EXAFS Analysis with FEFF, LARCH, ARTEMIS

Matthew Newville

Center for Advanced Radiation Sources
The University of Chicago

21-July-2018

Slides for this talk: https://tinyurl.com/Larch2018
https://millenia.cars.aps.anl.gov/gsecars/Data/Larch/2018Workshop

The XAFS Equation used with FEFF:

$$\chi(k) = \sum_{j} \frac{S_0^2 N_j f_j(k) e^{-2R_j/\lambda(k)} e^{-2k^2\sigma_j^2}}{kR_j^2} \sin[2kR_j + \delta_j(k)]$$

- $f(k)$ and $\delta(k)$ are *photo-electron scattering amplitude and phase*:
  - ▶ Energy dependent $k \sim \sqrt{(E - E_0)}$.
  - ▶ Depend on Z of the scattering atom(s).
  - ▶ Non-trivial: must be calculated or carefully extracted from measured spectra.
- $\lambda(k)$ tells how far the photo-electron can travel.
- The sum is over *Scattering Paths* of the photo-electron, from absorbing atom to neighboring atoms and back.
  - ▶ May include *multiple scattering*.

> If we know $f(k)$, $\delta(k)$, and $\lambda(k)$, we can get:
> - R – near neighbor distance.
> - N – coordination number.
> - σ^2 – mean-square disorder in R.

$$\chi(k) = \sum_j \frac{S_0^2 N_j f_j(k) e^{-2R_j/\lambda(k)} e^{-2k^2\sigma_j^2}}{kR_j^2} \sin[2kR_j + \delta_j(k)]$$

Step 1 Calculate theoretical XAFS spectra with FEFF, starting with a rough guess of the local structure.

Step 2 *Refine* R, N, and σ^2 to best match experimental data.

Step 3 Compare lots of refined models.

Topics to try to cover:

- How do we run FEFF to generate $f(k)$, $\delta(k)$, and $\lambda(k)$?
- What Correction Factors do we need to worry about?
- How do we fit experimental data?
- How do we interpret the results?
- What tips and tricks can help fitting EXAFS data with FEFF calculations?

The normal scheme for using FEFF for EXAFS data analysis is:

1. Start with a structure close to the *local* atomic structure of your sample, and generate x,y,z coordinates for the atoms. Often a crystal structure is close – it does not have to be perfect!

2. Run FEFF. This creates a `feffnnnn.dat` files for each path in your structure.

3. Use these Path Files in to model measured XAFS in an analysis program.

ARTEMIS and SIXPACK makes this even easier: they run and manages FEFF for you.

> Having good starting structures can be important, but you do not need the exact structure to use FEFF.

> You can (and may need to) mix structural models to model real data.

Anatomy of `feff.inp`

FEFF is an old program that runs with a text input file that **must** be called `feff.inp`. No Kidding! Each run should be in its own subdirectory.

`feff.inp` file:

```
TITLE   FeO, rock salt structure

HOLE    1   1.0      # K Edge, S02=1.0
CONTROL 1 1 1 1      # which parts of code to run
PRINT   1 0 0 3      # which output files to write

POTENTIALS           # list of Atomic Potentials
* potential  z label
     0      26  Fe   # Absorbing Atom: Potential=0
     1       8  O    # 1 Potential for each Z
     2      26  Fe   # Fe scatterer (not absorber)

ATOMS  # list of Atomic X, Y, Z, Potential, Label
  0.00000    0.00000    0.00000    0   Fe
  0.00000    0.00000   -2.13870    1   O
 -2.13870    0.00000    0.00000    1   O
  0.00000   -2.13870    0.00000    1   O
  0.00000    0.00000    2.13870    1   O
  2.13870    0.00000    0.00000    1   O
  0.00000    2.13870    0.00000    1   O
  0.00000    2.13870    2.13870    2   Fe
  0.00000   -2.13870   -2.13870    2   Fe
 -2.13870    2.13870    0.00000    2   Fe
  2.13870   -2.13870    0.00000    2   Fe
  0.00000   -2.13870    2.13870    2   Fe
```

`feff.inp` includes:

- A list of unique Atomic Potentials:
  - Pot 0 = Absorbing Atom
  - 1 Potential per atomic species (Z)

- List of atomic coordinates:
  x, y, z, Pot
  for the cluster of atoms (in Å). The cluster does not need to be crystalline!

- There must be one absorbing atom. It does not have to be at (0, 0, 0).

- You can edit this file to alter the cluster. You may *need* to edit this file.

ATOMS and *atoms.inp*

One way to create *feff.inp* is with the ATOMS program which converts Crystal Structure to Atomic Cluster:

atoms.inp file:

```
TITLE    FeO, rock salt structure

space f m 3 m    # Space Group

a  = 4.2774      # Lattice Constant
rmax = 6.00      # Cluster Size
core = Fe1       # Central Atom

atom    # List of Cell Parameters
Fe   0.0000   0.0000   0.0000  Fe1
O    0.5000   0.5000   0.5000  O1
------
```

Using ATOMS is convenient, but not necessary!

ATOMS does not allow "doping", or fractional substitution – you can edit the *feff.inp* file for that, or combine paths from multiple FEFF runs.

The cluster used by FEFF does not need to be a crystal!

Atomic clusters may also come from:

Protein Data Bank, CIF Format data, Density Functional Theory Calculations, ...

FEFF writes out a separate file for each Scattering Path:
feff0001.dat, *feff0002.dat*, ...

feff.dat file:

```
FeO, rock salt structure                                    Feff XX 6.10
Abs   Z=26 Rmt= 1.197 Rnm= 1.402 K shell
Pot 1 Z= 8 Rmt= 0.937 Rnm= 1.103
Pot 2 Z=26 Rmt= 1.203 Rnm= 1.418
Gam_ch=1.325E+00 H-L exch
Mu=-4.210E+00 kf=2.067E+00 Vint=-2.048E+01 Rs_int= 1.755
Path   1      icalc     2
-------------------------------------------------------------------------
2   1.000   2.1387    2.4356   -4.21002 nleg, deg, reff, rnrmav(bohr), edge
      x        y        z    pot at
   0.0000   0.0000   0.0000  0  26 Fe      absorbing atom
   0.0000   2.1387   0.0000  1   8 O
 k   real[2*phc]  mag[feff]  phase[feff] red factor  lambda    real[p]
  0.000  2.9505E+00  0.0000E+00 -3.0508E+00  0.1009E+01  2.3773E+01  2.0669E+00
  0.200  2.9485E+00  9.9421E-02 -3.8738E+00  0.1009E+01  2.3876E+01  2.0762E+00
  0.400  2.9424E+00  1.9477E-01 -4.6322E+00  0.1009E+01  2.4141E+01  2.1037E+00
  0.600  2.9320E+00  2.8278E-01 -5.3276E+00  0.1010E+01  2.4444E+01  2.1488E+00
  0.800  2.9172E+00  3.6142E-01 -5.9621E+00  0.1011E+01  2.4599E+01  2.2105E+00
  1.000  2.8976E+00  4.2991E-01 -6.5379E+00  0.1012E+01  2.4410E+01  2.2878E+00
  1.200  2.8723E+00  4.8834E-01 -7.0570E+00  0.1014E+01  2.3748E+01  2.3798E+00
```

This file contains:

1. Path geometry
2. Data for $f(k)$, $\delta(k)$, and $\lambda(k)$.

You *never* have to read this file – the analysis program will do that.

You may have to move or rename files.

Analysis programs will use these files to model EXAFS as a *Sum of Paths*.

FEFF: what *should* you worry about?

Using FEFF in a real analysis, we'll be *refining* bond lengths and coordination numbers. For this to work, here are some "rules of thumb":

1. refining distances by **more than** $0.1\,\text{Å}$ probably means the calculation should be re-run – the overlap of atomic potentials may not be accurate.

2. refining energy origins E_0 by **more than** $10\,\text{eV}$ may mean the calculation – or the selection of E_0 in the data reduction – is wrong or needs to be re-examined.

3. FEFF can set S_0^2 and σ^2 values. **don't do this**.

4. You **should set** S_0^2 for a particular set of data (central atom, beamline energy resolution, etc), based on an experimental standard, or you will have uncertain values for N.

Structural Disorder and the Pair Distribution Function

An EXAFS measurement averages billions of *snapshots* of the local structure:

- Each absorbed x-ray generates 1 photo-electron.
- the photo-electron / core-hole pair lives for about 10^{-15} s – much faster than the thermal vibrations (10^{-12} s).
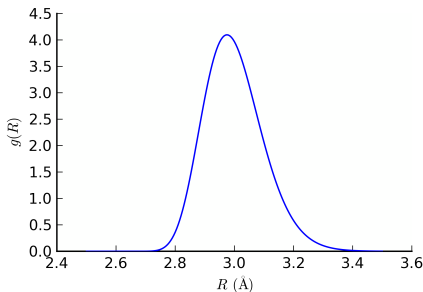- An EXAFS measurement samples 10^4 (dilute fluorescence) to 10^{10} absorbed x-rays for each energy point.

So far, we've put this in the EXAFS Equation as $\chi \sim N \exp(-2k^2\sigma^2)$

More generally, EXAFS samples the

Partial Pair Distribution Function

$g(R) =$ probability that an atom is a distance R away from the absorber.

EXAFS and The Pair Distribution Function

To fully account for a highly disordered local structure, we should use

$$\chi(k) = \left\langle \sum_j \frac{f_j(k) e^{i2kR_j + \delta_j(k)}}{k R_j^2} \right\rangle$$

where $\langle x \rangle = \int dR\, x\, g(R) / \int dR\, g(R)$ – averaging over the billions+ of snapshots.

R won't change too much, so we'll neglect the changes to $1/R^2$:

$$\chi \approx \sum_j f_j(k) \frac{e^{i\delta_j(k)}}{k R_j^2} \left\langle e^{i2kR_j} \right\rangle$$

each path in the sum now has a $g(R)$ with respect to the absorbing atom.

The *the cumulant expansion* relates $\langle e^x \rangle$ to $\langle x \rangle$, the moments of $g(x)$:

$$\left\langle e^{i2kR} \right\rangle = \exp\left[\sum_{n=1}^{\infty} \frac{(2ik)^n}{n!} C_n \right].$$

The Cumulants and Moments of a Distribution Function

The cumulants C_n of $g(R)$ are related to the moments of $g(R)$: $\langle r^n \rangle$, with $r = R - R_0$ and R_0 is the centroid of the distribution:

$$C_1 = \langle r \rangle = \Delta R$$
$$C_2 = \langle r^2 \rangle - \langle r \rangle^2 = \sigma^2$$
$$C_3 = \langle r^3 \rangle - 3\langle r^2 \rangle\langle r \rangle + 2\langle r \rangle^3$$
$$C_4 = \langle r^4 \rangle - 3\langle r^2 \rangle^2 - 4\langle r^3 \rangle\langle r \rangle + 12\langle r^2 \rangle\langle r \rangle^2 - 6\langle r \rangle^4$$

C_3 (the *third cumulant*) can be important For many EXAFS problems.

Sometimes, the cumulant expansion isn't good enough. One can build more complicated models . . . .

Example:

Use many closely spaced paths to allow for nearly any distribution.

Model the amplitude of each Path.
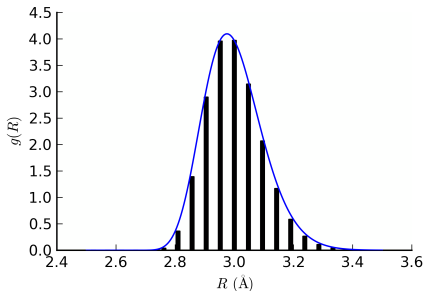
To model XAFS as a Sum of Paths:

$$\chi(k) = \sum_j \frac{S_0^2 N_j f_j(k) e^{-2R_j/\lambda(k)} e^{-2k^2\sigma_j^2}}{kR_j^2} \sin[2kR_j + \delta_j(k)]$$

we may refine these Parameters *For Each Path*:

In XAFS Equation	IFEFFIT Parameter	Physical Meaning
NS_0^2	amp	Amplitude Factor: N and S_0^2
E_0	e0	Energy Shift (where $k = 0$)
ΔR	delr	Change in path length $R = \Delta R + R_{\text{eff}}$
σ^2	sigma2	Mean-square-displacement in R

- R_{eff} is the starting R value for the FEFF Path.
- Other Parameters: higher order cumulants, energy broadening, . . .
- In principle, any parameter for any path could be refined.

FeO has a rock-salt structure.

To model the Fe K edge EXAFS of FeO, we'll calculate the `feffNNNN.dat` files (with $f(k)$ and $\delta(k)$), for Fe-O based on the FeO crystal structure.

We'll use these functions to *refine* the values R, N, σ^2, and E_0 so our model EXAFS function matches our data.

Fe-O octahedra,
$R = 2.14\,\text{Å}$.

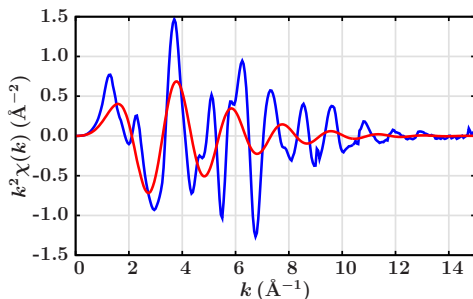Results:

$$
\begin{aligned}
S_0^2 &= 0.75 \text{ (fixed)} \\
N &= 4.7 \pm 0.4 \\
R &= 2.10 \pm 0.01\text{Å} \\
\Delta E_0 &= \text{-3.5} \pm 1.0 \text{ eV} \\
\sigma^2 &= 0.011 \pm 0.001 \ \text{Å}^2.
\end{aligned}
$$
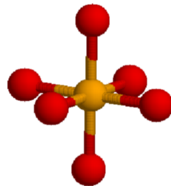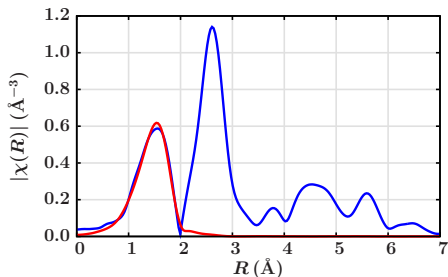
$|\chi(R)|$ for FeO data and 1st shell fit.

1st shell fit in k space.

There is clearly another component in the XAFS!
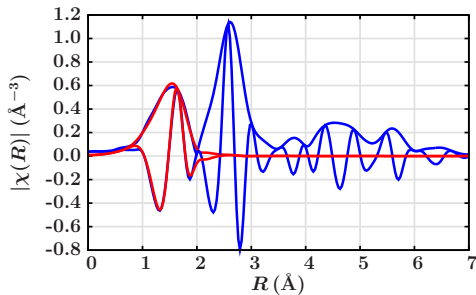
1st shell fit in R space.

$|\chi(R)|$ and $\mathrm{Re}[\chi(R)]$ for FeO (blue), and a 1st shell fit (red).

Though the fit to the magnitude isn't perfect, the fit to $\mathrm{Re}[\chi(R)]$ looks very good.

Fitting Strategies

Data analysis seeks a *Model* that best matches a *Measurement*.

We'll use χ^2 (don't confuse with EXAFS χ!!) to describe how good the match is:

$$\chi^2 = \sum_i^{N_{\rm fit}} \frac{[\chi_i^{\rm measured} - \chi_i^{\rm model}(x)]^2}{\epsilon^2}$$

where

- $N_{\rm fit}$ = number of points in the data to fit.
- ϵ = the estimated noise level in the data.
- x is the set of parameters to be varied in the analysis

The Best Fit is the one with lowest χ^2.

Questions:

1. How do I know how many independent measurements I have?
2. What is ϵ for my data?
3. What parameters can/should I vary?

The number of parameters we can reliably extract from our data is limited:

$$N_{\mathrm{idp}} \approx \frac{2 \Delta k \Delta R}{\pi}$$

where Δk and ΔR are the k- and R-ranges of the usable data.

For a typical range of $k = [3.0, 12.5]\,\text{Å}^{-1}$ and $R = [1.0, 3.0]\,\text{Å}$, there are ~ 12 parameters that can be determined from EXAFS. That's not much!

The Fit statistics and confidence in the measured parameters need to reflect this. But we usually oversample our data ($N_{\mathrm{fit}} > N_{\mathrm{idp}}$) so we have

$$\chi^2 = \frac{N_{\mathrm{idp}}}{\epsilon^2 N_{\mathrm{fit}}} \sum_{i}^{N_{\mathrm{fit}}} [\chi_i^{\mathrm{measured}} - \chi_i^{\mathrm{model}}(x)]^2$$

Note: I also assumed ϵ is a constant.

Other "Goodness of Fit" Statistics

chi-square:

$$\chi^2 = \frac{N_{\mathrm{idp}}}{\epsilon^2 N_{\mathrm{fit}}} \sum_i^{N_{\mathrm{fit}}} [\chi_i^{\mathrm{measured}} - \chi_i^{\mathrm{model}}(x)]^2$$

reduced chi-square: scale N_{varys} by the "degrees of freedom" :

$$\chi_\nu^2 = \chi^2 / (N_{\mathrm{idp}} - N_{\mathrm{varys}})$$

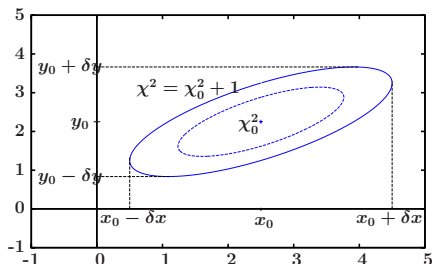For a "Good Fit", χ_ν^2 should be ~ 1. This *never* happens!

R-factor: $\mathcal{R}$ gives a "fractional misfit" (and not scaled by the data uncertainty ϵ):

$$\mathcal{R} = \frac{\sum_i^{N_{\mathrm{fit}}} [\chi_i^{\mathrm{measured}} - \chi_i^{\mathrm{model}}(x)]^2}{\sum_i^{N_{\mathrm{fit}}} [\chi_i^{\mathrm{measured}}]^2}$$

A fit finds a set of values x_0 that are the "best fit" of the variables x – they give the lowest value of χ^2.

Uncertainties in x are estimated by increasing the χ^2 by 1:

Some Parameters are *Correlated*:

Changing x away from its best value will change the best value for y.

For EXAFS, (R, E_0) and (N, σ^2) are usually very highly correlated (> 0.85).

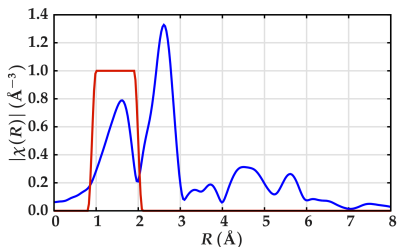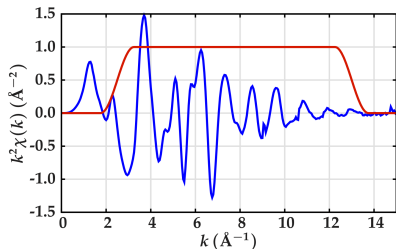*Increasing χ^2 by 1 assumes we have a "Good Fit", with $\chi^2_\nu \approx 1$.*

We typically have $\chi^2_\nu \gtrsim 50$ (Yikes!).

We increase the best χ^2 by χ^2_ν to estimate error bars.

The χ^2 definition didn't say anything about what our data $\chi_i^{\mathrm{measured}}$ actually is ...

We usually fit in R-space, so that we can select which "shells" to ignore:

Fitting $\chi(R)$ (both real and imaginary parts!) gives more meaningful fit statistics – we know that we're not fitting all the spectral features.
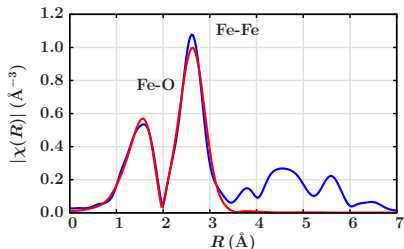
Plus: We can have $\chi_i^{\mathrm{measured}}$ extend over

<p align="center">multiple data sets, multiple k-weightings, etc,</p>

as long as we generate the corresponding $\chi_i^{\mathrm{model}}(x)$ to match these data.

Adding the 2nd shell Fe $-$ `feffNNNN.dat` for Fe-Fe $-$ and refining R, N, σ^2:

$|\chi(R)|$ data for FeO (blue), and fit of 1st and 2nd shells (red).

These results are consistent with the known values for FeO:
6 O at 2.13Å, 12 Fe at 3.02Å.

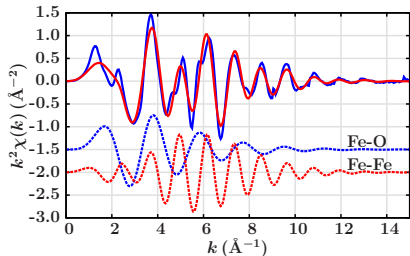Fit results: Statistics: $R \approx 0.015$ $\chi^2_\nu \approx 165$.

Shell	N	R (Å)	σ^2 (Å^2)	ΔE_0 (eV)
Fe-O	5.1(0.8)	2.10(.01)	0.012(.003)	-1.4(0.9)
Fe-Fe	12.6(1.6)	3.07(.01)	0.013(.001)	-1.4(0.9)

These are typical even for a "very good fit" on known structures.
The calculation for $f(k)$ and $\delta(k)$ are good, but not perfect!
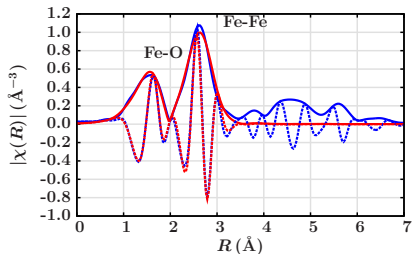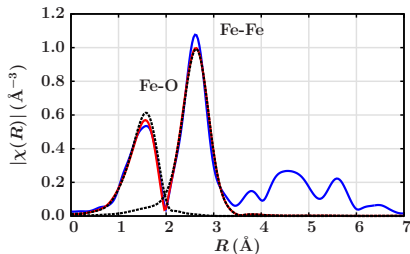
EXAFS Analysis: Second Shell of FeO

Other views of the data and fit:

The Fe-Fe EXAFS extends to higher-k than the Fe-O EXAFS.

Even in this simple system, there is some *overlap* of shells in R-space.

The fit in $\mathrm{Re}[\chi(\mathrm{R})]$ look especially good – this is how the fits are done.

The EXAFS Equation has at least 4 adjustable parameters *Per Path*: E_0, NS_0^2, R, and σ^2.   But N_{idp} is low:

$$N_{\mathrm{idp}} = 8 \ \text{ for } \Delta R = 1\,\text{Å} \text{ and } \Delta k = 12.5\,\text{Å}^{-1}$$

For simple crystalline structures with well-isolated, single-scattering path (like FeO), it's OK to fit N, R, σ^2, and E_0 for every path.

For more complicated problems, we need a way to limit the number of parameters varied.

We might *want* to impose relationships between parameters to get more meaningful results. . .

Constraints and Generalized Variables

Instead of varying the Path Parameters directly, we write them in terms of *Generalized Variables*. This allows simple *Constraints* and model building:

Parameter=Variable

```
# one e0 for 2 paths
guess e0      = 1.0
path(1,  feff= feo.dat,   e0 = e0)
path(2,  feff= fefe.dat,  e0 = e0)
```

Mixed Coordination Shell

```
# mix O and S in 1st coordination shell
set  SO2    = 0.80
guess xSulfur = 0.5

path(1,  feff= feo.dat, amp=SO2* xSulfur )
path(2,  feff= fes.dat, amp=SO2* (1-xSulfur) )
```

Einstein Temperature

```
# 1 parameter to set sigma2 for all paths
set   factor = 24.254337   #= (hbar*c)^2/(2 k_boltz)
# mass and reduced mass in amu
set   mass1 = 63.54,  mass2 = 63.54
set   r_mass =  1/ (1/mass1 +  1/mass2)

# the Einstein Temp will be adjusted in the fit!
guess thetaE = 200

# use for data set 1, T=77
set   temp1 = 77
def ss2_path1 = factor*coth(thetaE/(2*temp1))/r_mass )
path(101,  sigma2 = ss2_path1   )

# use for data set 2, T=300
set   temp2 = 300
def ss2_path2 = factor*coth(thetaE/(2*temp2))/r_mass )
path(201,  sigma2 = ss2_path2   )
```

This allows us to use *Prior Knowledge* into the data analysis, and consider more complicated problems:

- force one R for the same bond for data taken from different edges.
- model complex distortions (height of a sorbed atom above a surface).

Example: Cu metal at 3 temperature

A very simple example of a Multi-Data-Set Fit:
Cu metal, at 3 different temperatures: 10K, 50K 150K.

Path Parameters:

- E_0: Same for all T
- S_0^2 Same for all T
- R: expands linearly with T (slope + offset).
- σ^2: goes as Einstein temperature (as before).

12 parameters become 5.

Fit range:

$R = [1.60, 2.75]\,\text{Å}$

$k = [1.50, 18.50]\,\text{Å}^{-1}$

Cu at three temperatures

```
guess s02    = 0.93
guess theta  = 240
guess e0     = 3.5
guess alpha  = 0.0
guess a_t    = 0.00

path(index  = 101, feff  = feffcu01.dat,
     label  = "Cu metal first shell, for 10K",
     s02    = s02,
     sigma2 = eins(10, theta),
     delr   = reff * (alpha + 10.0 * a_t) ,
     e0     = e0)

path(index  = 201, feff  = feffcu01.dat,
     label  = "Cu metal first shell, for 50K",
     s02    = s02,
     sigma2 = eins(50, theta),
     delr   = reff * (alpha + 50.0 * a_t) ,
     e0     = e0 )

path(index  = 301, feff  = feffcu01.dat,
     label  = "Cu metal first shell, for 150K",
     s02    = s02,
     sigma2 = eins(150, theta),
     delr   = reff * (alpha + 150.0 * a_t) ,
     e0     = e0 )
```
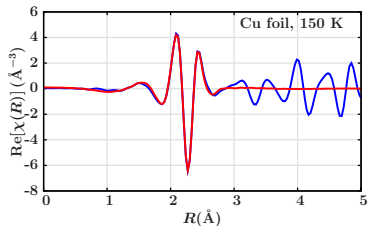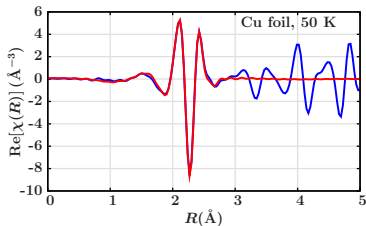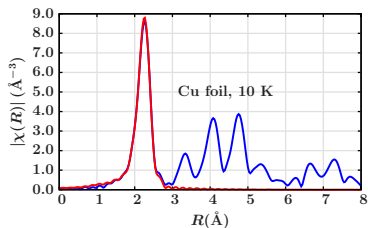
Example: Cu metal Results

$S_0^2 = 0.91(0.08)$
$\theta = 233.5(19.6)\,\text{K}$
$e_0 = 0.4(1.3)\,\text{eV}$
$\alpha = 0.002(0.003)\,\text{Å/K}$
$a_t = 0.5(1.8) \times 10^{-5}\,\text{Å}$

Cu foil, 10 K

Cu foil, 50 K

Cu foil, 150 K

Conclusions

EXAFS Analysis is challenging:

1. Lots of Paths.
2. $f(k)$, $\delta(k)$ are not trivial.
3. There's not much information in a real measurement: $N_{\mathrm{idp}} \approx \frac{2\Delta k \Delta R}{\pi}$

We need to:

1. reduce the number of Paths to consider (Fourier analysis).
2. re-use FEFF calculations for $f(k)$ and $\delta(k)$.
3. recycle parameters to cut down the number of independent variables in the fit, while keeping a meaningful analysis.

Build physical models with *Constraints*

Study the error bars and correlations.

Find the simplest interpretation of your data.