Utility

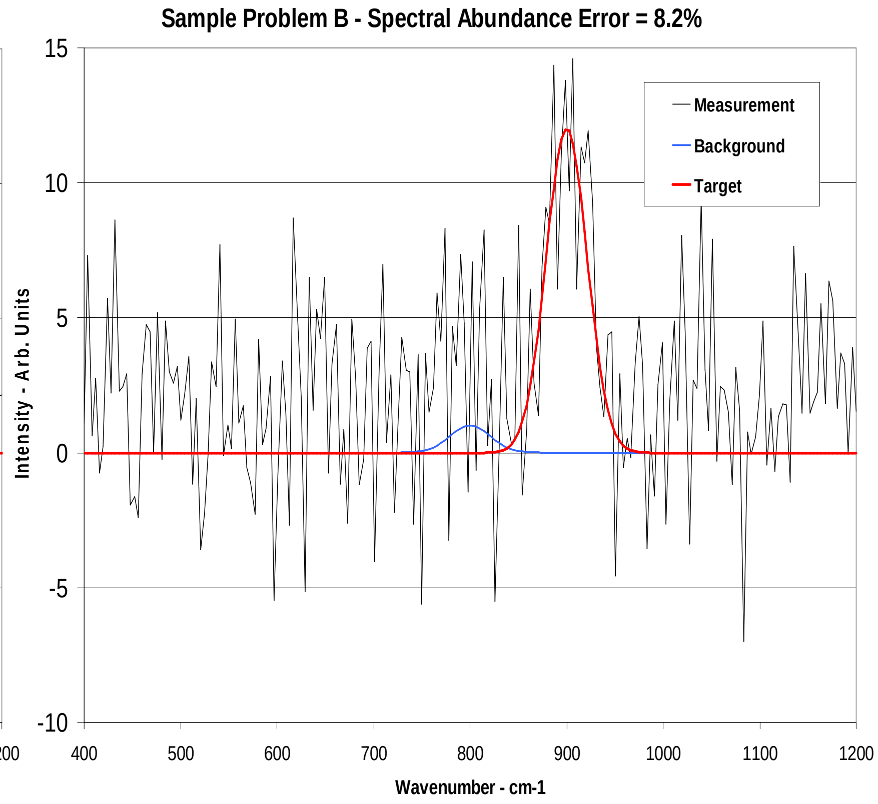
```
void Mproject(double *a, int n, double *b, double *c)
{
    //    Input  matrix "a" is nxn      a[k][j] = a[n*(k-1)+j]
    //    Input  matrix "b" is 1xn      b[j][l] = b[p*(j-1)+l]
    //    Output matrix "c" is 1xn      c[k][l] = c[p*(k-1)+l]
    int        j, k, l;

    for (k=0; k<n; k++)
    {
          l=0;
          c[k] = 0.0;
          for (j=0; j<n; j++)
          {
               c[k] += a[n*k+j]*b[j];
          }
    }
}
```

Applications and Examples

Big and Small Background Spectra
Target Signal-to-Noise = 1.0





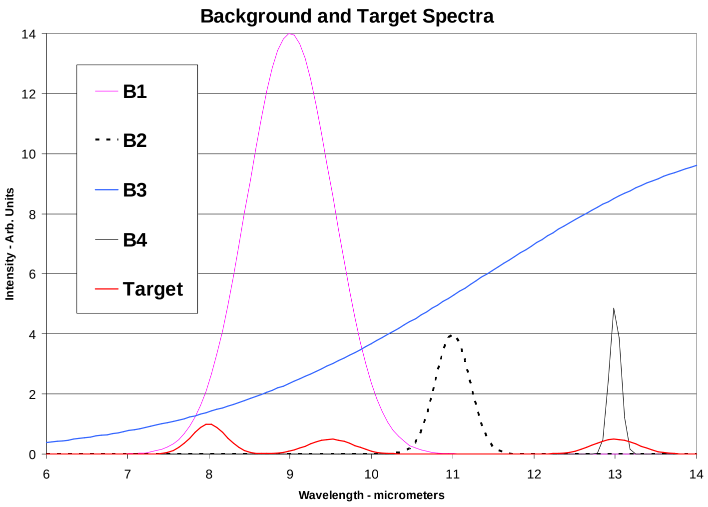
Notes

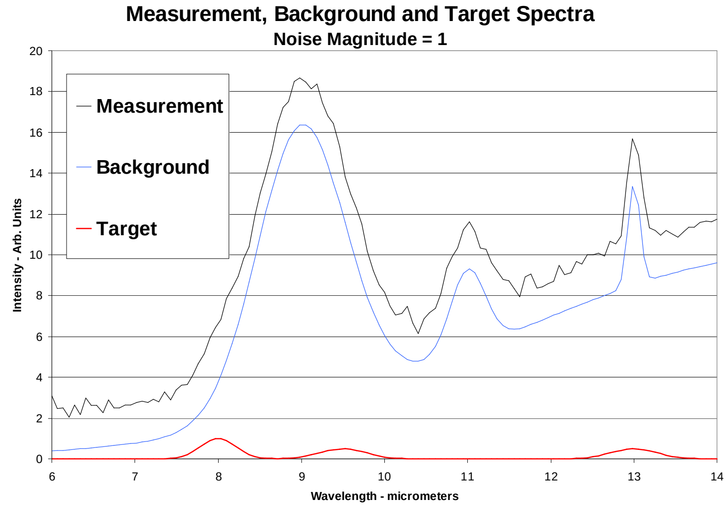
- Data are not “algebraic” but sequences of intensities, so analysis is not tied to any functional form.

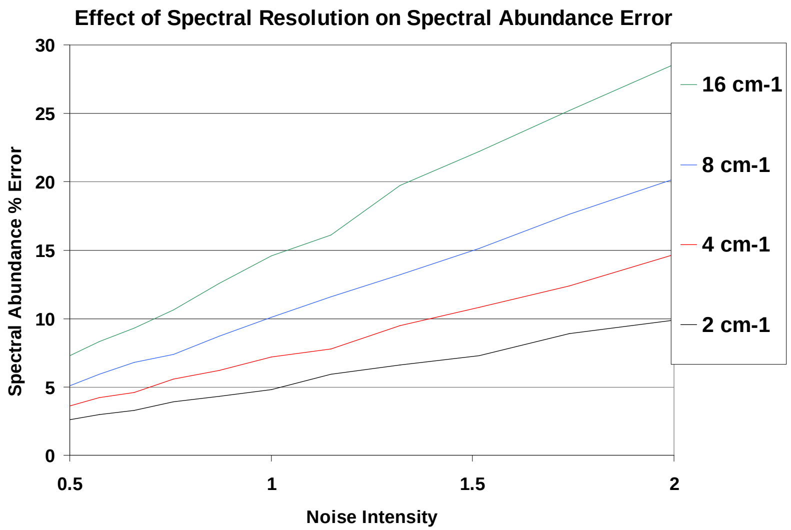
- Summary information below is presented as Standard Deviation of Abundance Error vs. Noise Magnitude where the Target Signal maximum magnitude = 1.0 !

Spectral Resolution

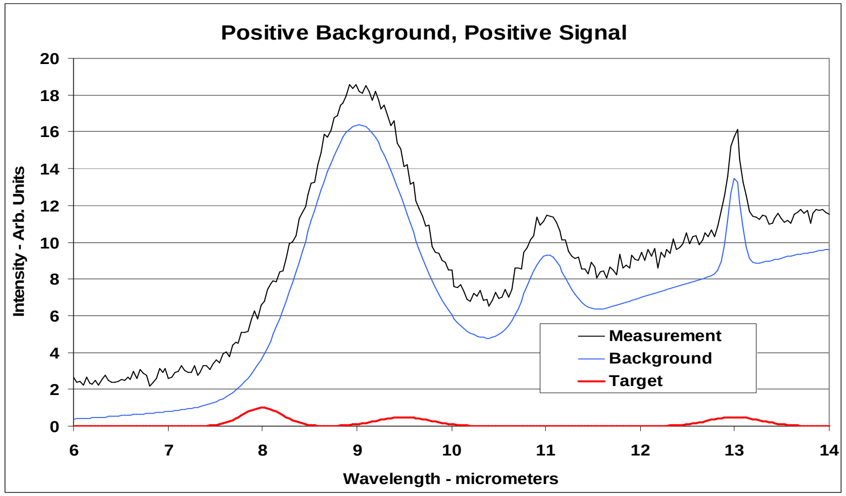
Target Signal Intensity = 1.0

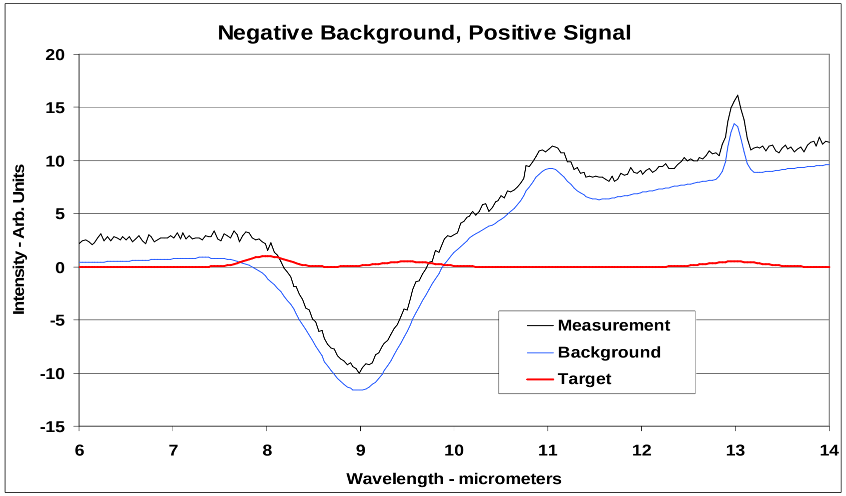


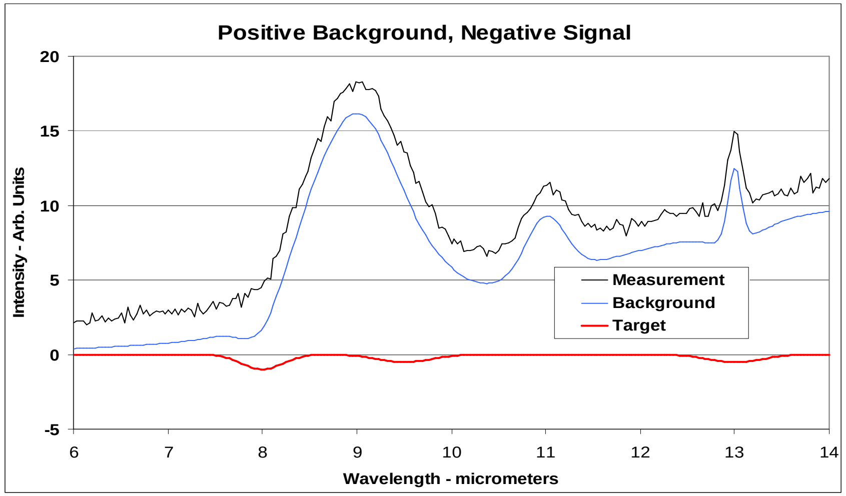


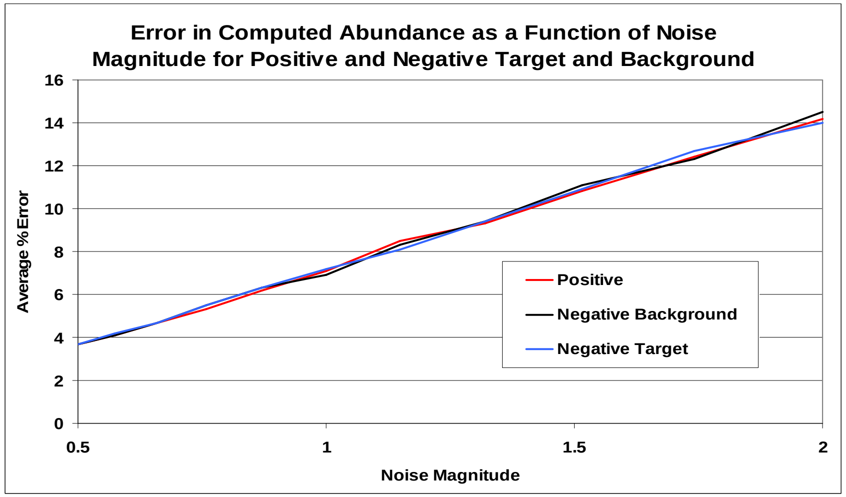


Works on Positive or Negative Signals

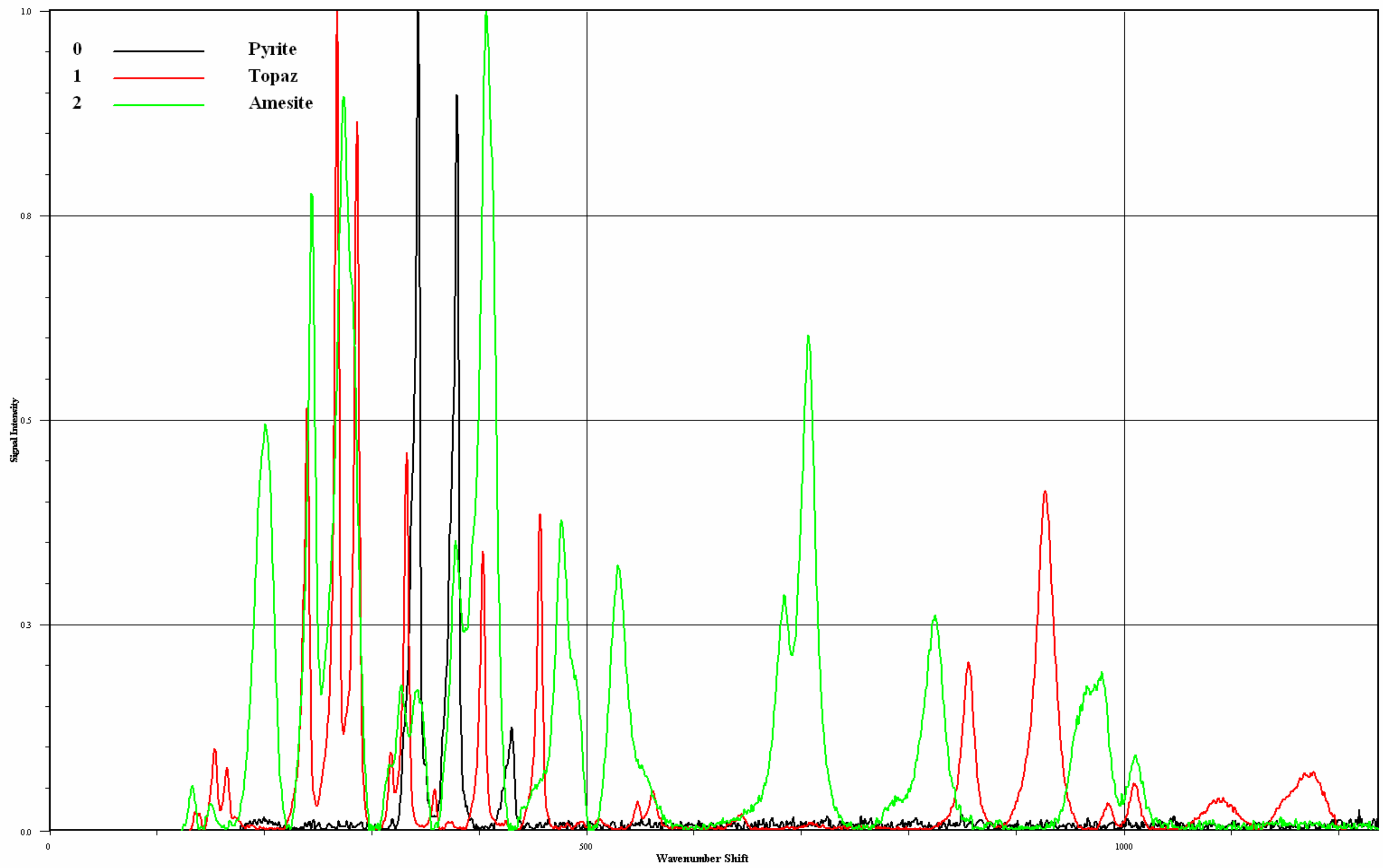




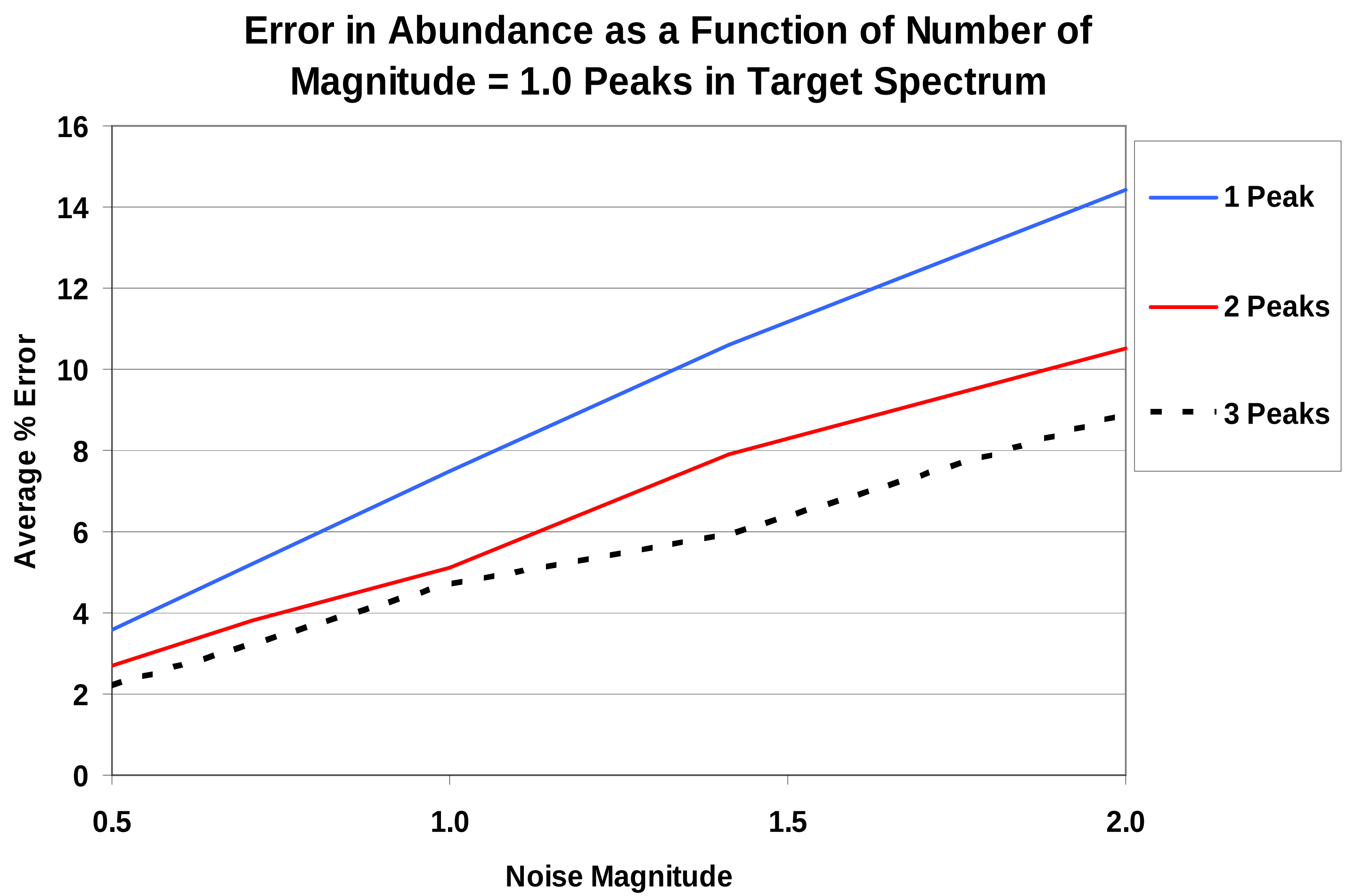




Raman Spectroscopy
Importance of those Other Peaks



Raman Spectroscopy
Importance of those Other Peaks



Cautions

- Potential Problems – Unknown Contaminants! Error proportional to:

  $$(\mathbf{OSP} * \mathbf{N}(\lambda)) \bullet \mathbf{d}(\lambda)$$

- Spectral Abundance may need to be Transformed into Species Abundance

- Condensed matter spectra rather than narrow emission/absorption lines.

Summary

- Powerful, Easy to Implement
- Removes Interference from Known Species including Fluorescence in Raman Spectroscopy
- No need to Compute Background Abundances
- Secondary Features More Useful
- High Spectral Resolution is Helpful ! !

References

- Duyen Nguyen and Eric Wu, “More Raman Analyses Less Fluorescence Interference”, Enwave Optronics, Inc.
- Harsanyi, J.C. and C.-I Chang (1994), "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," IEEE Trans. on Geoscience and Remote Sensing, vol. 32, no. 4, pp. 779-785, July, 1994.
- H. Ren and C.-I Chang, “Automatic spectral target recognition in hyperspectral imagery,” IEEE Trans. Aerosp. Electron. Syst., vol. 39, pp.1232–1249, Oct. 2003.
- Chang, C.-I. (2006) Information-Processed Matched Filters for Hyperspectral Target Detection and Classification, in *Hyperspectral Data Exploitation: Theory and Applications* (ed C.-I. Chang), John Wiley & Sons, Inc., Hoboken, NJ, USA.
- William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 1997
- *Infrared (2.1-25 um) Spectra of Minerals*, John W. Salisbury, Louis S. Walter, Norma Vergo and Dana M. D'Aria, Johns Hopkins University Press, 1992
- Downs R T (2006) “The RRUFF Project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals”, Program and Abstracts of the 19th General Meeting of the International Mineralogical Association in Kobe, Japan. O03-13, rruff.info