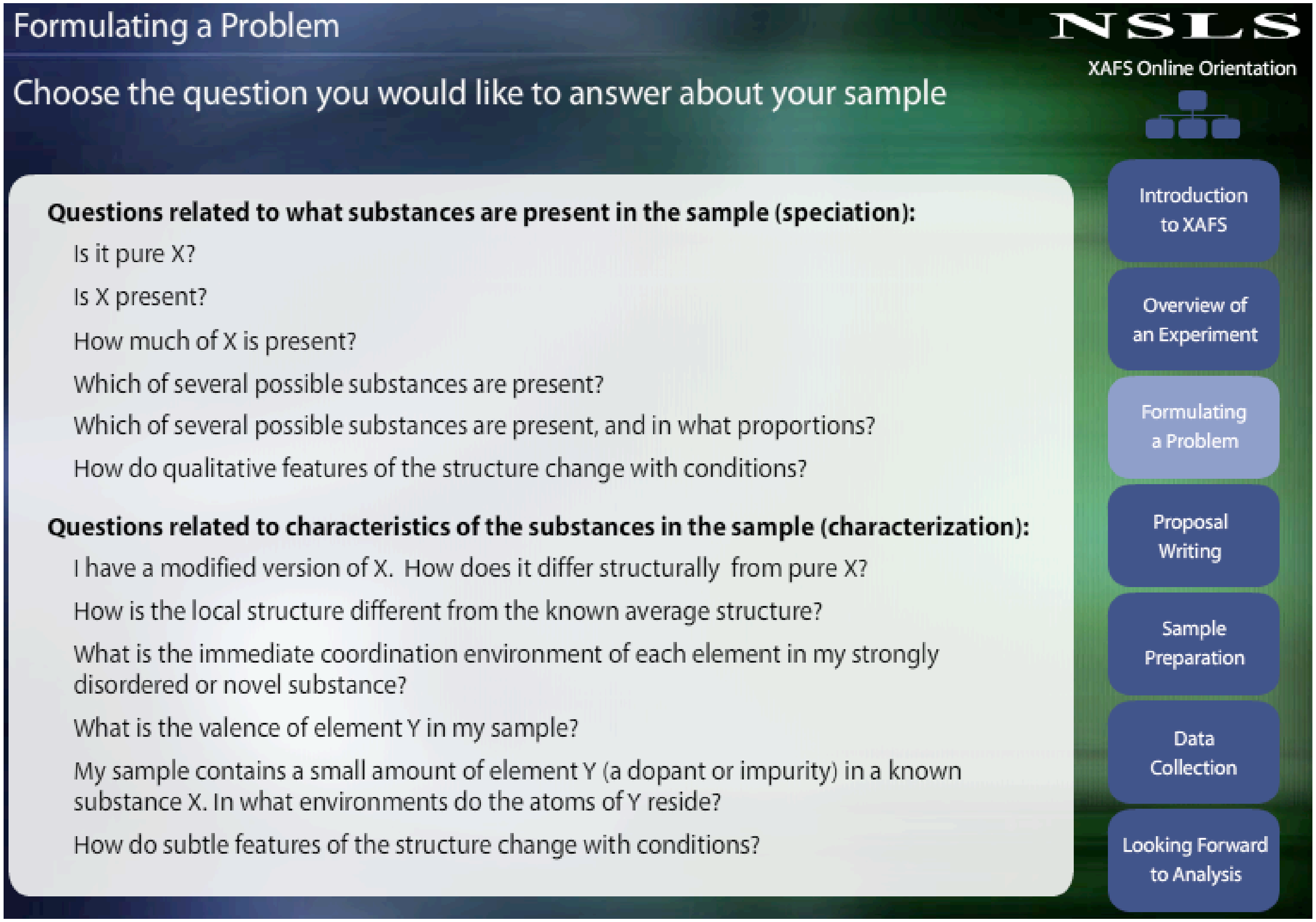
SARAH · LAWRENCE · COLLEGE

EXAFS is not a Black Box or How I Learned to Stop Worrying and Love the Fit

Scott Calvin

Sarah Lawrence College
Stanford Synchrotron Light Source

Some Typical Questions



Formulating a Problem

Choose the question you would like to answer about your sample

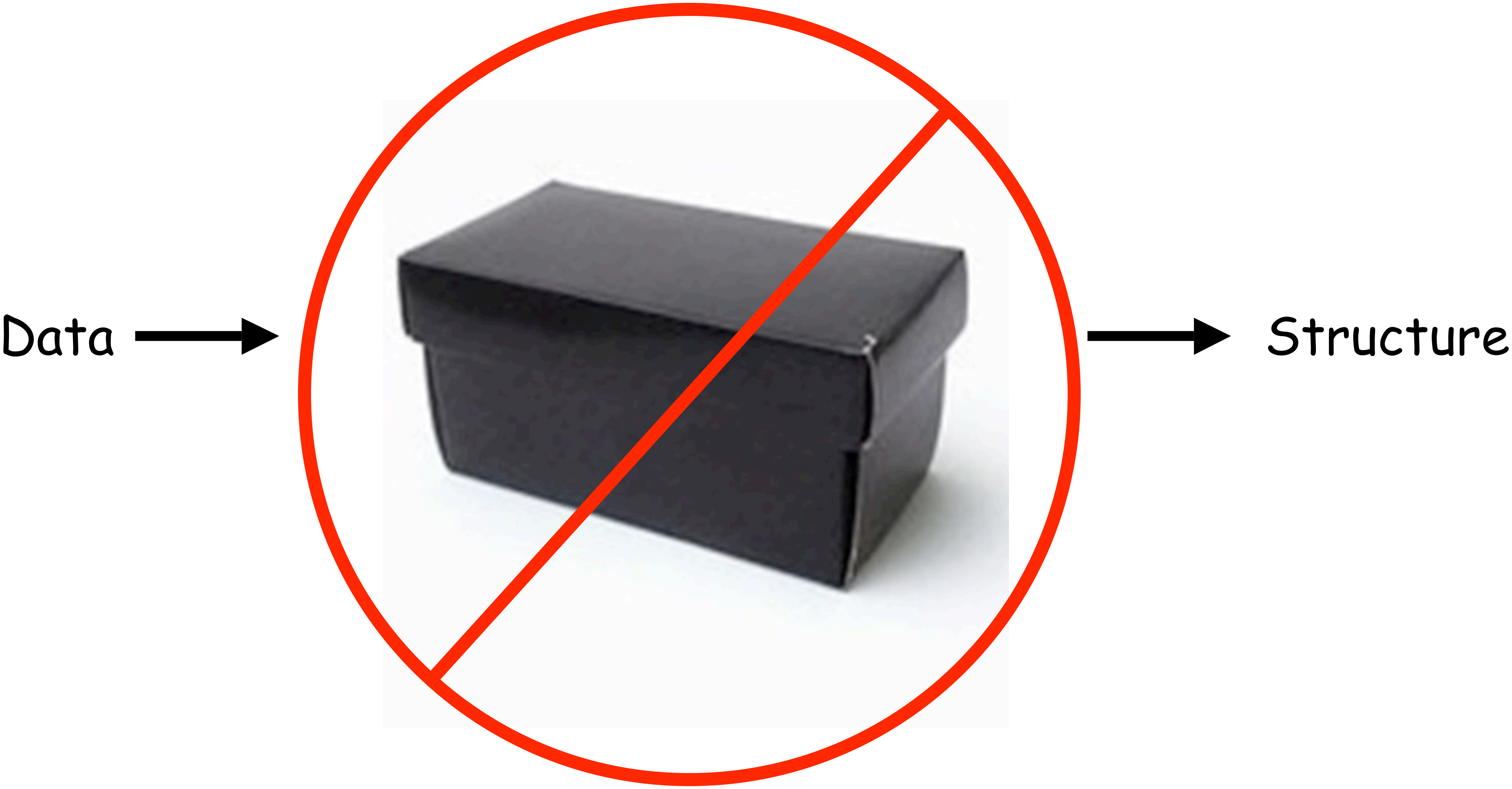
Questions related to what substances are present in the sample (speciation):

- Is it pure X?
- Is X present?
- How much of X is present?
- Which of several possible substances are present?
- Which of several possible substances are present, and in what proportions?
- How do qualitative features of the structure change with conditions?

Questions related to characteristics of the substances in the sample (characterization):

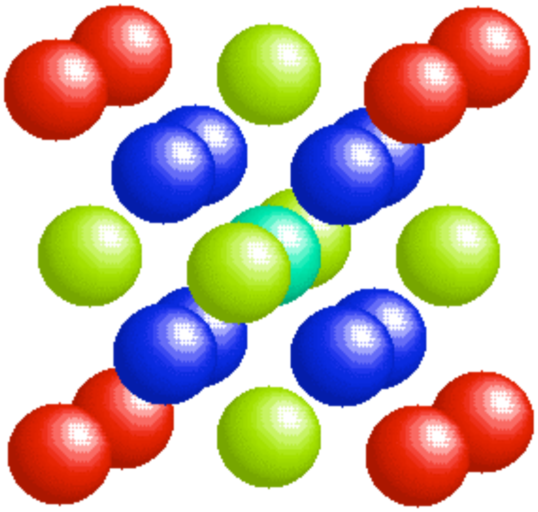
- I have a modified version of X. How does it differ structurally from pure X?
- How is the local structure different from the known average structure?
- What is the immediate coordination environment of each element in my strongly disordered or novel substance?
- What is the valence of element Y in my sample?
- My sample contains a small amount of element Y (a dopant or impurity) in a known substance X. In what environments do the atoms of Y reside?
- How do subtle features of the structure change with conditions?

Why XAFS is not a Black Box

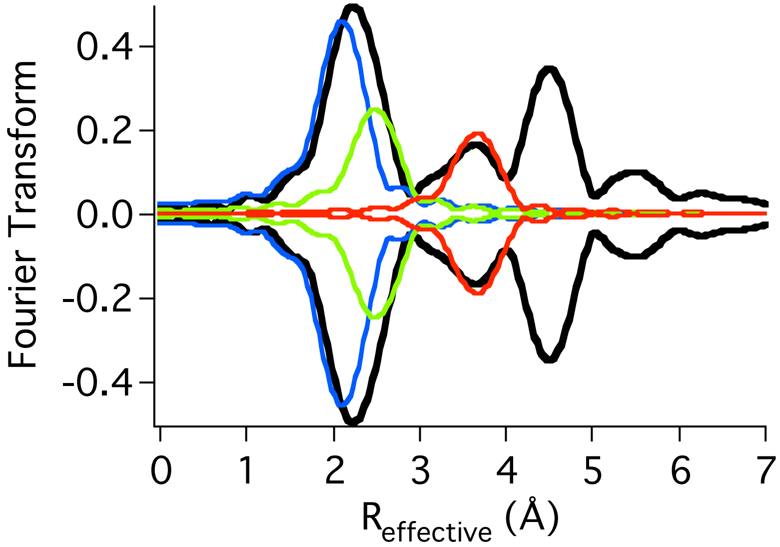


The Inverse Problem

This is "easy":



Structure

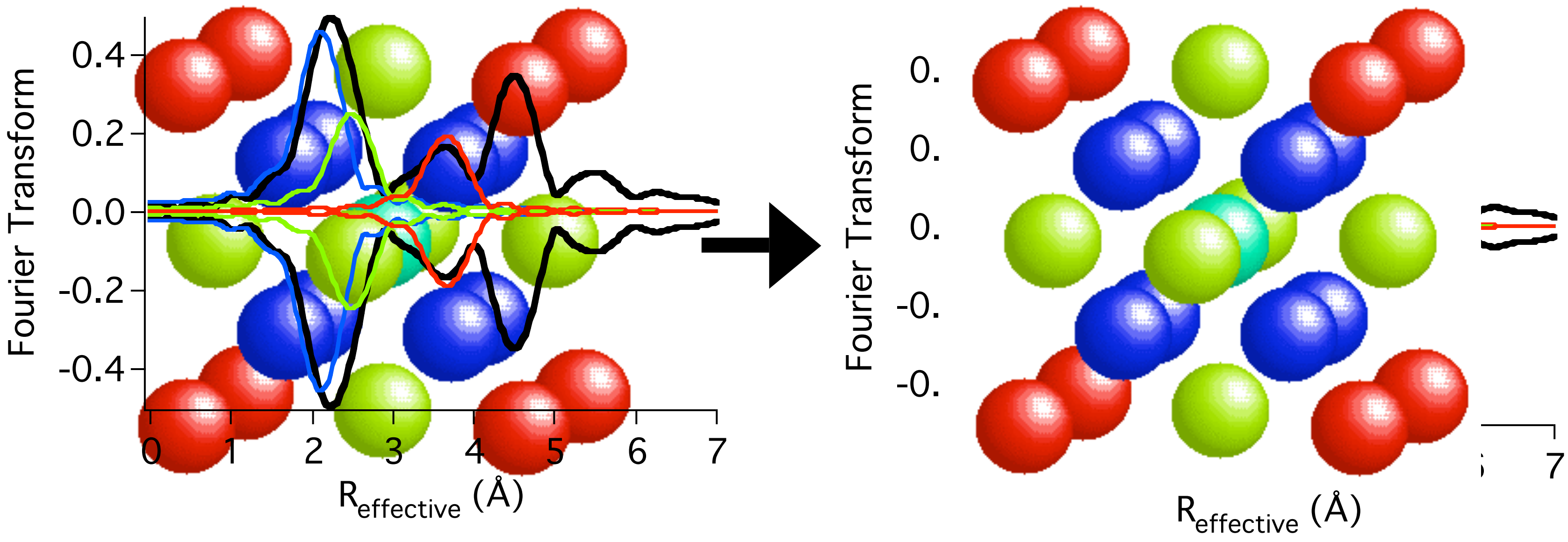


Fourier Transform of EXAFS

FEFF does this

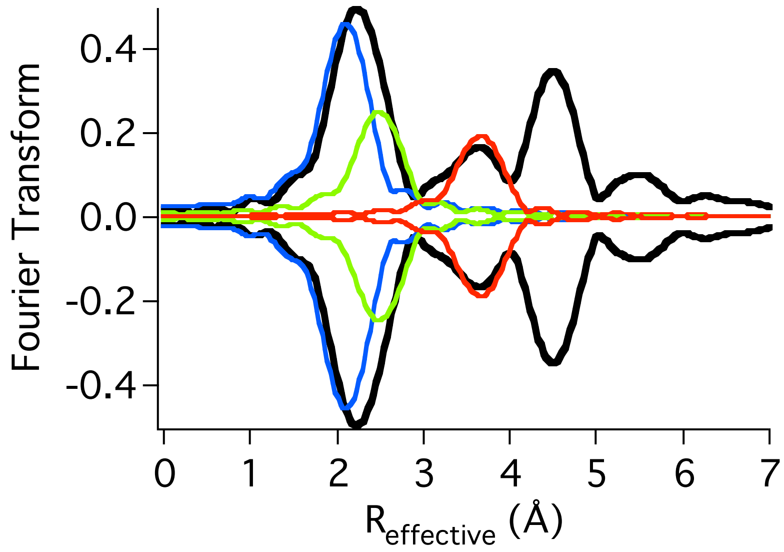
The Inverse Problem

This is "hard":



What Makes it Hard?

- The peaks of the Fourier Transform are shifted down from the actual absorber-scatterer distances, typically by 0.3 to 0.5 angstroms.

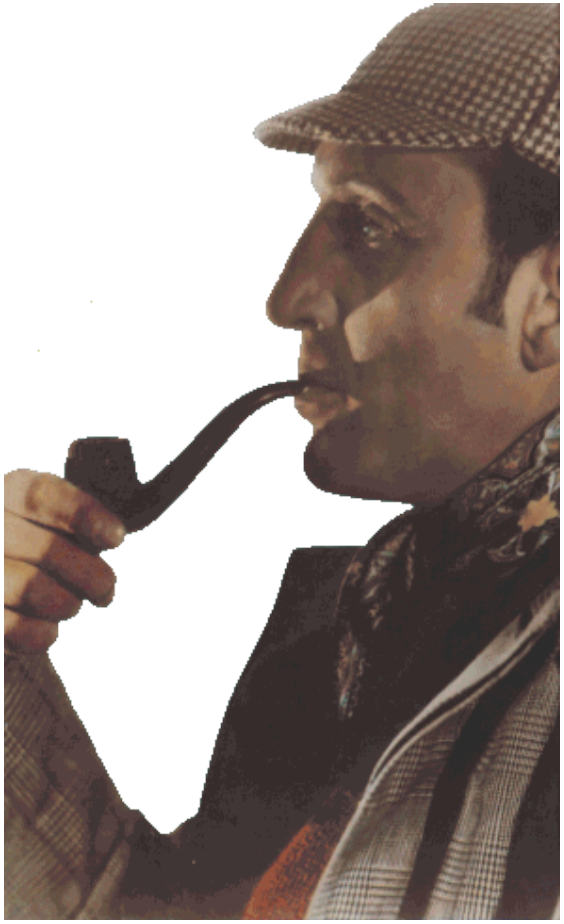


- Multiple-scattering may be significant
- **More than one structure may correspond to the same spectrum!**

In a sense, the task is not just hard, it is impossible!

SARAH·LAWRENCE·COLLEGE

Detective Work



While XAFS cannot yield a structure absent other information, it can rule out structures, so:

Make an **educated guess** as to possible structures, then use XAFS to eliminate those that are incorrect.

"When you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth."--Sherlock Holmes

XAFS Analysis Techniques

There are four common techniques for analyzing XAFS data. All rely, in one way or another, on making educated guesses as to what is in the material:

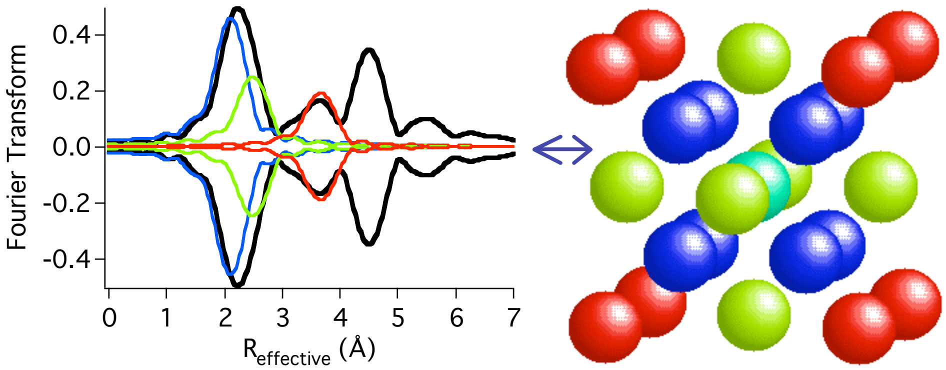
- Fingerprinting
- Linear combination analysis
- Principle component analysis
- Modeling ("Curve fitting to a theoretical standard")

Analysis Technique: Modeling (Curve Fitting to a Theoretical Standard)

In this technique, a computer calculates a theoretical spectrum based on a guess as to the structure of the material, and then adjusts the guessed structure so as to match the measured spectrum.

This is the only technique which can provide information about structures for which there are no good standards, and is the only way of reliably measuring quantitative information such as bond lengths.

This is also the most difficult technique to learn and apply.



Useful for novel materials, particularly modifications or variants of well-known structures

The EXAFS Equation

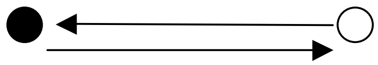
$$\chi(k)=\sum_j \frac{S_0^2 N_j f_j(k) e^{-2R_j/\lambda(k)} e^{-2k^2\sigma_j^2}}{kR_j^2}\sin[2kR_j+\delta_j(k)]$$

In the EXAFS region, the oscillations can be thought of as a sum of individual scattering paths *j*.

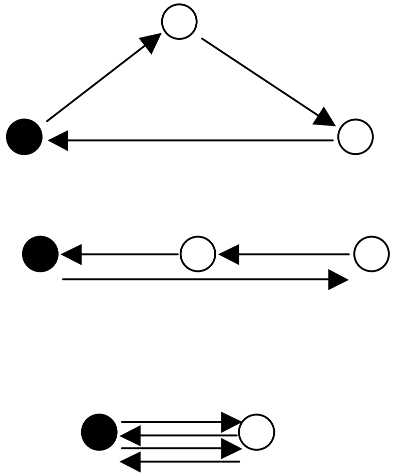
- R_j : the absorber-scatterer distance
- σ^2_j : the variance in the absorber-scatterer distance for path j
- N_j : the coordination number for path j
- S_o^2 : the “amplitude reduction factor”
- f_j, δ_j: scattering parameters based on identity of scattering atom(s)
- λ: mean-free path of the photoelectron

For multiple-scattering paths, the parameters are defined in analogous fashion; i.e. R_j is half of the total path length)

Single-scattering:

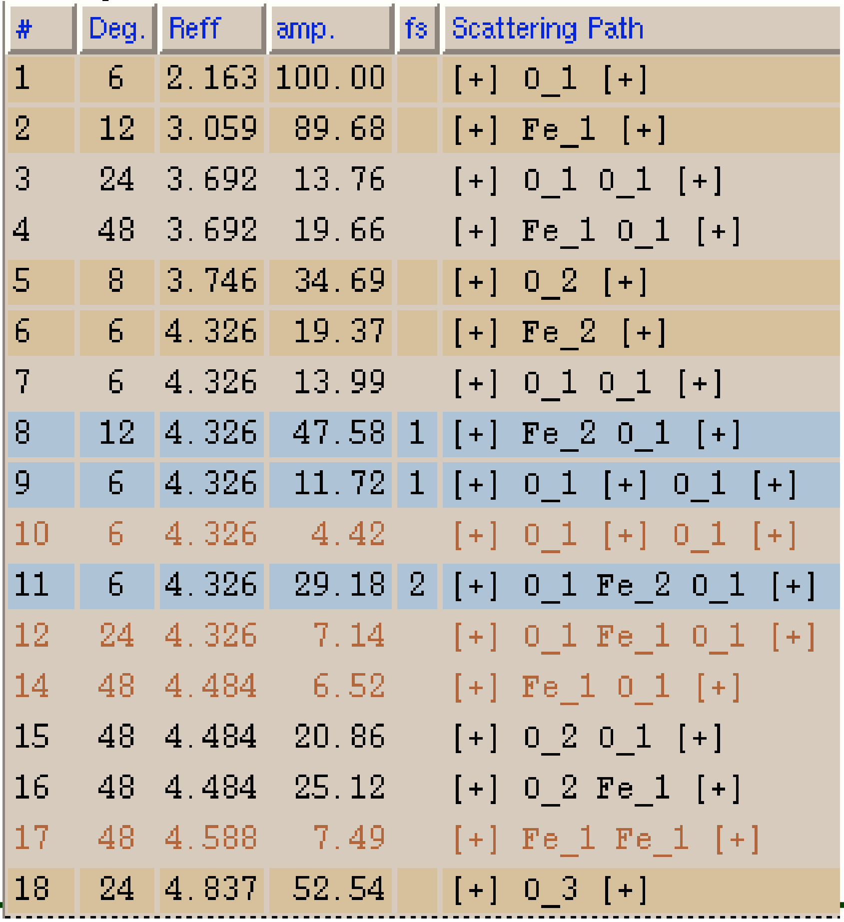


Multiple scattering:



Example: FeO

For FeO, FEFF generates the following list of paths out to a distance of 5 angstroms:



#	Deg.	Reff	amp.	fs	Scattering Path
1	6	2.163	100.00		[+] O_1 [+]
2	12	3.059	89.68		[+] Fe_1 [+]
3	24	3.692	13.76		[+] O_1 O_1 [+]
4	48	3.692	19.66		[+] Fe_1 O_1 [+]
5	8	3.746	34.69		[+] O_2 [+]
6	6	4.326	19.37		[+] Fe_2 [+]
7	6	4.326	13.99		[+] O_1 O_1 [+]
8	12	4.326	47.58	1	[+] Fe_2 O_1 [+]
9	6	4.326	11.72	1	[+] O_1 [+] O_1 [+]
10	6	4.326	4.42		[+] O_1 [+] O_1 [+]
11	6	4.326	29.18	2	[+] O_1 Fe_2 O_1 [+]
12	24	4.326	7.14		[+] O_1 Fe_1 O_1 [+]
14	48	4.484	6.52		[+] Fe_1 O_1 [+]
15	48	4.484	20.86		[+] O_2 O_1 [+]
16	48	4.484	25.12		[+] O_2 Fe_1 [+]
17	48	4.588	7.49		[+] Fe_1 Fe_1 [+]
18	24	4.837	52.54		[+] O_3 [+]

Note that there are five single scattering paths, and a number of additional multiple scattering paths with significant amplitude

Since IFEFFIT can fit four parameters per path ($S_0^2N_j$, ΔR_j, σ^2_j, *and* ΔE_o), that's 20 parameters for the single scattering paths alone. Include all the multiple scattering paths shown and the total balloons to 72 parameters. Does EXAFS really contain that much information?

Information Content of EXAFS

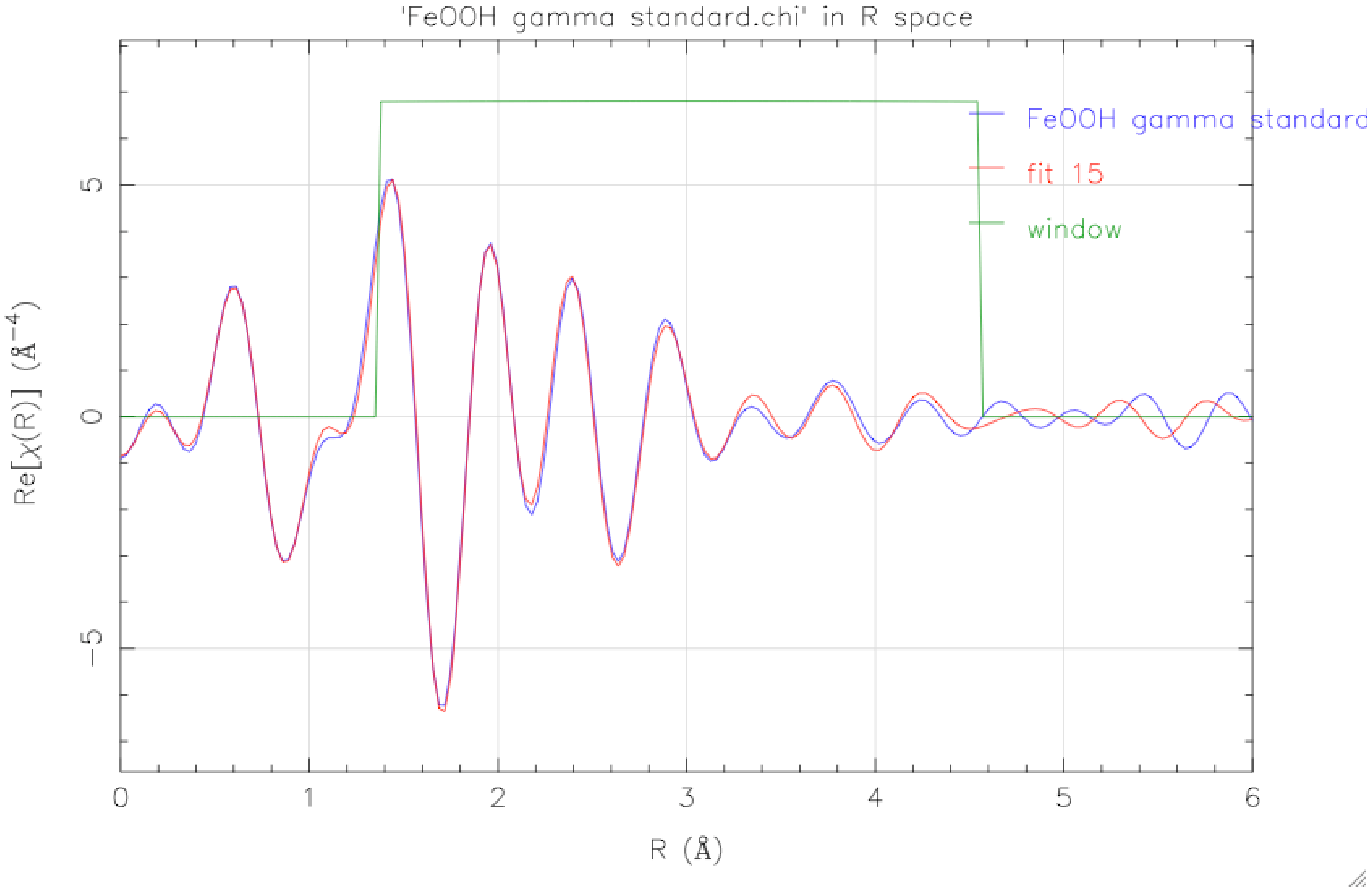
According to information theory, the maximum amount of information contained in an EXAFS Fourier transform is given by:

$$N \approx \frac{2\Delta k \Delta R}{\pi}$$

Where Δk is the range of data selected in *k*-space and ΔR is the range of the Fourier transform being fit.

For good (but not great) EXAFS data, Δk might typically be 10 Å^{-1} and ΔR might be 4 Å, yielding about 25 independent points. So, including multiple scattering paths, we do not have enough information to independently vary every parameter in the FeO example on the previous page.

An Example of a False Fit



The Solution: Science!

We're not mathematicians, and we're not data technicians. We're all scientists of one type or another. That means we have additional information about our system and our world. We can use that information to cut down on the number of parameters being fit.

There are at least two ways to do this: **constraints** and **restraints**. Today we'll talk a little about constraints; restraints are less commmonly used; ask me at the end if you'd like to know more about them.

SARAH · LAWRENCE · COLLEGE

A Very Simple Example of a Constraint

It seems plausible that the energy origin of the photoelectron does not depend on what the electron then scatters off of; i.e.

ΔE_o should be the same for every path.

If 18 paths are being fit, that reduces the number of free parameters by 17!

But How Do I Know A Constraint is Justified?

It has been argued, for instance, that it may actually help to use different ΔE_o's for different paths, perhaps as a way of compensating for approximations made by FEFF.

Using analysis software, it is not difficult to try fits with and without a given constraint. If the constraint is not appropriate, the statistical quality of the fit (as measured by something called "reduced chi-squared"), should be significantly better without the constraint.

More Examples of Simple Constraints

- Perfect crystals: if a crystal is not defective, the coordination number for all paths may be known *a priori*
- Lattice expansion: for substances with high symmetry, it's possible that the sample differs from the model structure by a constant expansion factor
- Similar paths behave similarly: oxygen paths "a long way" from the absorbing atom may all have similar values of σ^2, for instance

More Complicated Constraints

More complicated mathematical expressions to be used as constraints. Just a few examples:

- Geometric: for example, an orthorhombic distortion will change the lengths of paths in fairly complicated but specific ways. As another example, individual ligands may be fairly rigid, but number and orientation may be unknown.
- Thermal: known patterns of temperature dependence (e.g. Debye laws) may reduce the number of free parameters when data is collected as a function of temperature
- Doping: principles like Vegard's law may be helpful dealing with doped crystals
- Stoichiometry: if the chemical composition of a sample is known, it may sometimes be used to constrain average coordination numbers
- Size/morphology of nanomaterials can also be used to constrain average coordination numbers
- Almost any relationship between parameters that can be expressed mathematically can be used as a constraint!

Common Fitting Strategy #1: Top Down

This one works well for substances which are expected to be fairly ordered and well known.

For instance, a material scientist might be investigating ferrite nanoparticles, with an interest in how they differ from the bulk.

- Start with a highly constrained model with a large ΔR.
- This allows a quick determination as to whether the sample is essentially "as advertised."
- Experiment with releasing constraints to probe possible differences from the model structure.

Common Fitting Strategy #2: Bottom Up

This one works well for substances which do not correspond to known crystals.

- Start with nearest neighbor paths and relatively few constraints.
- As the structure becomes more clear, attempt to add reasonable constraints and perhaps more distant paths.

The Two Fitting Strategies

- Notice that both fitting strategies aim at the same target: a partially constrained multi-shell fit.

- The two strategies approach the desired fit from opposite directions, depending on the starting knowledge.
  - Top-down starts with lots of shells and adds free parameters by relaxing constraints
  - Bottom-up starts with most parameters free, and adds shells and constraints

The Point of Fitting

Remember, you are trying to answer one or more questions about your sample by EXAFS. The fit is not an end in itself! It is often helpful to divide questions into two types:

- **Qualitative**. For example, "Is my sample Fe_2O_3?" Qualitative questions are often addressed by trying to fit different starting structures.
- **Quantitative**. For example, "What is the iron-oxygen bond length in my sample?" Or "What fraction of my sample is oxidized?" These questions are generally answered by allowing parameters you are interested in to by varied during fitting.

When Am I Done Fitting?

- Fitting almost always follows a pattern of diminishing returns; as the fit improves, the continued improvement gets smaller and smaller per hour you put in.
- Remember the initial purpose of your investigation. While more refined fits are always nice for publication, a question as to whether a sample is composed of a given compound generally demands less work than determining a second-nearest-neighbor bond length to 0.01 angstrom accuracy.

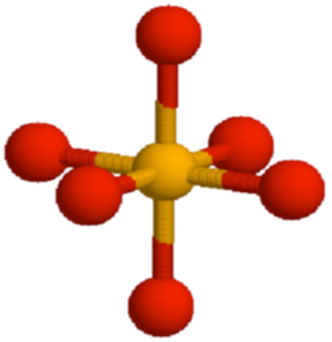
Are Single-Shell Fits the Most Reliable?

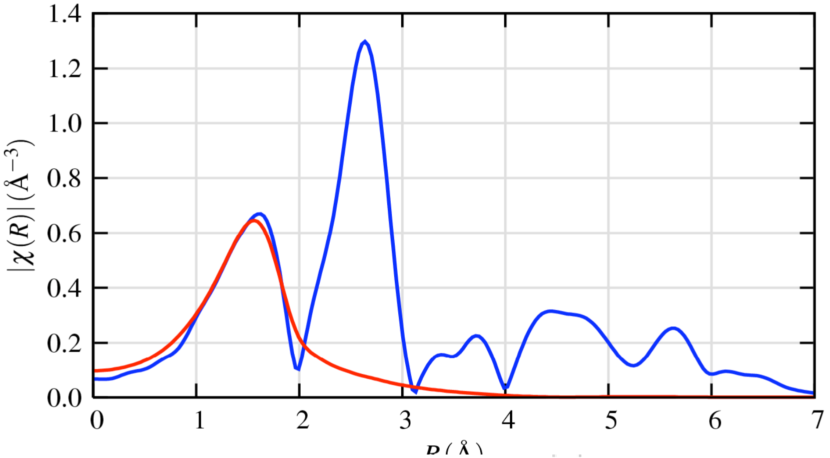
No!

If you know something about scatterers beyond the nearest-neighbor, fitting only the first shell means you're throwing information away. Even if you're only interested in questions about the first shell, you'll get a more reliable fit if you can include more shells.

Data Modeling: 1st Shell of FeO

- FeO has rock-salt structure
- Calculate f(k) and $\delta(k)$ using FEFF based on a a guess of structure, with Fe-O distance R=2.14 Å in a regular octahedral coordination
- Use the calculated functions to refine values of R, N, σ^2, and E_0 to match experiment



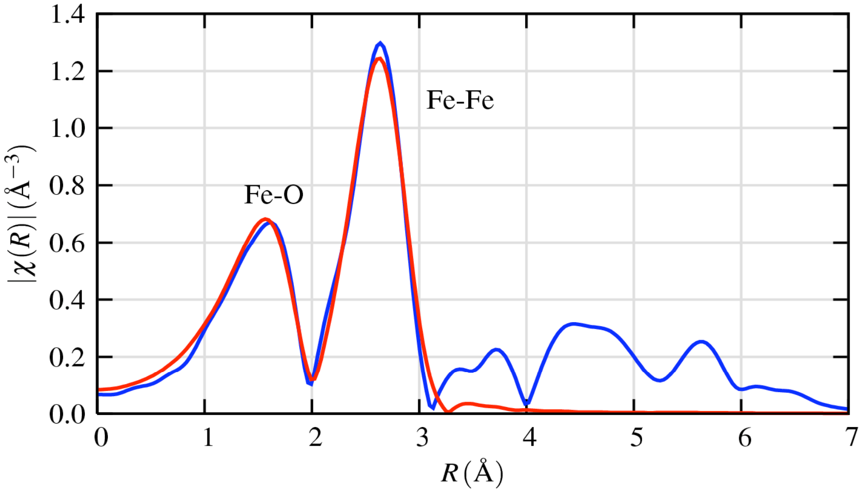


$|\chi(\mathrm{R})|$ for FeO (blue), and a 1st shell fit (red).

Fit results:

N = 5.8 ± 1.8
R = 2.10 ± 0.02Å
$\Delta\mathrm{E}_0$ = -3.1 ± 2.5 eV
σ^2 = 0.015 ± 0.005 Å^2.

Data Modeling: 2nd Shell of FeO



- Results are consistent with the know values for crystalline FeO:
    - 6 O at 2.13 A
    - 12 Fe at 3.02 A

Fit results (uncertainties in parentheses):

Shell	N	R (Å)	σ^2 (Å^2)	ΔE_0 (eV)
Fe-O	6.0(1.0)	2.10(.02)	0.015(.003)	-2.1(0.8)
Fe-Fe	11.7(1.3)	3.05(.02)	0.014(.002)	-2.1(0.8)

data=blue fit=red

Fancy Fitting

Here are some additional things that can be done when fitting with analysis software such as IFEFFIT:

- **Restraints**. This forces a parameters to stay *near* a value while still allowing it to vary a bit. A special case is a "penalty" restraint, which forces a parameter to stay within a specified range.
- **Multiple data set fits**. Data from more than one sample,more than one edge of a given sample, or a sample under more than one set of conditions can be refined simultaneously. Why is this useful? Because some parameters may be the same for all data sets; for example, in a temperature series, the chemical composition might remain the same.
- **Multiple phases**. An element in a sample might be in more than one environment (for instance, metal and oxide). A FEFF calculation can be run for each phase and combined to make a fit. Doping and compounds with nonequivalent crystallographic sites are also handled in this way.

Do I Have a Good Fit?

Criterion 1: Statistical Quality

- Statistical quality of fit is measured by "reduced chi-square," χ_ν^2. Oddly, this is *not* related to $[\chi(k)]^2$…we're just out of symbols!

$$\chi_\nu^2 = \frac{1}{\text{degrees of freedom}} \sum_i \left(\frac{\text{difference between data and fit}_i}{\varepsilon_i} \right)^2$$

- ε_i is the uncertainty in measurement *i*. This quantity is difficult to estimate. Ifeffit (and thus SixPack and Artemis), for instance, takes a shot at it by looking at noise high in the Fourier transform, but it's not necessarily a good estimate within the fitting range.

- If the difference between fit and measurement is entirely attributable to measurement uncertainty, then the reduced chi-square should be around 1.

- Unfortunately, because ε_i is difficult to estimate, the actual value of the reduced chi square is not very useful for judging the quality of the fit.

- It is very helpful, however, for comparing two fits on the same data set! Corwin will talk more about this in a few minutes!

Do I Have a Good Fit?

Criterion 2: Closeness of Fit

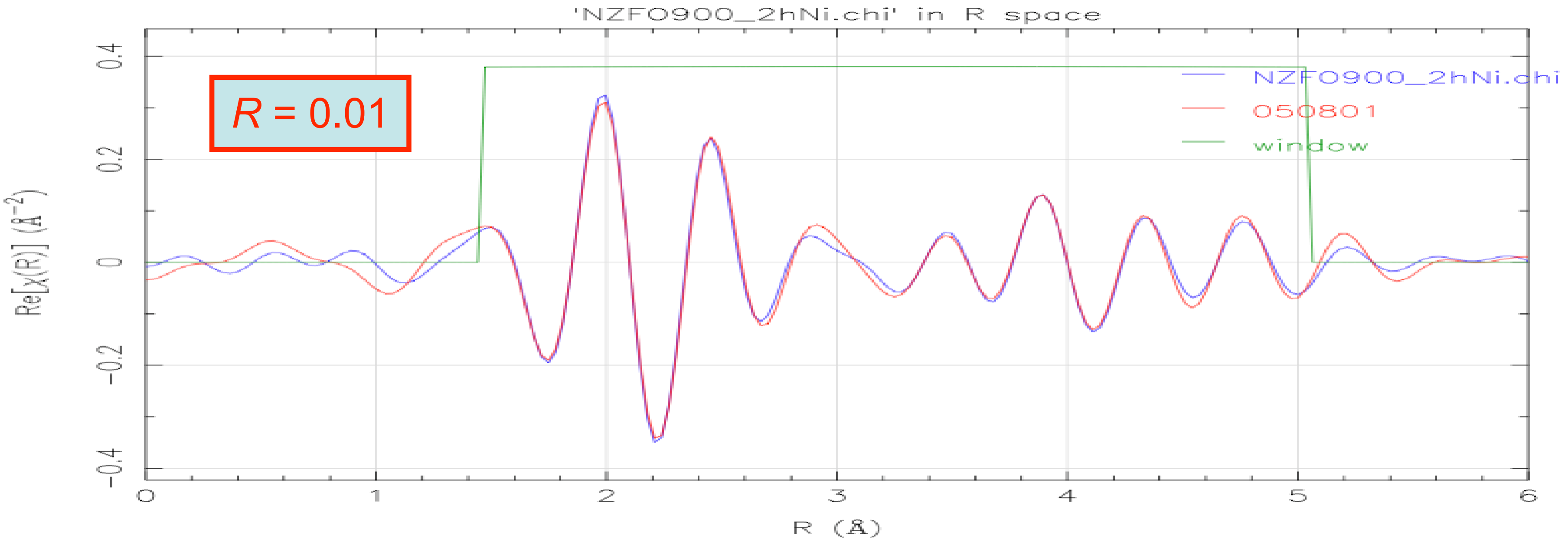
- This is not the same thing as a statistically good fit. If you have *really* good data, for instance, it can pick up very subtle variations between your constrained model and the sample. That can make your fit statistically poor. And yet the constrained model is "pretty good," even though the data would support better. To reflect this notion of closeness of fit, we use the EXAFS *R*-factor:

$$\mathscr{R} = \frac{\sum_i \left(\text{difference between data and fit}_i\right)^2}{\sum_i \left(\text{data}\right)^2}$$

- My off-the-record guidelines for interpreting the *R*-factor for data of moderate quality:

Range of *R*	Interpretation
< 0.02	Good enough
0.02-0.05	Either model has some details wrong, or data is low quality. Nevertheless, consistent with a broadly correct model.
0.05-0.10	Serious flaws in model or very low quality data.
> 0.10	Model may be fundamentally incorrect.

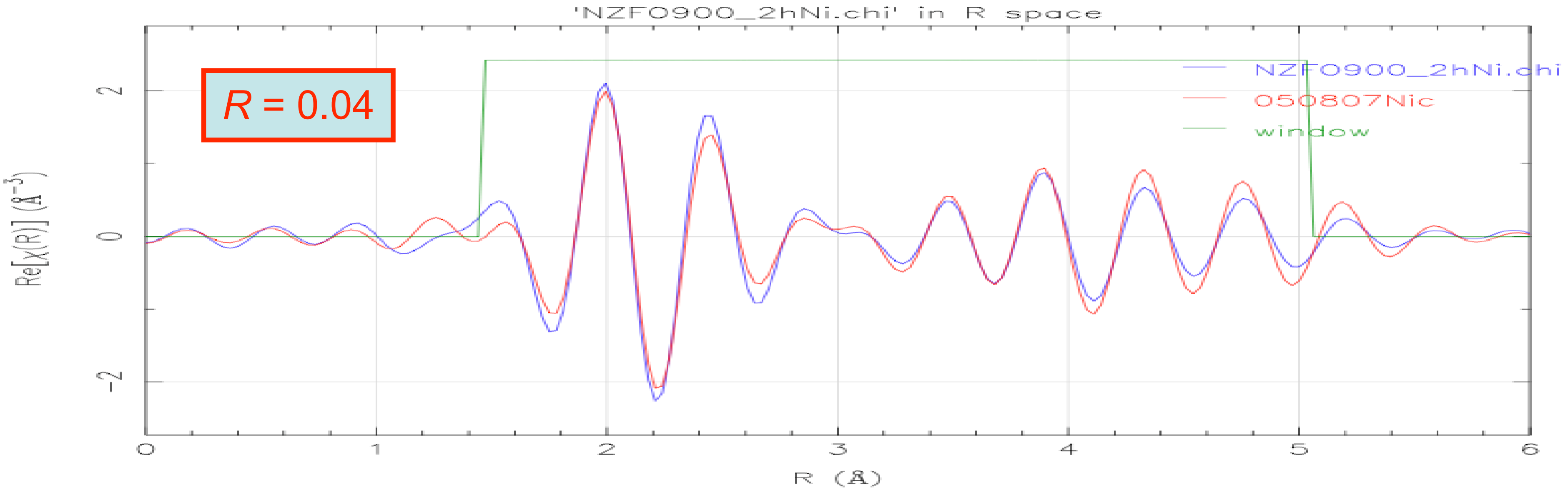
Do I Have a Good Fit?



Range of *R*	Interpretation
< 0.02	Good enough
0.02-0.05	Either model has some details wrong, or data is low quality. Nevertheless, consistent with a broadly correct model.
0.05-0.10	Serious flaws in model or very low quality data.
> 0.10	Model may be fundamentally incorrect.

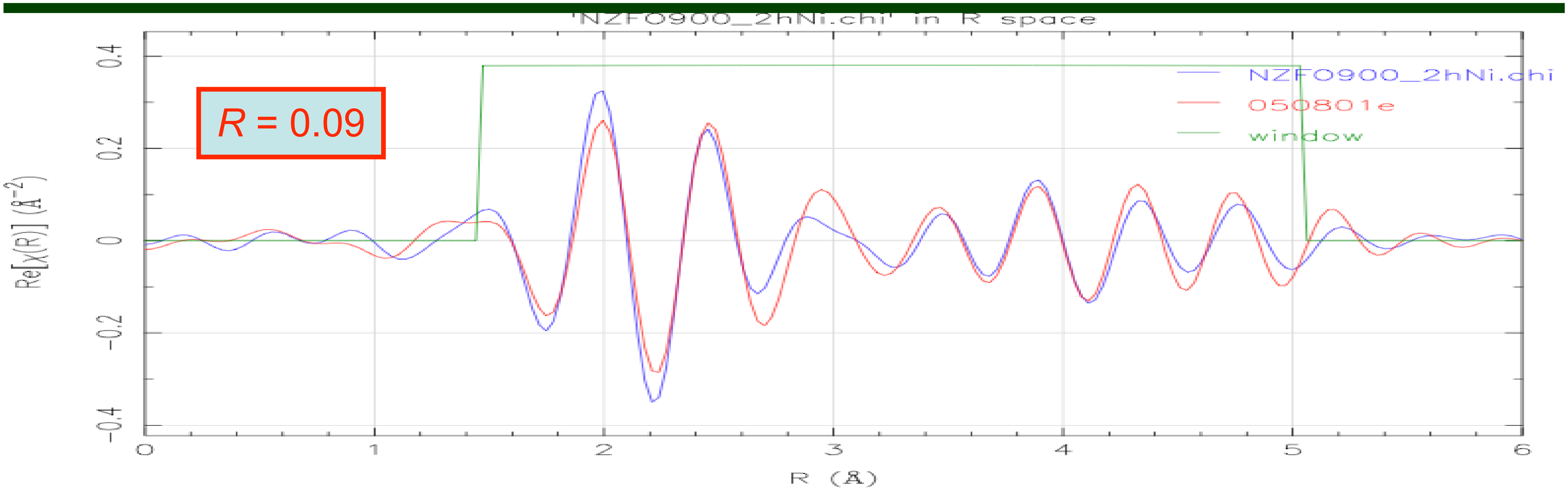
SARAH • LAWRENCE • COLLEGE

Do I Have a Good Fit?



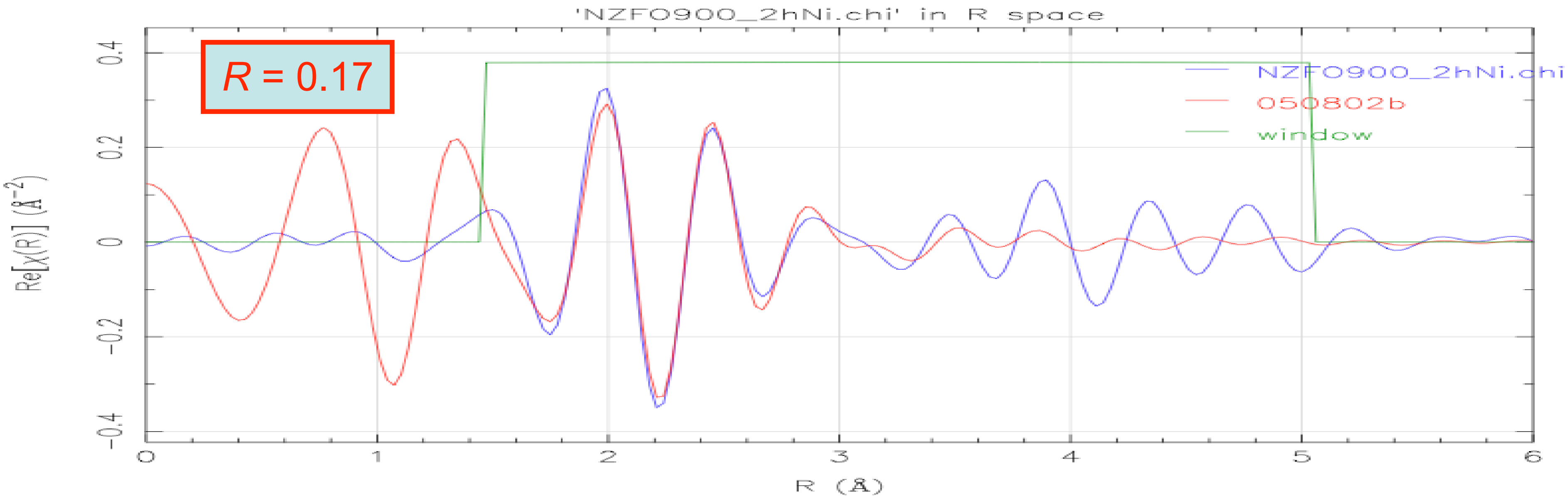
Range of *R*	Interpretation
< 0.02	Good enough
0.02-0.05	Either model has some details wrong, or data is low quality. Nevertheless, consistent with a broadly correct model.
0.05-0.10	Serious flaws in model or very low quality data.
> 0.10	Model may be fundamentally incorrect.

Do I Have a Good Fit?



Range of *R*	Interpretation
< 0.02	Good enough
0.02-0.05	Either model has some details wrong, or data is low quality. Nevertheless, consistent with a broadly correct model.
0.05-0.10	Serious flaws in model or very low quality data.
> 0.10	Model may be fundamentally incorrect.

Do I Have a Good Fit?



Range of *R*	Interpretation
< 0.02	Good enough
0.02-0.05	Either model has some details wrong, or data is low quality. Nevertheless, consistent with a broadly correct model.
0.05-0.10	Serious flaws in model or very low quality data.
> 0.10	Model may be fundamentally incorrect.

Do I Have a Good Fit?

Criterion 3: Do the Results Make Sense?

- The danger of "false fits" means that IFEFFIT may generate a close match of fit to data that is utter nonsense. This can generally be detected by examining the values of parameters allowed to vary in the fitting process. False fits are often marked by one or more "unreasonable" values for parameters.

- "Isn't this a circular argument? Aren't you just rejecting fits that don't give you the results you want?" No, although care must be taken to guard against cherry-picking fits. The key is to reject fits that are physically highly implausible, not ones that support an alternative but plausible hypothesis. Typical examples of physically implausible parameters:
  - S_o^2 less than 0.50 or more than 1.20 (*should* be 0.80 to 1.0)
  - E_o not on or very near rising portion of the edge (more in a later talk in this workshop)
  - Unheard of bond lengths for the species involved
  - Negative values of σ^2
  - Excessive coordination numbers (e.g. 20)
  - Site occupancies negative or greater than one
  - Any parameter greatly at odds with a "known" value (probably determined by another experimental technique)

- "OK, so those fits are false. How do I know a false fit wouldn't happen to generate reasonable parameters?" It might, but it's less likely. It is a good idea, though, to "stress" your fit, which leads us to the next criterion…

Do I Have a Good Fit?

Criterion 4: Is the Model Defensible?

- For a given set of data, it might be possible to obtain a statistically better fit by some strange choice of constraints: perhaps, for instance, the third, fourth, and sixth nearest-neighbor oxygens are constraned to have the same σ^2, but the fifth nearest-neighbor oxygen has a σ^2 that is fit separately. If there is no *a priori* reason to expect the fifth nearest-neighbor to behave differently from the others, then choosing this fit is not statistically valid. After all, there might be a hundred or more oddball variations on constrant schemes that could be tried. By *luck*, several are likely to show statistically significant improvement at the 0.05 confidence level!

- Models and constraint schemes, therefore, should make enough physical sense that it is possible to imagine someone with knowledge of the system picking them *a priori*. It is fine to choose between two "sensible" models on the basis of statistics, but that doesn't mean a model that appears arbitrary or capricious should be selected in that way.

- All else being equal, it is good to choose a model which is either:
  - **Simple**: a model with very few free parameters, and the parameters that are varied incorporate simplifying assumptions. For instance, there might be a single parameter representing an overall lattice expansion. For many materials, such a model is likely to be a simplification of the actual behavior of the material as it changes temperature, but it *might* work as a first approximation.
  - **General**: a model which could fit a wide variety of candidate structures. Such a model will have few constraints, but many free parameters.
  - **Physical**: a model constrained by a physical understanding of the material. This understanding may be derived from other characterization techniques or from theoretical tools such as DFT.

- When trying to decide whether a model is "defensible," imagine how you will describe it in a scientific paper. If you find yourself tied in knots trying to justify why you chose the model you did, it is probably best to see if you can find a model which is either simpler, more general, or more physical.

Do I Have a Good Fit?

Criterion 5: Stability

- A good fit should be stable, meaning that the key results should not be sensitive to details of the fitting strategy.

- Always test the stability of your final fit by trying to "stress" it. That may be done in the following ways:
  - Change the *k*-range of data being used
  - Change the *R*-range of the Fourier transform being fit
  - Change the *k*-weight
  - Remove a constraint

- If the fit is good, modest changes in the above should not change the answer(s) to the scientific questions you are trying to address.

Do I Have a Good Fit?

Criterion 5: Stability (continued)

- The following often occur during stability checks, but are *not* cause to reject the initial fit:
  - Some of the stability checks yield very large uncertainties for some parameters
  - Parameters which are not of interest drift outside of their original uncertainties
  - When the details of the fitting strategy are changed sufficiently, the fit may get "lost" and replaced by a patently false fit
  - The *R*-factor degrades
- On the other hand, any of the following suggest more work needs to be done:
  - Parameters which answer your scientific questions drift outside of their original uncertainties without other indications that the fit has become "bad"
  - The fit flips to another set of values which also appear reasonable and with a comparable R-factor

Do I Have a Good Fit?

Criterion 6: Precision

A fit that tells you a bond length is 2.24 ± 1.45 angstroms doesn't tell you much about that bond length.

If it's something you're interested in, you'd like a fit with better precision.

Do I Have a Good Fit?

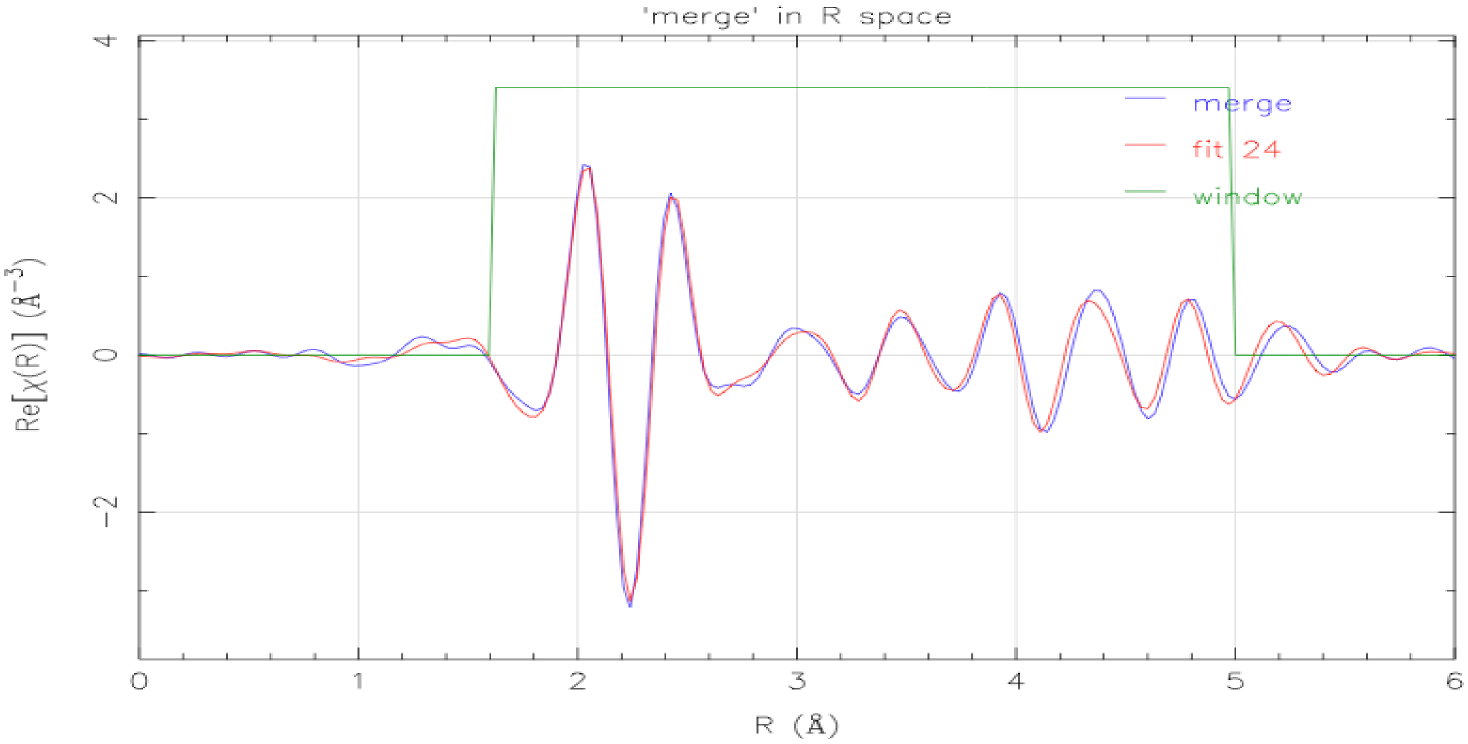
Criterion 7: More data is better

All else being equal, a fit which uses more of your data (a wider *k*-range and a wider *R*-range) is more convincing.

Do I Have a Good Fit?

Criterion 8: Agreement outside the fitted range

If the fit agrees with the data fairly well outside the range being fitted, that's a good sign.



Do I Have a Good Fit? Summary.

Criterion 1: Statistical quality
Criterion 2: Closeness of fit
Criterion 3: Sensible results
Criterion 4: Defensible model
Criterion 5: Stability
Criterion 6: Precision
Criterion 7: More data is better
Criterion 8: Agreement outside the fitted range