

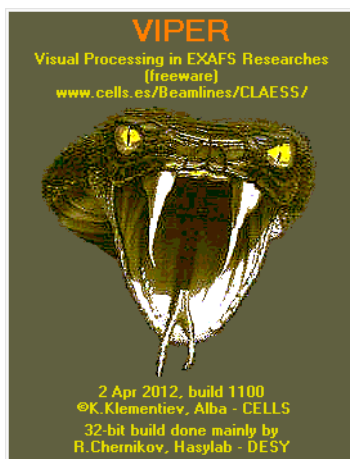
VIPER

(Visual Processing in EXAFS Researches) for Windows

Users Manual and Tutorial
with comments on analysis methods in EXAFS

version of manual 2.20
version of program 11.00

2 April 2012



K. Klementiev

CELLS-ALBA, Carretera BP 1413, km. 3, E-08290 Cerdanyola del Vallès, Barcelona SPAIN

www.cells.es/Beamlines/CLAESS

Contents

1 Introduction.....	4
1.1 What is VIPER?.....	4
1.2 What makes VIPER special?.....	4
1.3 System requirements.....	4
1.4 About this manual.....	4
2 Opening data files.....	4
3 Short tips to the program interface.....	5
4 Working with experimental signals.....	7
4.1 Energy calibration.....	8
4.1.1 Constant angle shift.....	8
4.1.2 Constant lattice shift.....	9
4.1.3 Constant energy shift.....	10
4.1.4 Which shift to use?.....	11
4.2 Absolute absorption coefficient.....	11
4.3 Marking glitchy regions.....	12
4.4 Getting μ and χ	12
4.5 How to go back to the Currents window?.....	12
5 Working with μ and χ.....	13
5.1 Deglitching.....	13
5.1.1 Step (jump) glitches.....	13
5.1.2 Sharp glitches.....	14
5.1.3 Switching off glitch coloring.....	15
5.2 Pre-edge background.....	15
5.2.1 Corrections of pre-edge background.....	16
5.2.2 How the far end of μ should behave?.....	16
5.2.3 Which polynomial to choose for the pre-edge of transmission spectra?.....	16
5.2.4 Show μ normalized.....	17
5.3 Setting E0.....	17
5.4 Setting k mesh.....	17
5.4.1 How to choose kmin?.....	17
5.4.2 How to choose kmax?.....	18
5.4.3 How to choose dk?.....	18
5.5 Construction of post-edge background μ_0	18
5.5.1 μ_0 as a spline drawn through varied knots.....	18
5.5.2 μ_0 as a smoothing spline.....	20
5.5.3 μ_0 as a Bayesian smoothing curve.....	22
5.6 Normalization of χ	23
5.7 Corrections to μ and χ	23
5.7.1 Self-absorption correction.....	23
5.7.2 Description of self-absorption correction.....	24
5.7.3 Realization in VIPER.....	24
5.7.4 Example of self-absorption correction.....	26
5.8 k-weighting of χ	27
5.8.1 What k-weighting to use?.....	27
5.9 Subtracting EXAFS due to another closely situated absorption edge.....	28
5.10 Deconvolution of life-time and experimental broadening.....	29
5.10.1 How to select the regularizer?.....	30
5.11 Combining several spectra together.....	30
6 Fourier analysis.....	31
6.1 Forward Fourier transform (FT).....	31
6.2 Back Fourier transform (BFT).....	31
6.3 Notes on the program interface.....	32
6.4 Selection of kmin and kmax.....	32
6.5 Which windowing function to use?.....	33
6.6 FT pre-correction.....	34
6.7 Extraction of amplitudes and phases.....	34
7 Experimental errors in EXAFS curve.....	35
7.1 Using the high-r portion of $\chi(r)$	35

7.2 Using FT filtering.....	36
7.3 Using μ_0 obtained by Bayesian smoothing.....	36
7.4 Using Bayesian deconvolution.....	37
7.5 Using μ_0 drawn through varied knots.....	37
7.6 Using standard deviation of multiple data.....	37
7.7 Calculated from a user formula.....	38
7.8 Comparison of different estimations of experimental errors.....	38
8 Fitting EXAFS.....	39
8.1 Fitting by ordinary EXAFS formula.....	39
8.1.1 How about multiple scattering (MS) fitting? Can VIPER do it?.....	40
8.1.2 Why the energy shift is not global but is different for different shells?.....	40
8.1.3 How to load a model and save the fitting results?.....	40
8.1.4 How to create a new model?.....	42
8.2 Fitting by a user-expanded EXAFS formula (cumulant expansion).....	42
8.3 Fitting using radial distribution function specified by user-defined formula.....	42
8.4 Fitting using oscillatory potential $U(r)$ of the absorber-scatterer pair.....	43
8.5 Multi-edge fitting.....	44
8.6 Fitting in k- and r-space.....	45
8.7 Is it possible to fit S02?.....	45
8.8 Details of the fitting algorithm.....	46
9 Statistical evaluations in fitting.....	46
9.1 χ^2 statistics.....	46
9.2 Posterior distribution.....	47
9.3 Simplest cases: partial correlations.....	47
9.4 General case: total correlations and a priori information.....	48
9.5 Most probable a priori weight.....	49
9.6 What if the experimental errors ϵ_i were determined incorrectly?.....	50
9.7 What if the experimental errors ϵ_i are unknown?.....	50
9.8 'Statistical evaluations' dialog.....	50
9.9 Statistical tests.....	51
9.9.1 χ^2 -test.....	51
9.9.2 F-test.....	51
9.10 Appendix.....	52
10 Exporting data and saving project file.....	53
References.....	53

1 Introduction

1.1 What is VIPER?

VIPER is a program for data analysis of EXAFS spectra. It includes:

- pre-processing of raw data with energy calibration, deglitching, deconvolution, advanced self-absorption correction etc.,
- various procedures for extraction of the EXAFS part,
- merging of spectra,
- Fourier-analysis,
- fitting procedures for the first few coordination shells, including multi-edge fitting,
- advanced error analysis.

VIPER does not include calculation of scattering amplitudes and phases. I use FEFF for this or, for well isolated in r -space first coordination shells, I extract the amplitudes and phases from reference spectra in VIPER. VIPER also does not produce publication quality graphs. It only exports column files to be loaded by Matplotlib, QtiPlot, Origin etc.

1.2 What makes VIPER special?

Any time, all curves and their changes under processing are visual. The visualization is not only a matter of convenience; it serves for the ultimate quality check of experimental data and processing steps by the program user.

VIPER is also useful for quick quality check during your beam time at synchrotrons. A simple drag-and-drop action reveals in a second the spectrum quality and reproducibility in E -, k - and r -space.

1.3 System requirements

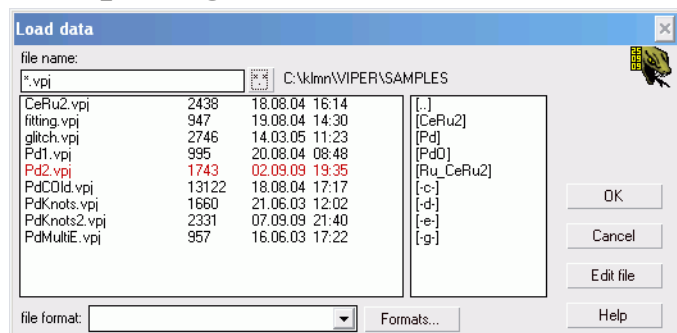
VIPER runs on all 32- and 64-bit Windows systems. It can run under Linux with Wine. The minimum screen resolution is recommended as 1024×768.

Originally, VIPER was a 16-bit program that could not run on 64-bit Windows. I thank Roman Chernikov (HASYLAB at DESY) for making the 32-bit build.

1.4 About this manual

It is essential to download and unpack the archive with several important examples (see the front page). I have tried to explain all the aspects of the program that may be useful to its user in setting up his or her analysis of EXAFS spectra. Some of the aspects are not quite standard or, being standard, are questionable. These are considered with higher attention.

2 Opening data files



You can select multiple files using Ctrl or Shift buttons or by mouse dragging. The name of the last opened file is colored by red. The design of the Load data dialog is old fashioned; the files are always sorted by name whereas frequently time sorting is more convenient. Therefore I recommend drag-and-drop technique combined with your favorite file commander or Explorer. This way is very useful at a beamtime, when you

quickly add a newly measured file to the already opened ones by simple drag-and-drop from your time-sorted directory. I use the Load data dialog mostly to set up new data formats and, sometimes, to manually select the file format. The latter is needed when the same file has transmission and fluorescence signals and one wants to load both. In this case one needs two formats described and, of course, only one of the two will be recognized automatically.

Important: you can load multiple files and do drag-and-drop only provided your file format is recognized automatically (i.e. when you see the format name updated correctly in the 'Load data' dialog after you have clicked a file name).

The number of the loaded spectra is restricted by your RAM. I used to work with a hundred of spectra. However, redrawing becomes slow. In this case you can 'Hide all other spectra' in the 'Spectra' menu.

Specify the file header. Give one or two sub-strings contained in the header for automatic recognition. If your file is recognized incorrectly, try to find other unique sub-strings or use button 'Up' to place your format earlier in the recognition queue.

In the description of the data columns one can use (almost) any function of variables Col1 ... Col52. For instance, one can load several fluorescence signals i1 as, say, Col5+Col6+... or, better, one can load these signals as separate spectra for better visual quality checking.

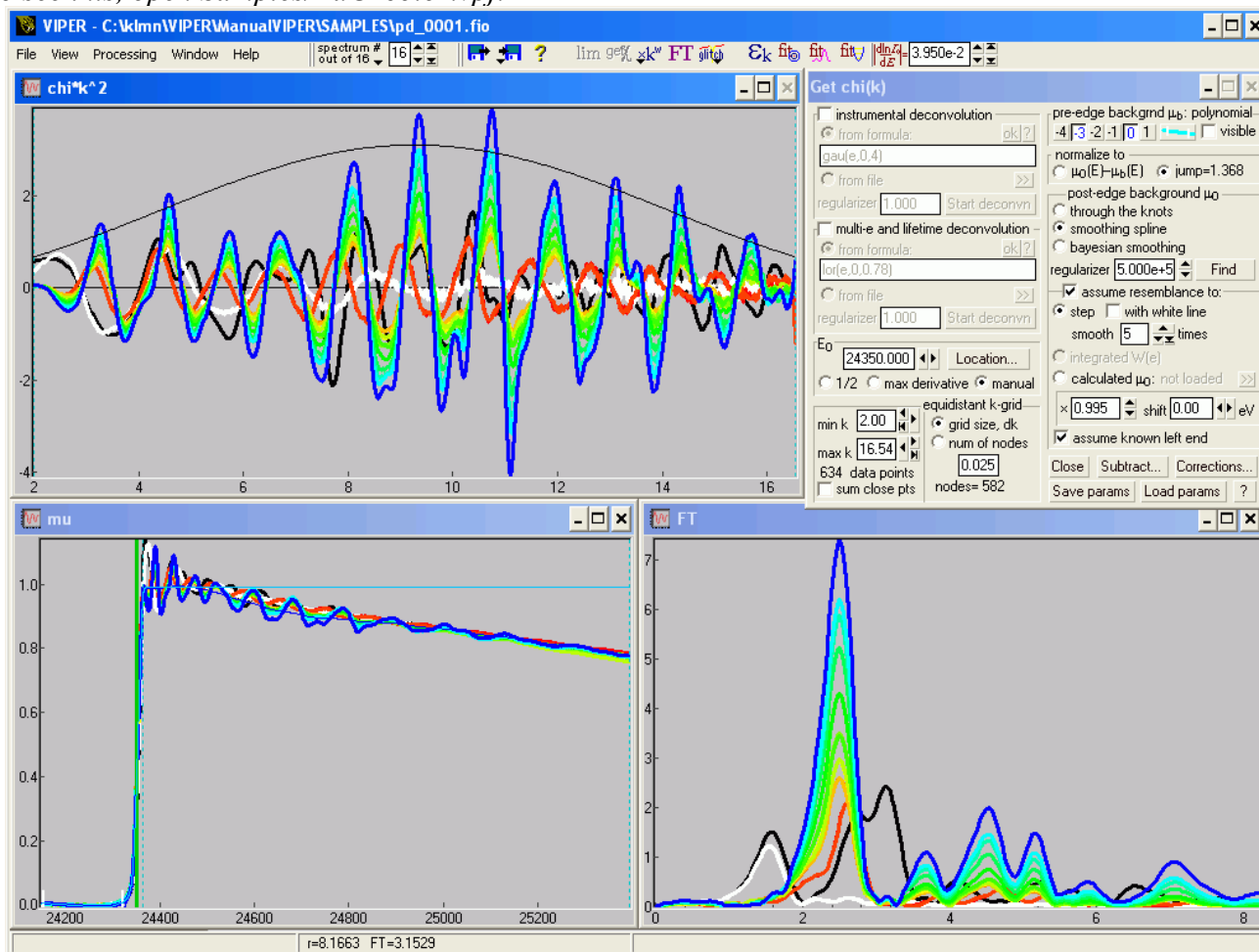
The internal energy unit is eV. Therefore if your energy unit is different, you should do a transform, like Col4*1e6. For keV unit there is a dedicated option.

The 'reference curve' is only needed for energy calibration and can be left empty. Usually, this is the absorption coefficient of a reference foil placed between the 2nd and the 3rd ionization chambers. Correspondingly, it is given by $\ln(i1/i2)$.

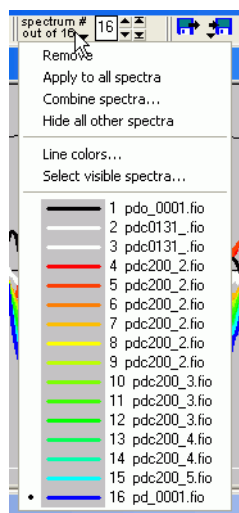
The format descriptions are saved in a text file formats.ini. If you want to transfer it to another computer, just copy it to the VIPER directory. You can manually merge various formats.ini files using a common text editor; re-number then the strings properly.

3 Short tips to the program interface

to see this, open Samples/PdC color.vpj:



In every window use the pop-up menu to access the most frequent commands, e.g. to restore the default zooming. Use the legend entries to access the line properties.

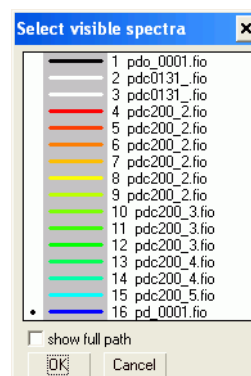


The 'Spectra' menu at the very top of the main window can be activated by flying the mouse cursor over it. Use the 'Spectra' menu to switch between the loaded spectra, change their sequence, add or remove spectra, access the line properties etc.

Note that all properties in many dialogs refer to the *current* spectrum which you can select here. Use this menu also to apply the selected options to the other spectra.

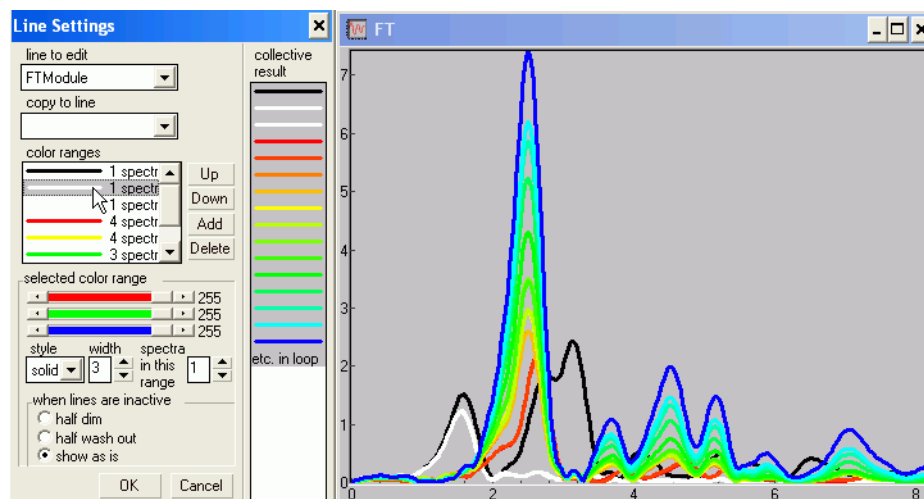
You can select spectra for visualizing/hiding.

Note that all the colors here refer to the current (active) window.



There are three ways for visual identification of spectra:

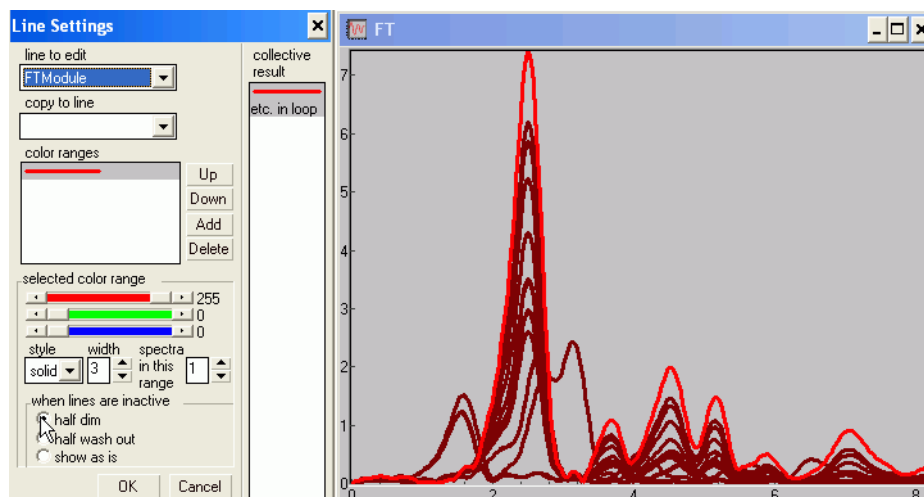
1) by colors (*load example Samples/PdC_color.vpj*):



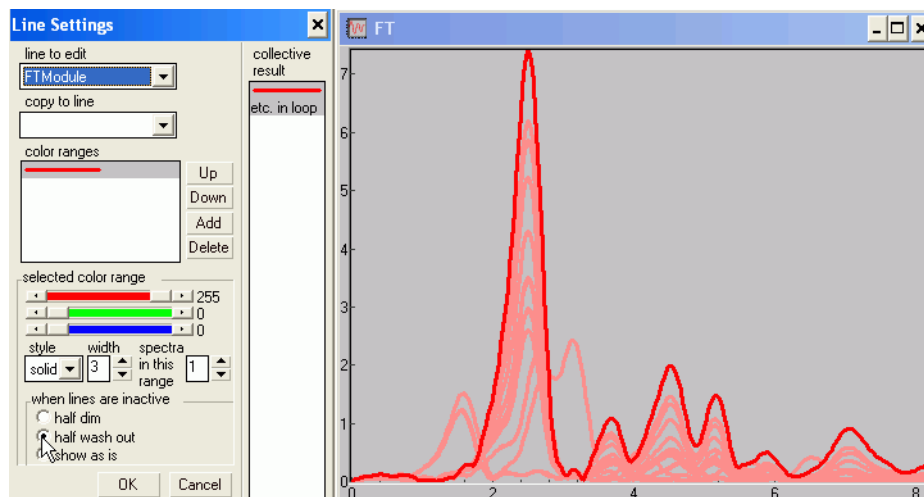
The line properties can be set collectively by specifying the color ranges. Within each range, the specified number of spectra are evenly spaced in the RGB color space towards the next color range. If the range is single, all the lines of a given kind will be equally colored for each spectrum. A single range is recommended for the lines which are only visible for the current spectrum such as μ_0 or μ_b lines.

The 'dots' line style puts dots not equidistantly but at the data points.

2) by dimming the inactive spectra (*load example Samples/PdC_dim.vpj*):

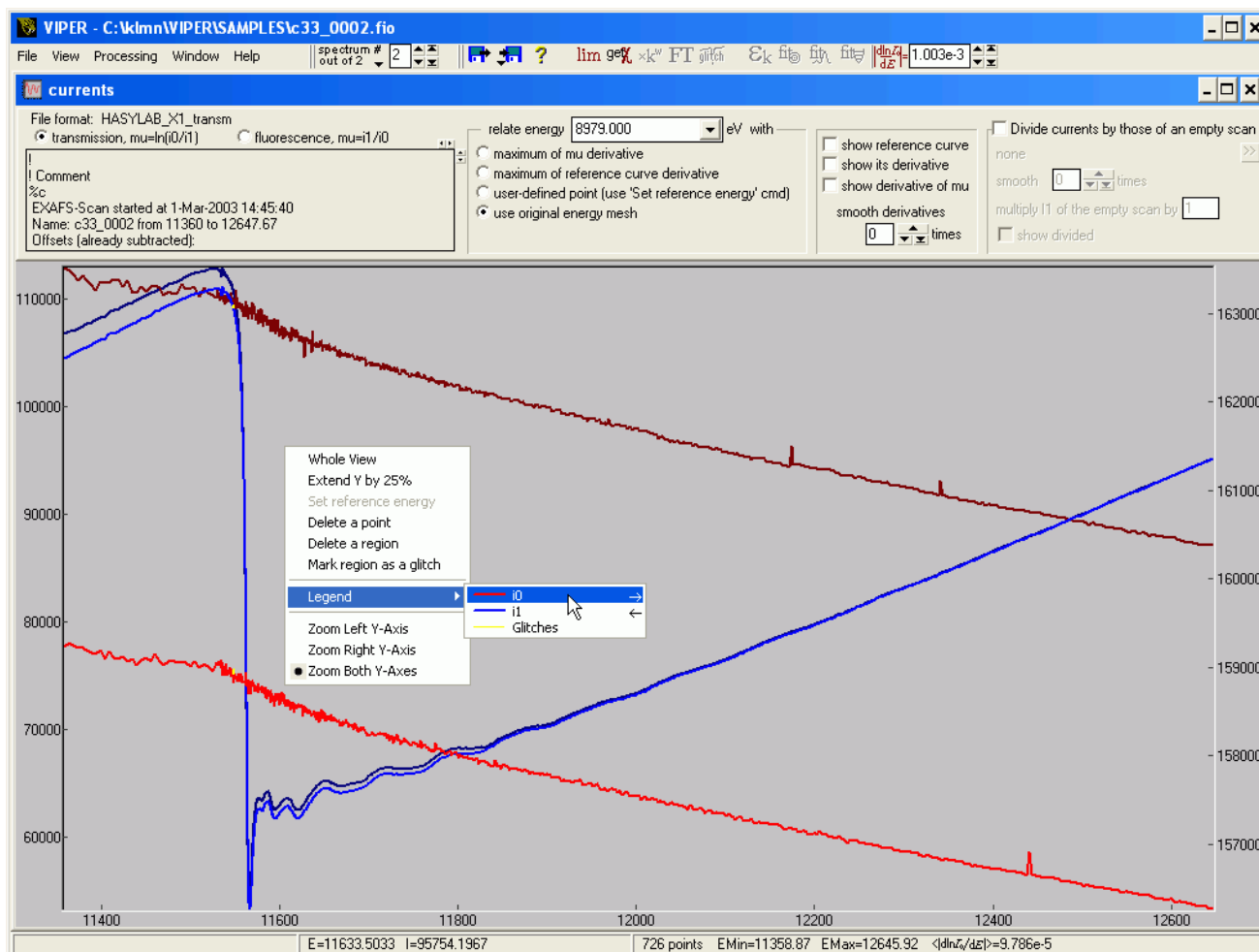


3) by 'washing out' the inactive spectra (load example Samples/PdC_wash.vpj):



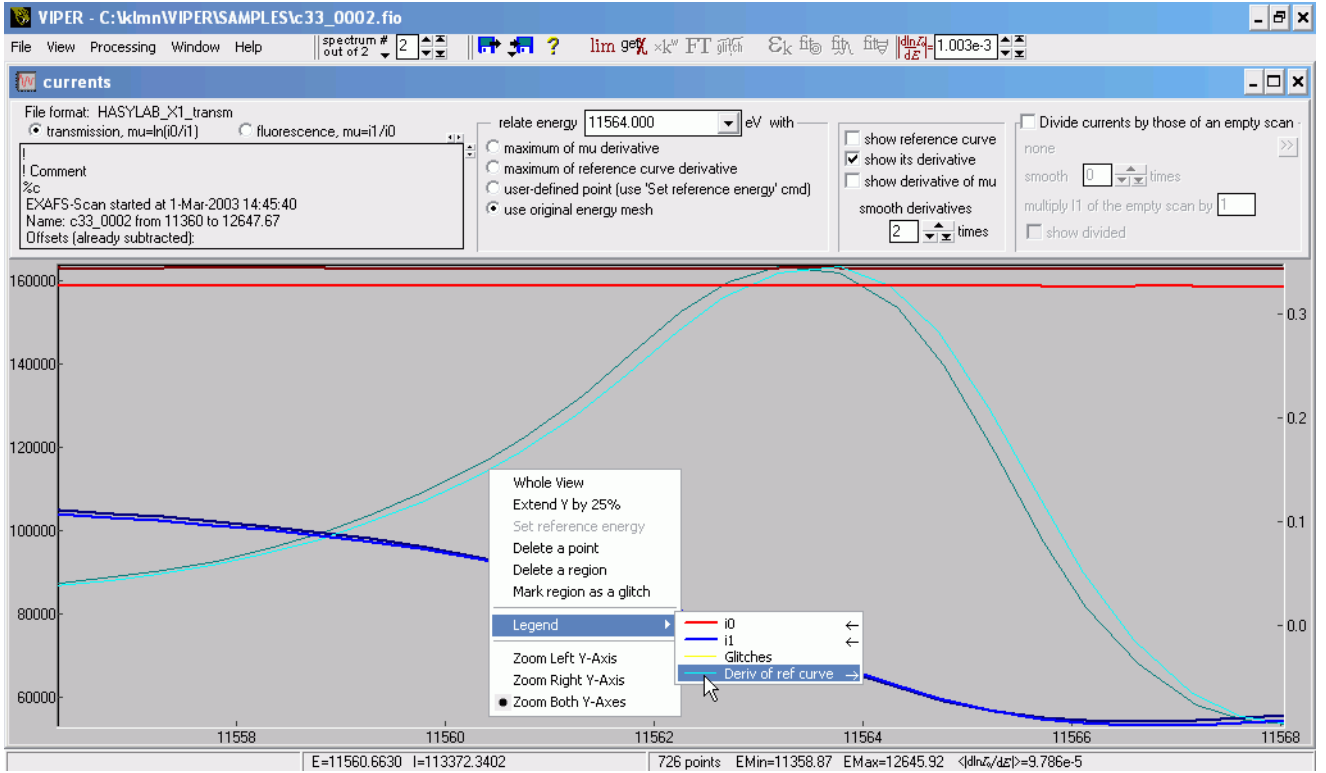
4 Working with experimental signals

to see this, open Samples/glitch.vpj and close 'mu' window by pressing the cross in its top right corner:



Select now transmission or fluorescence mode. If you did it false, you can redo this at a later stage, even with active χ , FT and BFT windows.

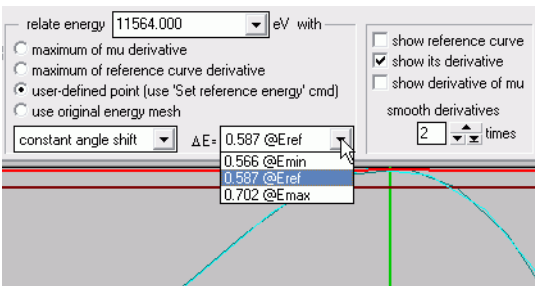
4.1 Energy calibration



Visualize the derivative of the reference curve. Now all the currents refer to the left Y-axis and the derivative refers to the right one. Zoom in the derivative peak, as seen on the screenshot.

Tip: derivative of μ is used for foils only. When the sample is a foil, usually you do not put the foil also at the reference position. In this case you use the derivative of μ for the sake of energy calibration but not the derivative of the reference spectrum.

Now select the reference energy. If your energy mesh around the absorption edge is fine, just use 'maximum of reference curve derivative'. In the example shown the mesh was rough, 0.5 eV. Therefore such a calibration will not improve the energy reproducibility between the two spectra (try it!). In this case much better calibration is given by manual positioning of the reference energy. For this, use 'user-defined point' and the pop-up menu command 'Set reference energy' (and put it somewhere close to the peak maximum) until you merge all the reference curves. You can see the resulting energy shift in the drop-down list:



[A note for future development] An option should be offered for the parabolic interpolation around the maximum of reference curve derivative. A similar interpolation is shown in Section 6.3.

4.1.1 Constant angle shift

The most frequent reason for energy shifts seems to be a Bragg angle shift. This can be due to (i) backlashes in the gearbox or (ii) elastic angular shift between the Bragg axis and its encoder or (iii) wrong energy calibration caused by a wrongly calculated angular offset (yes, this happens frequently).

How it works (here δ is a shift, ∂ is its error and Δ is a scan range):

$$\left(\frac{ch}{2dE_{\text{true}}} - \frac{ch}{2dE_{\text{exp}}} \right) \frac{1}{\cos\theta} = \delta\theta = \left(\frac{ch}{2dE_{\text{true}}^{\text{ref}}} - \frac{ch}{2dE_{\text{exp}}^{\text{ref}}} \right) \frac{1}{\cos\theta^{\text{ref}}},$$

$$\frac{1}{E_{\text{true}}} = \frac{1}{E_{\text{exp}}} + \left(\frac{1}{E_{\text{true}}^{\text{ref}}} - \frac{1}{E_{\text{exp}}^{\text{ref}}} \right) \frac{\cos\theta}{\cos\theta^{\text{ref}}} \approx \frac{1}{E_{\text{exp}}} + \left(\frac{1}{E_{\text{true}}^{\text{ref}}} - \frac{1}{E_{\text{exp}}^{\text{ref}}} \right) (1 - \tan\theta^{\text{ref}} \cdot \Delta\theta).$$

Now we neglect

$$\left(\frac{1}{E_{\text{true}}^{\text{ref}}} - \frac{1}{E_{\text{exp}}^{\text{ref}}} \right) \tan \theta^{\text{ref}} \cdot \Delta \theta ,$$

which finally gives

$$\frac{1}{E_{\text{true}}} = \frac{1}{E_{\text{exp}}} + \frac{1}{E_{\text{true}}^{\text{ref}}} - \frac{1}{E_{\text{exp}}^{\text{ref}}} \quad \text{and the error} \quad \partial(\delta E) \approx \delta E^{\text{ref}} \tan \theta^{\text{ref}} \cdot \Delta \theta .$$

Let us take $\delta\theta=10^{-4}$, $\theta=13^\circ$, i.e. the energy shift $\delta E \sim 4\text{eV}$ @ Cu *K*-edge, and $\Delta\theta=1^\circ$ (i.e. the energy range $\Delta E \sim 1\text{ keV}$). The error in the energy correction is then $\partial(\delta E) \sim 4\text{ meV}$ to the end of the spectrum and is really negligible.

Let us see what error we get when we by mistake do a "constant *E*" correction to a constant angle shift. We modify the final expression slightly:

$$\delta E \equiv E_{\text{true}} - E_{\text{exp}} \approx \delta E^{\text{ref}} \left(\frac{E}{E^{\text{ref}}} \right)^2 \approx \delta E^{\text{ref}} \left(1 + \frac{2\Delta E}{E^{\text{ref}}} \right).$$

δE differs from the constant shift δE^{ref} by

$$\partial \delta E = \delta E^{\text{ref}} \frac{2\Delta E}{E^{\text{ref}}}$$

For the example above this gives $\partial \delta E \approx 0.8\text{ eV}$ to the end of the spectrum, or $\partial k = k(\delta E_{\text{ref}}/E_{\text{ref}}) \approx 0.007\text{ \AA}^{-1}$, or $\partial r \approx 0.001\text{ \AA}$ for the distance determination in the 1st coordination shell. In most EXAFS applications this is negligible and a constant energy shift calibration can be acceptably applied. This is especially true for XANES spectra, as the mis-calibration is proportionally smaller for short spectral range ΔE .

4.1.2 Constant lattice shift

Another reason for energy shifts is a wrong assumption on the lattice constant of the monochromator crystal because it was taken for a temperature different from the working one.

How it works:

$$\frac{\delta E}{E} = \frac{\delta d}{d} = \frac{\delta E^{\text{ref}}}{E^{\text{ref}}} \quad \text{and hence} \quad E_{\text{true}} = E_{\text{exp}} \frac{E^{\text{ref}}}{E_{\text{exp}}^{\text{ref}}} .$$

Let us see what error we get when we by mistake do a "constant *E*" correction to a constant lattice shift.

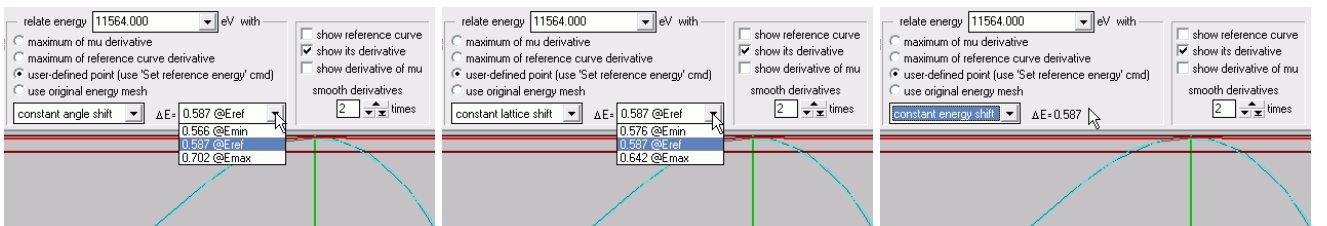
$$\delta E \equiv E_{\text{true}} - E_{\text{exp}} = \delta E^{\text{ref}} \frac{E}{E^{\text{ref}}} = \delta E^{\text{ref}} \left(1 + \frac{\Delta E}{E^{\text{ref}}} \right).$$

δE differs from the constant shift δE^{ref} by

$$\partial \delta E = \delta E^{\text{ref}} \frac{\Delta E}{E^{\text{ref}}}$$

This error is twice as smaller as compared to the constant angle shift, and for the example above this gives $\partial \delta E \sim 0.4\text{ eV}$ to the end of the spectrum.

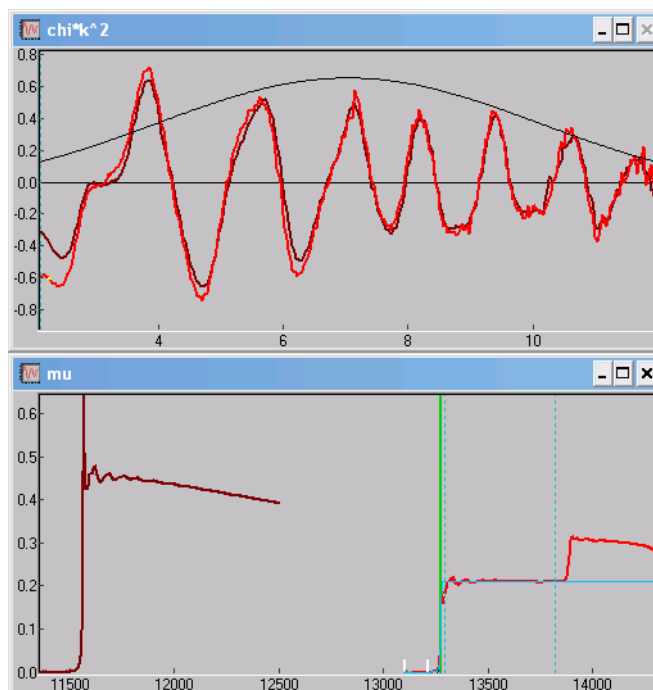
The following screenshots demonstrate the three energy shifts along the spectrum:



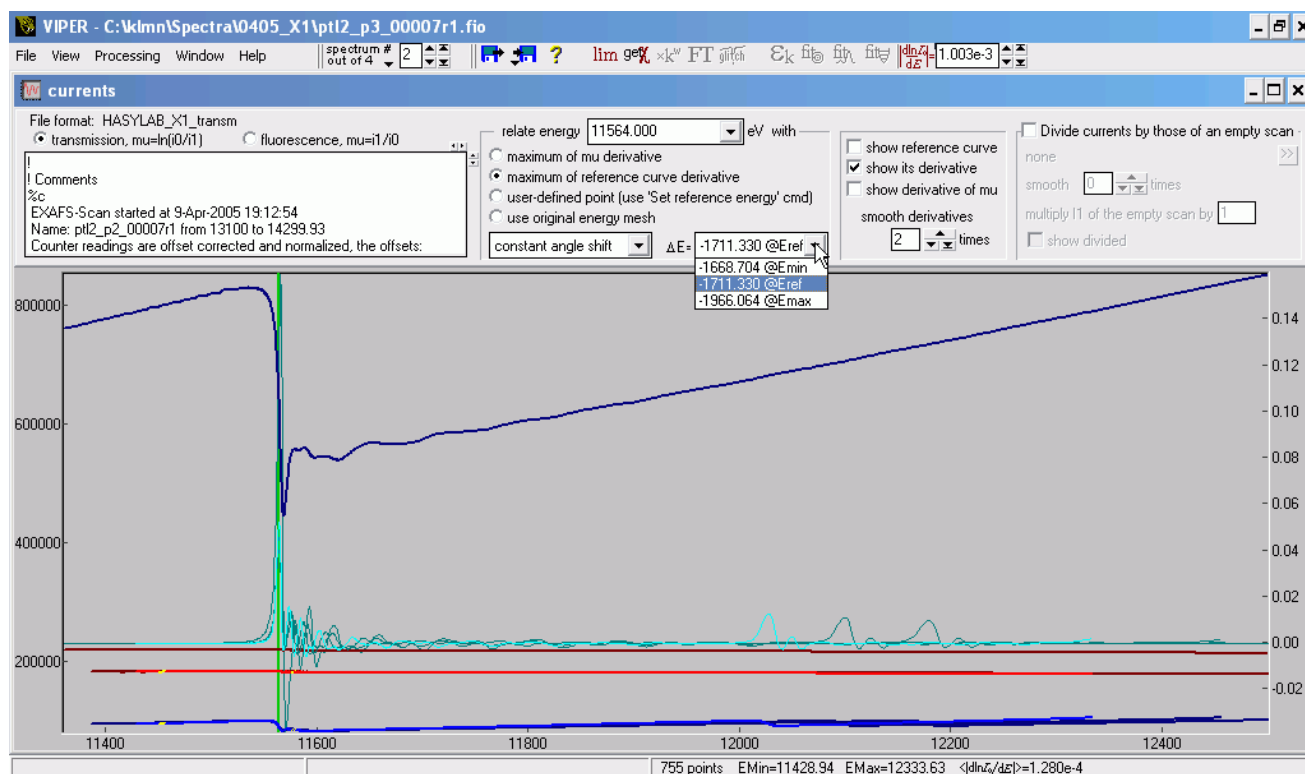
4.1.3 Constant energy shift

Such a shift is usually implemented in other EXAFS analysis programs. Here it is implemented as well. Its main application is for the alignment of two various absorption edges, as in the following example of the Pt L₃ and L₂ edges.

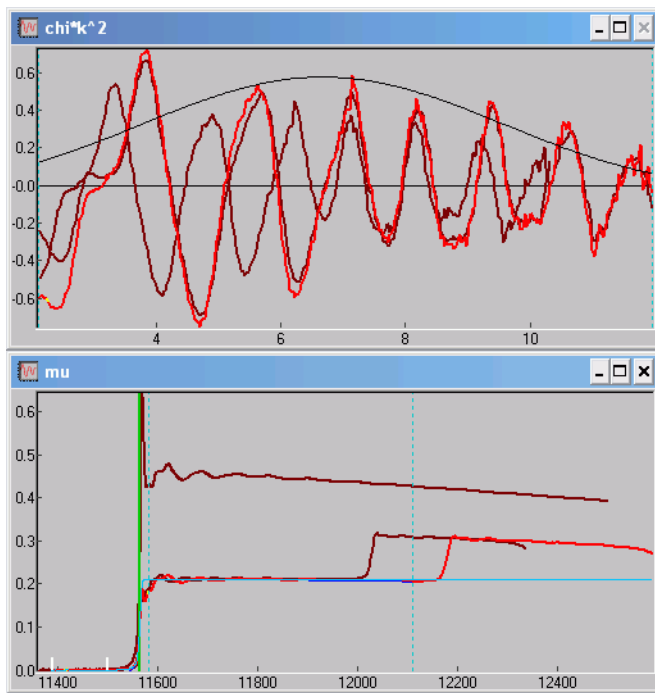
As is well known, EXAFS of L₃ and L₂ edges are very similar:



If we now want to compare XANES, we shift L₂ to L₃ or in the reverse way. This shift has nothing to do with wrong Bragg angle and should be of constant energy. What happens if we shift it with a constant angle shift? This would lead to a non-uniform shift along the spectrum (see the drop-down list):



and to false contraction or expansion of the shifted spectrum (here the two L₂ edges are shifted towards L₃(dark): one with constant energy shift (red), the other with constant angle shift (also dark); the latter one gives the false out-of-phase EXAFS):



4.1.4 Which shift to use?

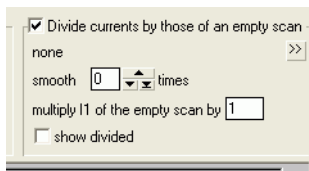
The answer should be explored for every XAS beamline. Ask the beamline scientists for possible reasons of energy drifts, they must know their instrument. As mentioned at the end of Section 4.1.1, constant energy shift works well in most cases. It does not work well when the reference edge is not the same as the one of the sample. This is a frequent case for L edges of rare earths when 3d metal foils are used for energy calibration.

4.2 Absolute absorption coefficient

The absolute absorption coefficient is useful for the determination of an unknown elemental concentration, see XAFSmass.

The material absorption μd is given by x-ray intensities before and after the sample, I_0 and I_1 : $\mu d = \ln(I_0/I_1)$, where μ is the linear absorption coefficient, d is the sample thickness. However, the quantities measured in a transmission XAS experiment are not I_0 and I_1 themselves, but some values proportional to them, e.g. currents of ionization chambers, $i_0 = \kappa_0 I_0$ and $i_1 = \kappa_1 I_1$. By taking the logarithm of their ratio, $\ln(i_0/i_1) = \mu d + \ln(\kappa_0/\kappa_1)$, one obtains a vertically displaced value of absorption. The second term, $\ln(\kappa_0/\kappa_1)$, is only slightly energy dependent and usually implicitly included into the background. This term can directly be measured by taking a spectrum of an empty experimental apparatus (hereafter referred to as "empty spectrum"), when μd is known to be zero. The absolute absorption is given then by $\mu d = \ln[(i_0/i_0^{\text{empty}}) \cdot (i_1^{\text{empty}}/i_1)]$. It should be noticed that the experimentally obtained i_0 and i_0^{empty} are not necessarily identical: the empty spectrum is a smooth function, it therefore can be measured on a sparse grid with short sampling time and then mathematically smoothed and interpolated. Moreover, in synchrotron radiation the beam instabilities are often much stronger than the statistical noise. Therefore the signals i_0 and i_1 (or i_0^{empty} and i_1^{empty}) are positively correlated and should be kept together in the ratio. Thus, even being measured on the same grid, the ratio i_0/i_0^{empty} must not be canceled.

VIPER provides a procedure for dividing transmission spectra by empty spectra. It may happen that i_1 of the empty spectrum was recorded with a different amplification. The corresponding factor must be remembered.



4.3 Marking glitchy regions

For visualizing suspicious regions and for subsequent deglitching of μ or χ , it is useful to mark glitches on the primary signals, i.e. currents, and to retain the marking on μ or χ curves. This can be done by clicking on the i_0 curve:



or by manual leveling of the logarithm derivative in the button bar: the lower is the level the more regions are marked by color.

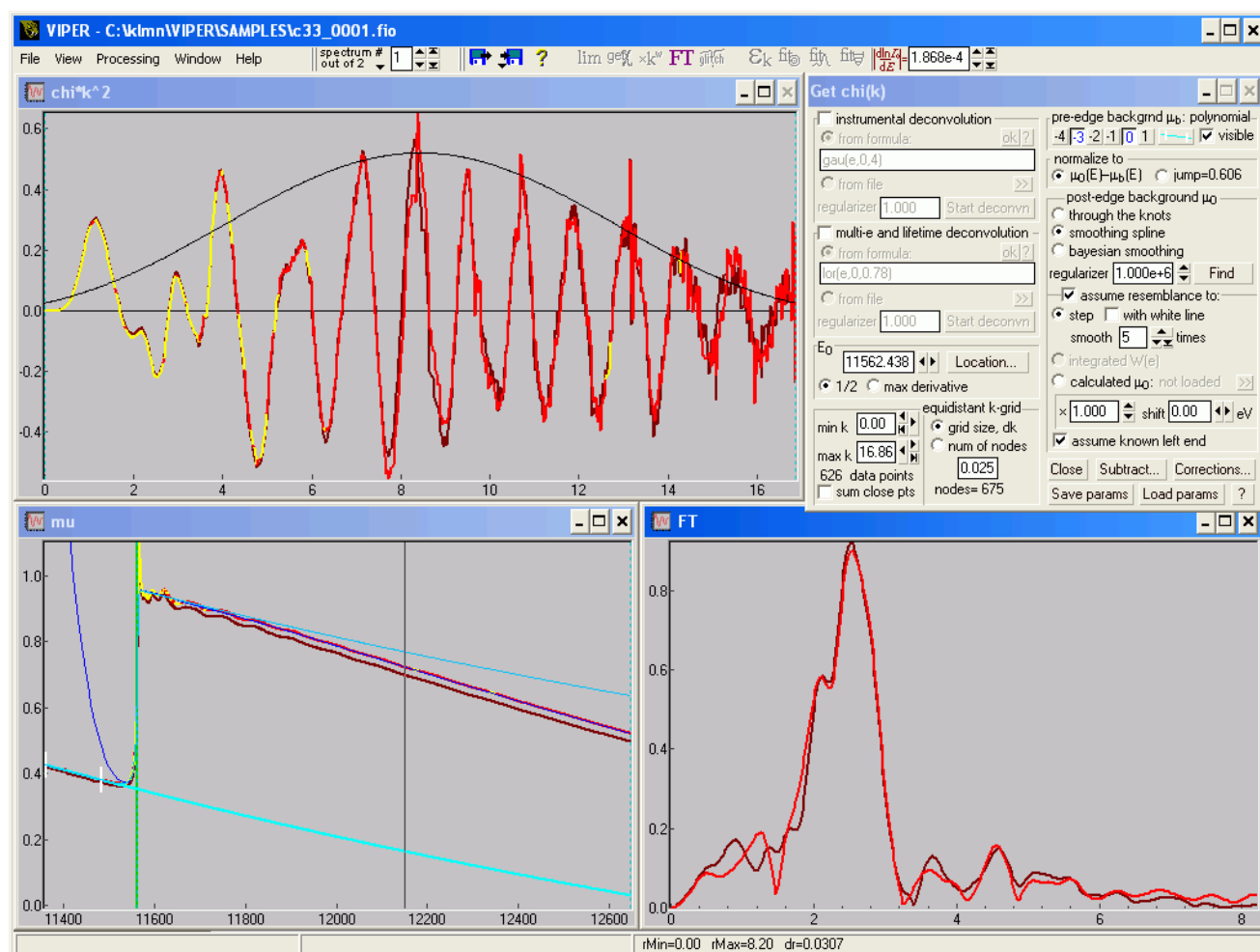
Alternatively, you can use the pop-up menu command "Mark region as a glitch".

4.4 Getting μ and χ

Press the 'get χ ' button or F2:



to get this:



4.5 How to go back to the Currents window?

You may need this sometimes, e.g. if you want to redo energy calibration or to switch the acquisition mode (fluorescence or absorption) when you did it wrong. The Currents window is hidden behind the other windows. To switch to it, use the top menu 'Window' or the standard Windows combination Ctrl+F6. You can also close the μ window; all the downstream windows will close as well. Then you can do 'get χ ' again with all the parameters in 'Get chi(k)' dialog restored automatically.

5 Working with μ and χ

If you start right from here, load Samples/glitch.vpj project.

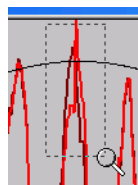
5.1 Deglitching

Note that any manipulation with the data will *not* change the original files. If you want to save the deglitched curves, export them to new column files, see Section 10. Even if you save a project file [a project file describes all the loaded files and actions on them], the deglitching steps will not be repeated when you load the project file.

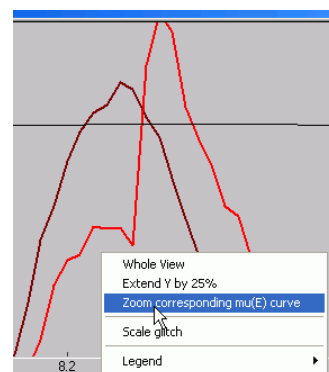
Also note that there is no undo for the deglitching manipulations. If unsatisfied with the result, you have to load your data again.

5.1.1 Step (jump) glitches

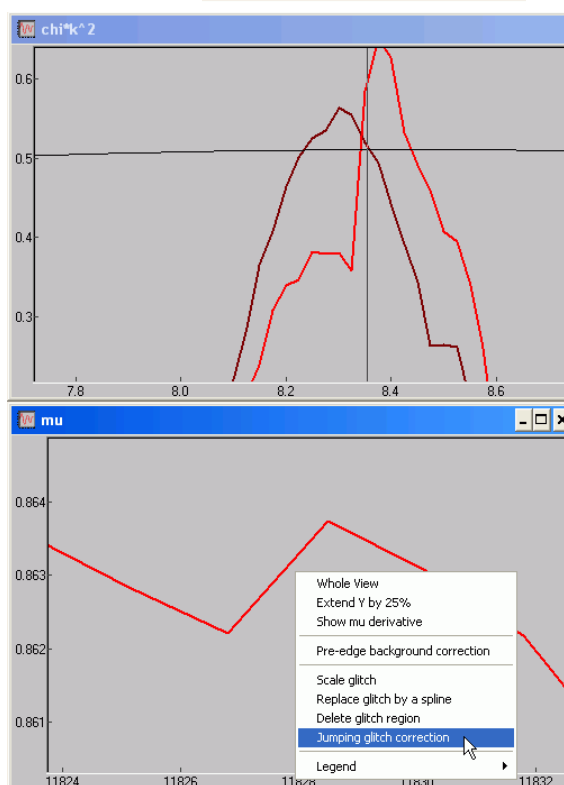
1) zoom the step glitch on χ spectrum #2 at $\sim 8.3 \text{ \AA}^{-1}$:



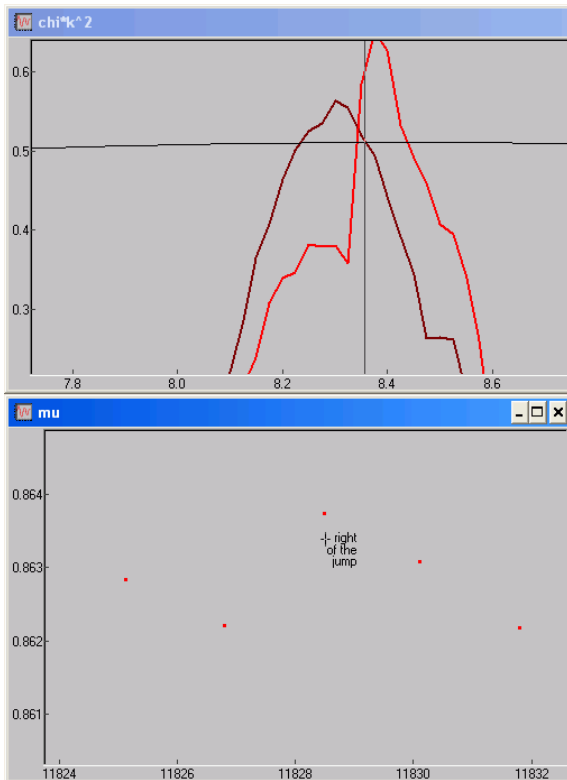
2) In χ window use the pop-up menu command 'Zoom corresponding $\mu(E)$ curve',



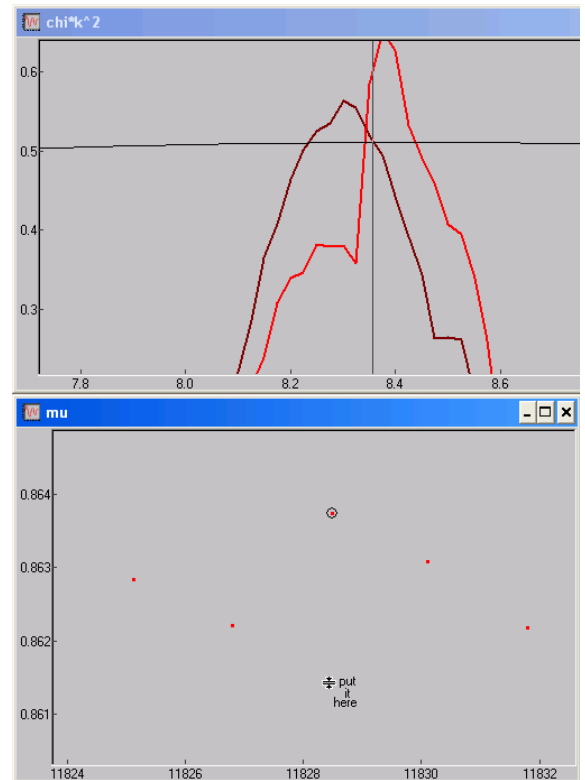
3) and click on the glitch. Now the glitch is zoomed also in μ window. Zoom it further as convenient. In μ window use the pop-up menu command 'Jumping glitch correction':



4) Now click on the right of the jump...



5) ...and then on the desired position for it:

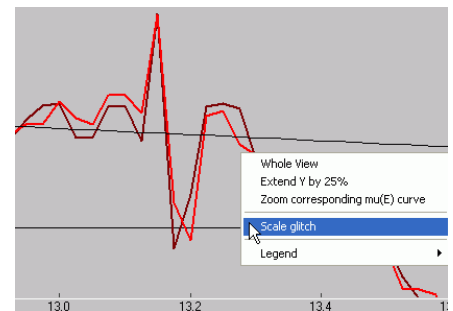


5.1.2 Sharp glitches

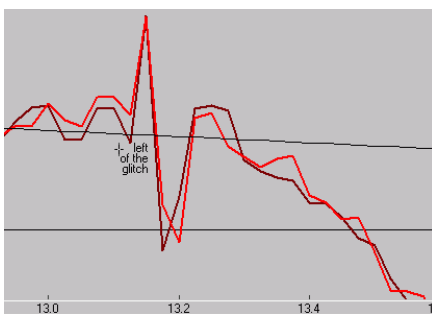
"Usual" sharp glitches are seen on both spectra at $\sim 13.2 \text{ \AA}^{-1}$.

5.1.2.1 Scale (compress) glitch

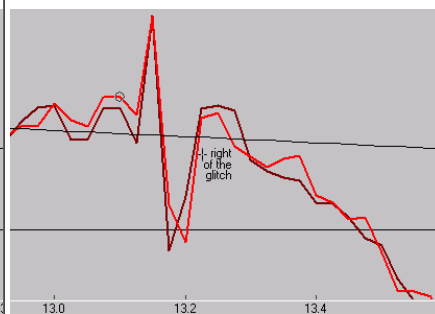
1) Zoom the glitch on χ curve by left click-and-drag and use the pop-up menu command 'Scale glitch':



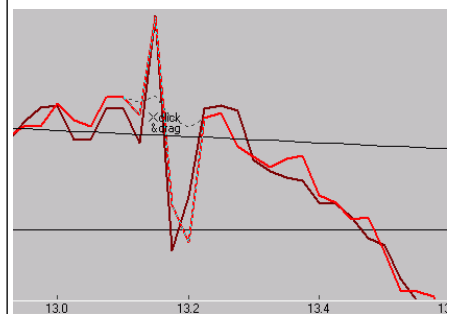
2) select the left of the glitch:



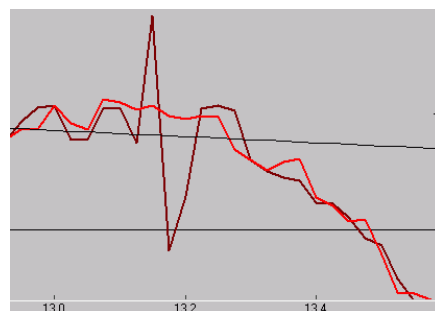
3) and the right of the glitch:



4) click somewhere below or above the glitch and drag into the glitch:



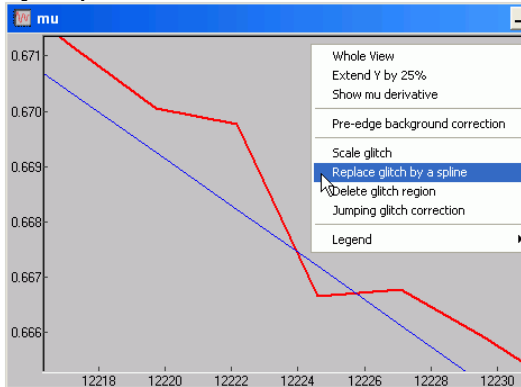
5) After mouse button release:



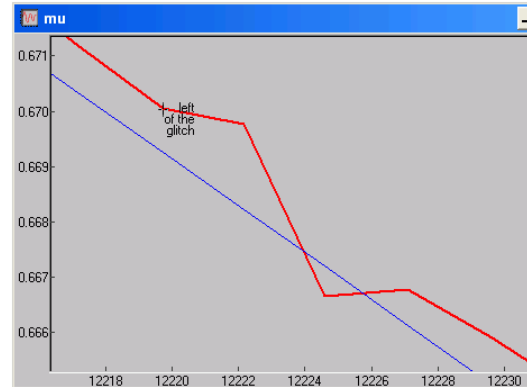
5.1.2.2 Replace glitch by a spline

Switch to the other spectrum and in χ window use the pop-up menu command 'Zoom corresponding mu(E) curve'. Magnify it further in μ window.

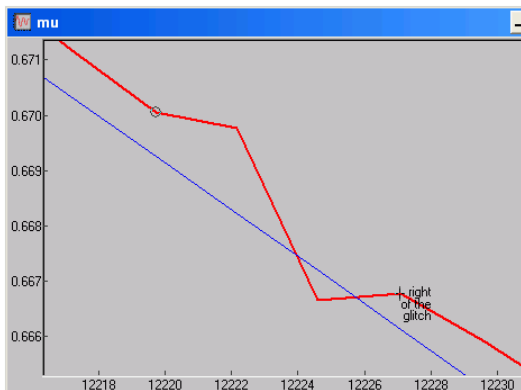
1) Use the pop-up menu command 'Replace glitch by a spline':



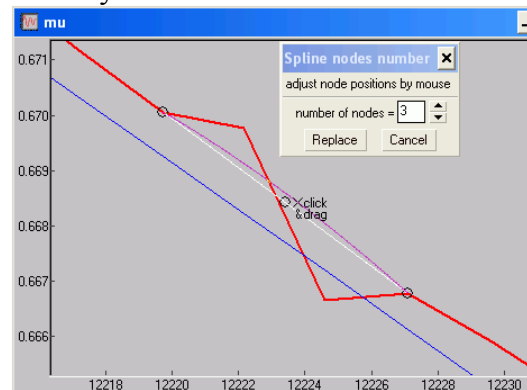
2) Select the left of the glitch:



3) the right of the glitch:



4) Specify the number of nodes and adjust them by mouse:



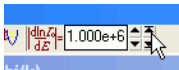
Watch how the χ curve is changing. Press 'Replace' when ready.

5.1.2.3 Delete glitch

Zoom the glitch on χ curve and then on μ curve (pop-up menu 'Zoom corresponding mu(E) curve'). In μ window do 'Delete glitch region'.

5.1.3 Switching off glitch coloring

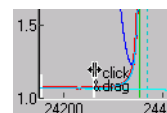
The glitch marking can be switched off by setting a big threshold value to the logarithm derivative in the button bar:



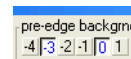
5.2 Pre-edge background

The pre-edge background is constructed by polynomial interpolation over the region specified by mouse:

or in a small dialog invoked via this small button:



The polynomial law is given by the power buttons:



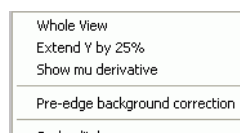
(modified Victoreen)

For absorption spectra measured in transmission mode, usually a Victoreen polynomial $aE^{-3}+bE^{-4}$ or a modified Victoreen polynomial $aE^{-3}+b$ is implied, where the coefficients are found by the standard least-squares method.

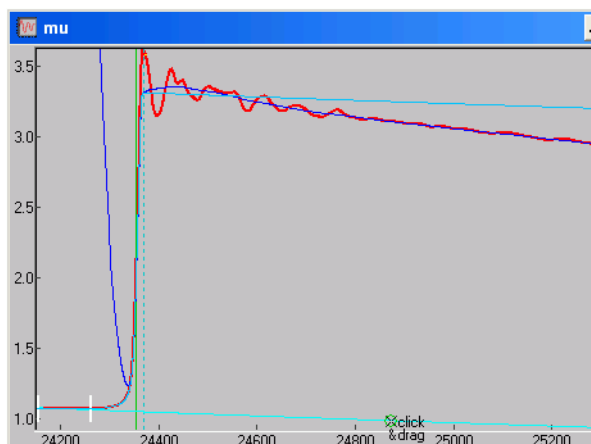
For absorption spectra measured in fluorescence mode, background subtraction is frequently not needed (unselect all the power buttons). More frequently a constant shift is sufficient (select button "0"). Sometimes the spectra exhibit a net growth with energy, which can be approximated by a linear law (select buttons "0" and "1").

5.2.1 Corrections of pre-edge background

On some low-quality spectra the background behaves strangely: it bends up or even may cross the μ curve. For example, a reason for this can be strong self-absorption. Whatever the reason is, you can correct this: do (i) normalization to a constant value (see Section 5.6) and (ii) apply corrections to the background. For the latter, use the pop-up menu 'Pre-edge background correction' in μ window:

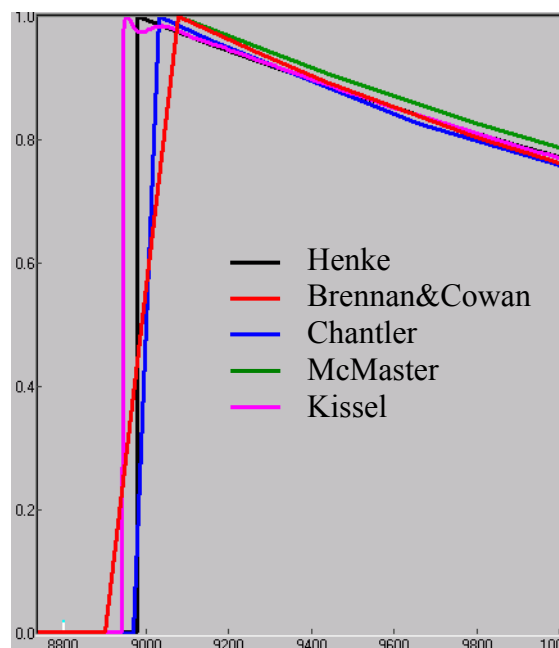


and click somewhere you believe the background must go through. You can then drag the correction node by mouse (right picture).



5.2.2 How the far end of μ should behave?

There is quite a widespread belief that μ above an absorption edge must in average go parallel to abscissa. Some EXAFS analysis programs offer dedicated procedures to achieve this behavior, e.g. a post-edge polynomial fit with subsequent subtraction. In order to check whether this is true, let us take various tabulations of atomic scattering factors using XAFSmass (its webpage provides the theoretical references used) and calculate the linear absorption coefficient as $\mu \propto f_2/E$ with the subsequent pre-edge subtraction and normalization using XANES dactyloscope. The graph on the right shows μ 's for Cu at the K-edge with a typical EXAFS length of ~ 1 keV. The curves look different at the edge because of different energy grids; some of them are very sparse. Important here is the far end behavior. As seen, μ 's are not constant and do decrease by $\sim 20\%$ at 1 keV.



5.2.3 Which polynomial to choose for the pre-edge of transmission spectra?

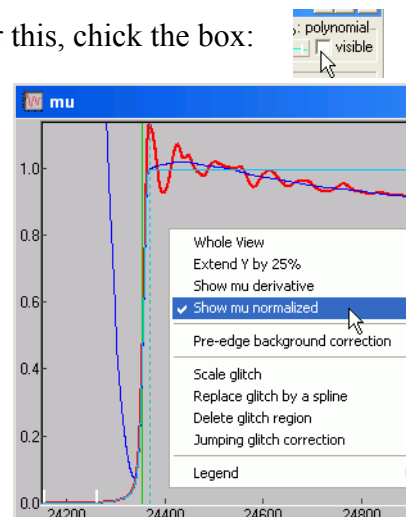
The usual options are Victoreen or modified Victoreen. Strictly speaking, the Victoreen polynomial is correct only for the true, not vertically displaced, absorption. The experimentally determined μ 's are displaced (see Section 4.2). Moreover, the shift is not constant but (weakly) energy dependent due to the energy dependence in (i) source/optics properties, (ii) efficiency of ionization chambers or PIN diodes, (iii) transmittance of windows, air paths etc. if any. To my experience, the different options on the pre-edge polynomial give EXAFS curves which differ from each other by a factor similar to those given by (i)–(iii). The total energy dependent displacement of absorptance is usually <0.1 at ~ 1 keV above the edge. This propagates to some uncertainty in Debye-Waller factors which is, in my experience, always smaller than the fitting errors.

Finally, the answer is: it does not really matter because there are more important sources for the errors in the sought structural parameters. I usually use the modified Victoreen.

5.2.4 Show μ normalized

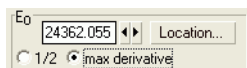
When you play with the pre-edge it is convenient to have it visible. For this, chick the box:

When you uncheck it, i.e. when you do the subtraction, you can normalize μ by the pop-up menu command 'Show mu normalized' (this command is not available with non-subtracted background):



5.3 Setting E_0

Usually E_0 energy (the origin for the photo-electron wave number) is set at the first inflection point of μ :



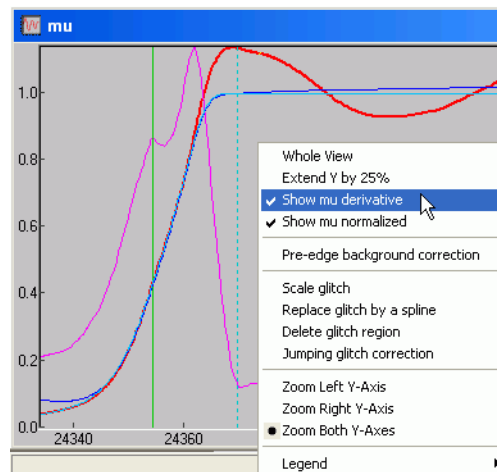
Frequently absorption edges have several peaks on the 1st derivative. The first one is not necessarily the highest. Therefore the automatic peak search may fail. Watch out: sometimes there are two peaks of almost the same height and in two different scan repetitions the automatically E_0 can suddenly jump so that EXAFS will look differently. This happened with K-edges of some 4d elements.

You can visualize the derivative curve (picture on the right) and move E_0 to the first peak manually (by mouse).

It is not important where to put E_0 because usually E_0 is one (or more if used independently for different coordination shells) of the fitting parameters (see Section 8.1.2).

Now, after E_0 has been defined, the energy dependence of μ is transformed to the wave number dependence as:

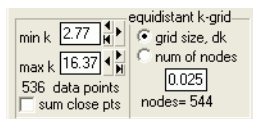
$$k = \sqrt{2m_e(E - E_0)/\hbar}$$



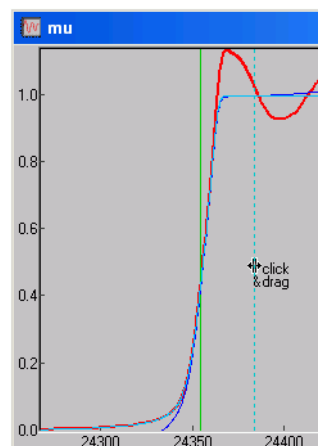
5.4 Setting k mesh

The limits k_{\min} and k_{\max} can be set by mouse in μ and in χ windows :

The equidistant k mesh can be specified by the grid size dk or by the number of nodes.



The absorption coefficient μ is transformed to this mesh by spline interpolation. If the option 'sum close points' is checked then close experimental points are summed when fall into a single dk segment.



5.4.1 How to choose k_{\min} ?

The answer depends on the way of how μ_0 is constructed. See Section 5.5.1 or Section 5.5.2.1. k_{\min} also affects Fourier transform, see Section 6.4.

5.4.2 How to choose k_{\max} ?

First set it to maximum (i.e. all available data are used). See where signal of χ becomes weaker than noise, either by eye or using one of the procedures given in Section 7. k_{\max} also affects Fourier transform, see Section 6.4.

5.4.3 How to choose dk ?

Read [Press et al., *Numerical Recipes* (1992-2007) chap. 12.1 Fourier Transform of Discretely Sampled Data]. In short: if your $\chi(r)$ is believed to contain nothing but noise above the critical distance r_c then $dk=1/(4r_c)$, where an additional factor 2 in the denominator is due to the doubled kr product in the EXAFS FT. For typical $r_c=10\text{\AA}$ this makes $dk=0.025\text{\AA}^{-1}$. If your dk is greater, the Fourier-transformed signal $\chi(r)$ at $r>1/(4dk)$ will be folded or "aliased" into the range $r<1/(4dk)$, which gives a distorted FT, especially at large r . For example, if your $dk=0.1\text{\AA}^{-1}$, you can analyze EXAFS only up to $r_c=2.5\text{\AA}$, i.e. only the 1st coordination shell, with $\chi(r)$, possibly, strongly aliased.

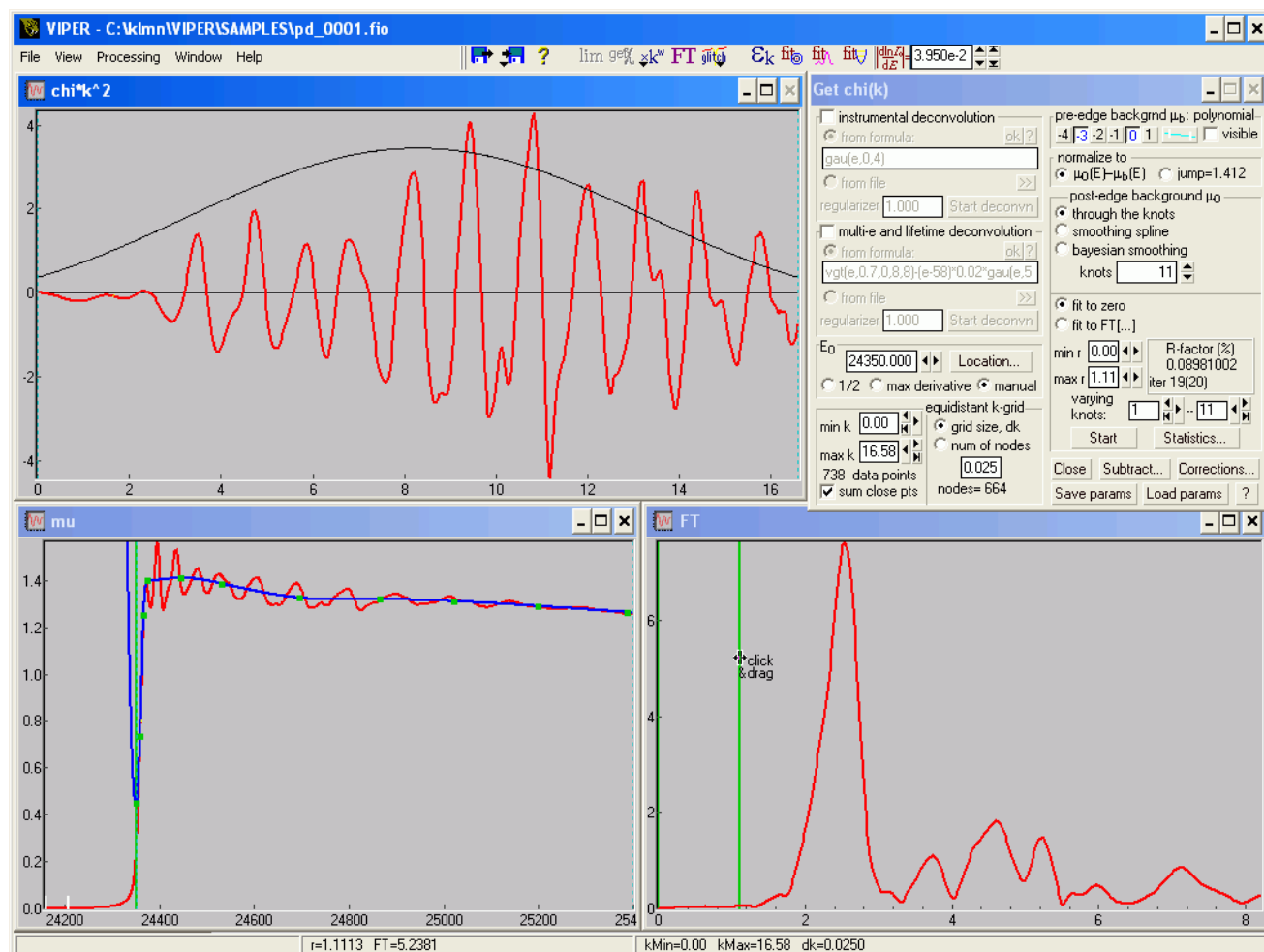
Note, you have to *measure* with small enough dk , it does not make sense to interpolate to a finer mesh afterwards!

5.5 Construction of post-edge background μ_0

VIPER offers three ways of constructing the post-edge background μ_0 (called also atomic-like absorption): (i) by a spline drawn through the knots varied to minimize the low- r EXAFS FT part; (ii) by a smoothing spline and (iii) by a Bayesian smoothing curve.

5.5.1 μ_0 as a spline drawn through varied knots

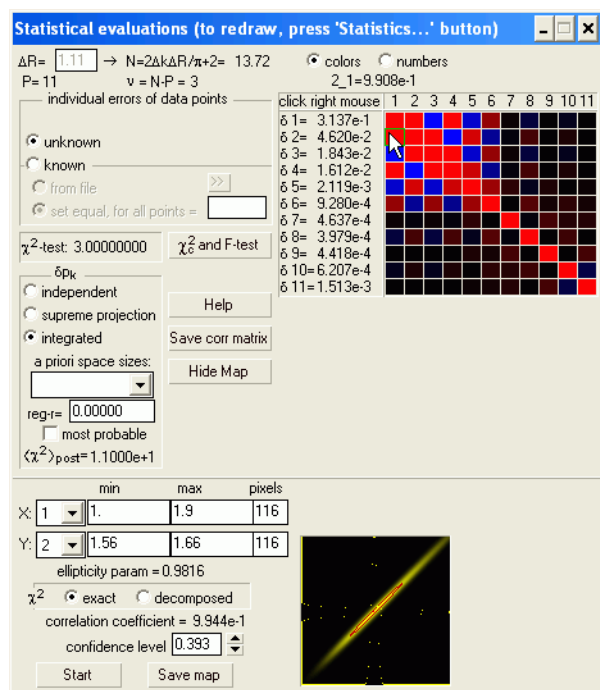
This example can be loaded as *Samples/PdKnots.vpj*:



In this procedure you can manually pre-adjust the knots, also horizontally (use the pop-up menu in μ window for this). By default, the knots are positioned equidistantly in k -space. Now choose the minim-

ized FT region and let the minimization start. I recommend to begin the fitting with a few first knots and then after several Start/Stop cycles gradually to include all knots. The maximum available number of knots can be seen in the Statistics dialog as denoted by N. See Section 9.8 for details on statistical evaluations.

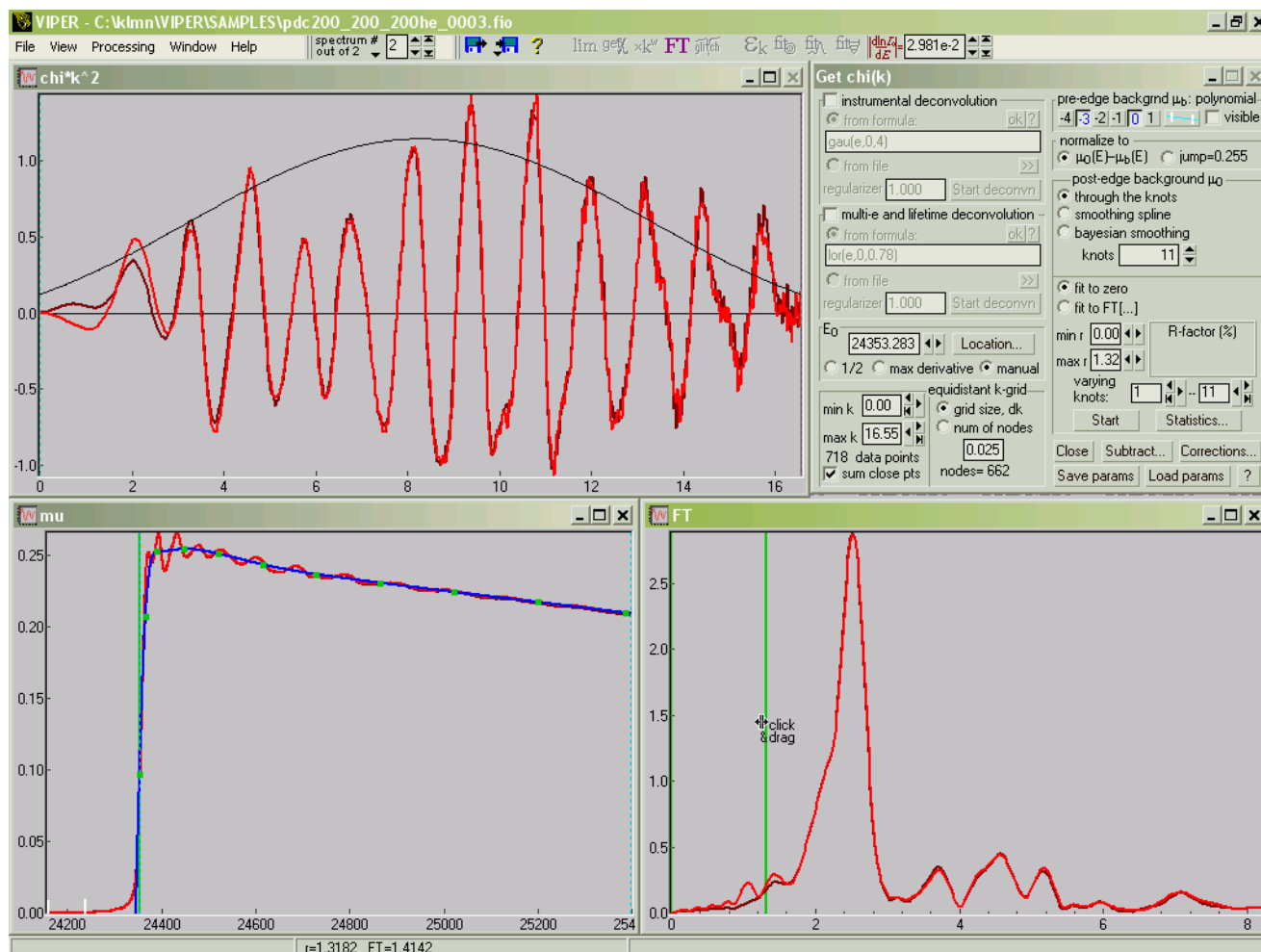
The provided example already has a well optimized set of knots, so you may play with the Statistics dialog right away:



The colored matrix shows the pair-correlation coefficients. Completely red and blue denote +1 and -1, black is 0. The coefficient r_{21} is pointed at by the cursor and its value (≈ 0.991) is displayed above the matrix. Such a strong correlation means that the first knot may acquire a variation and the thus induced difference in the target function (here, the power spectrum in the minimized FT region) can be almost fully compensated by a variation of the second knot. These words are illustrated by the black-yellow graph showing a constant-level map of the target function. The map is strongly stretched, which shows that uncertainty in one fitting variable is projected onto uncertainty in the other fitting variable. This leads to big fitting errors for the first few knots and therefore for $\chi(k)$ at $k < \sim 3 \text{ \AA}^{-1}$. The fitting errors are listed at the left of the correlation matrix.

This procedure is also not stable in respect to the variation in the number of the knots. In the following

examples the spline is drawn through 11 and 12 knots. The two χ 's differ at low- k :



Finally, the method works well when χ is cut off at low k . There is quite an advantage of this method: μ_0 contains only low-frequency oscillations whereas χ preserves all the true structural oscillations. This is not the case for the smoothing techniques (see below), where the resulting χ partially loses its signal and transfers it towards μ_0 .

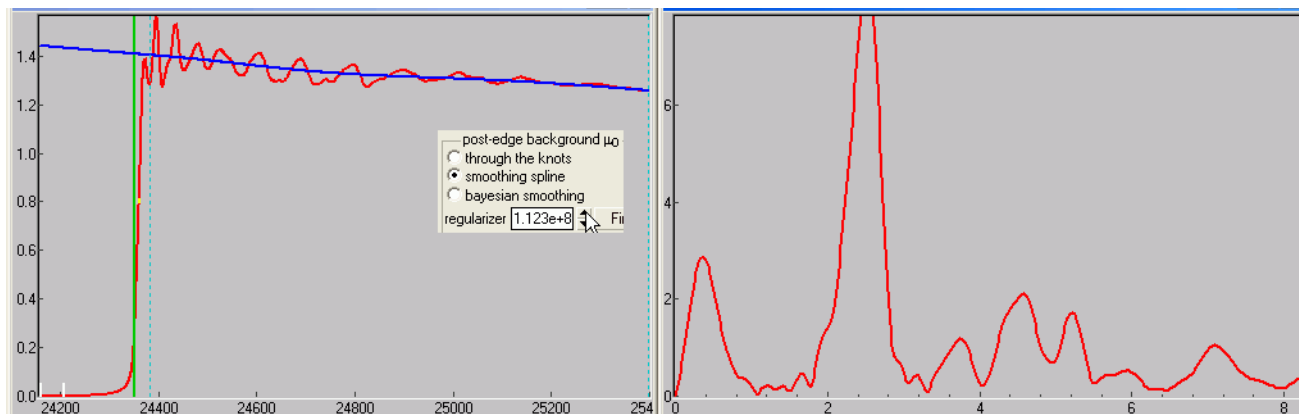
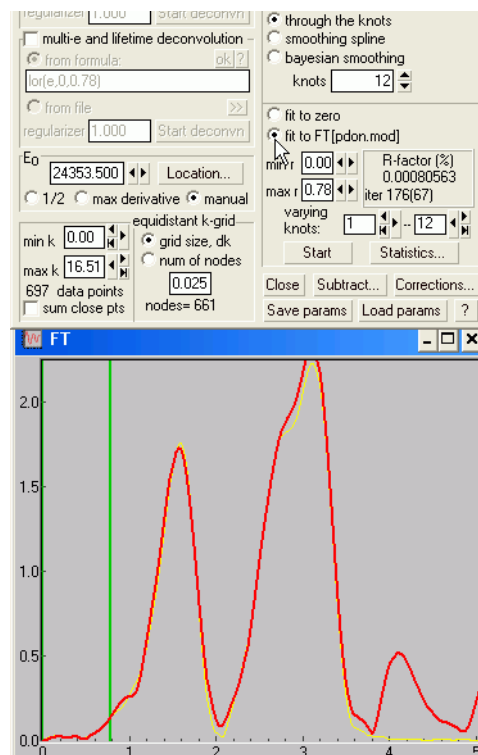
There is a possibility to fit the power spectrum in the minimized FT region not to zero but to the power of another, usually calculated, spectrum. On one hand, this procedure accounts for the leakage of some signal from the 1st coordination shell into the minimized FT region. On the other hand, a calculated spectrum assures that there is no unphysical signal in that region. This option is especially important for oxides where, contrary to metals, the 1st coordination shell is at relatively short distance:

5.5.2 μ_0 as a smoothing spline

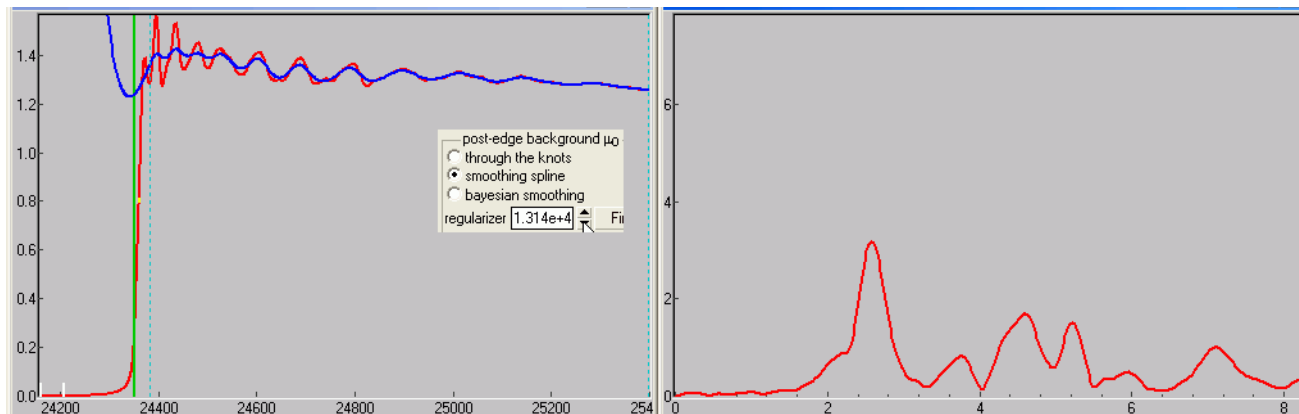
This method is the simplest and therefore, probably, the most frequently used. The μ_0 curve is constructed from the experimental μ spectrum by smoothing out the EXAFS wiggles.

μ_0 is weakly energy dependent as can be seen on absorption spectra of gases. Moreover, there can be weakly changing experimental factors to it, as listed in Section 5.2.3. Thus, the smoothing spline should allow for such slow variations.

The smoothing spline depends on a parameter that specifies its stiffness. If the parameter is big, the spline approaches a straight line and disallows the slow variation. This is manifested by the non-physical signal in the low- r FT:

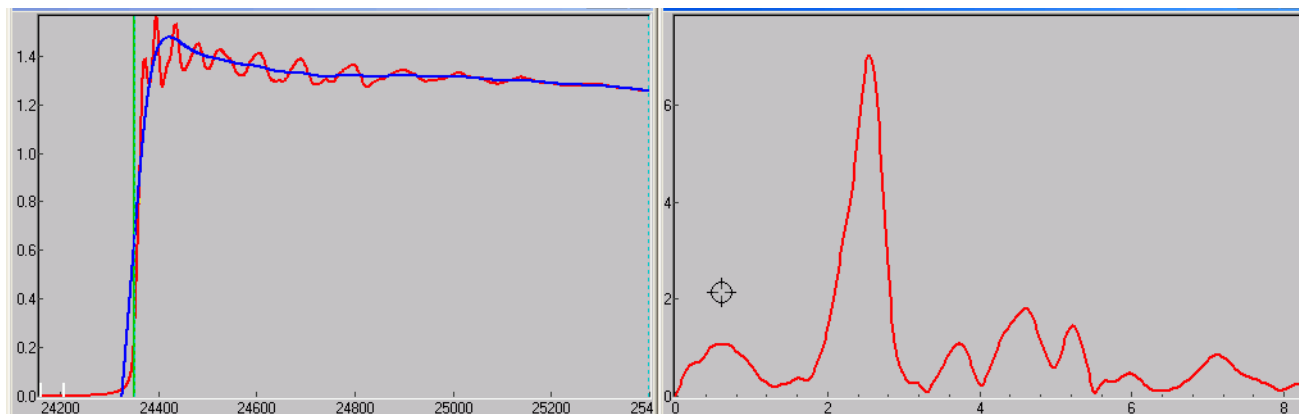


If the parameter is small, the spline approaches the μ curve and thus χ goes to zero:

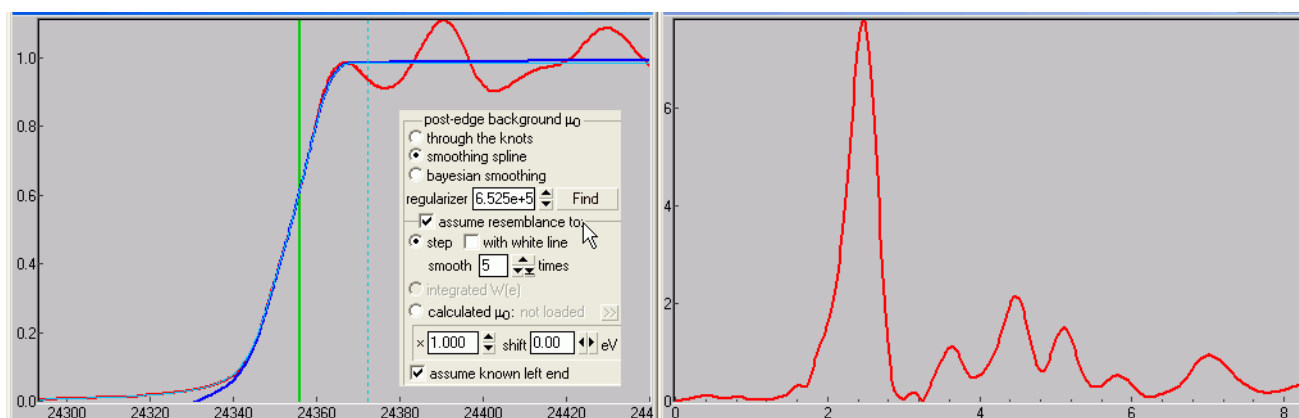


Hence, between the extreme inappropriate values there must be an optimum. VIPER offers an automatic procedure to find it, as described in [Klementev, 2001a]. To my experience, this procedure tends to somewhat underestimate the smoothing parameter. How to judge if it is good? See Section 5.5.2.1.

If you now set k_{\min} to zero, you see that because the spline is relatively stiff, it bends too high right after the edge and the low- r FT gets some false signal:



Nevertheless, one can go to lower k values due to the following advanced feature of the smoothing spline procedure. It can take into account *a priori* knowledge on the shape of the final μ_0 curve. The same also holds for the Bayesian smoothing procedure. See [Klementev, 2001a] for detail. We can safely assert that μ_0 looks like a step function:



This step function can be constructed from the experimental spectrum: it follows μ until it reaches the height of the edge, thereafter it is constant. This step can be smoothed and moved horizontally and vertically. The latter affects $\chi(r)$ at low- r quite strongly. See the next Section for how to set it.

5.5.2.1 Criterion for a good μ_0 . AXAFS.

How to set k_{\min} , smoothing parameter, step height etc? The criterion is simple: no false FT signal until the peak of the 1st coordination shell.

How about AXAFS?

Some EXAFS researchers claimed "atomic-XAFS" ("AXAFS") to give a significant contribution into low- r FT signal. Almost exclusively they belong(ed) to Koningsberger's group (Inorganic Chemistry in Utrecht University). There are some papers on AXAFS also by Baberschke et al. (Physics, FU Berlin).

The AXAFS phenomenon is explained as the usual EXAFS with the difference that the photoelectron scattering occurs not by the atomic potentials but rather by interstitial potential or charge density. Wende & Baberschke [1999] (also Rehr et al. [1995]) have calculated AXAFS by FEFF and found it to be quite strong. Several works by Koningsberger et al. [Ramaker et al., 2000] established a procedure on how to separate AXAFS from multi-electron excitations which may also contribute to the low- r $\chi(r)$ signal.

I have spoken to many EXAFS people on this topic and I think there is a strong belief that AXAFS is more artifact than a real phenomenon that can be used as a research instrument. My personal doubts are the following:

1) For a wave to be efficiently scattered by any object, the wave length must be of the order of the object size. For an electron to be efficiently back-scattered by a free electron (and electrons in the interstitial regions are almost free), its wave length must be of the order of the electron radius, $r_0 \sim 3 \cdot 10^{-5}$ Å. Typical photoelectron wave lengths in EXAFS regime are $\sim 10^4$ times longer than r_0 . Hence, if such scattering exists, it is by atomic potentials, not by local charge density, as claimed by Baberschke et al. Therefore the numerous conclusions on charge transfer made on the basis of AXAFS are questionable.

2) The amplitude of k -weighted AXAFS reported by Wende&Baberschke is stronger than 0.1 Å^{-1} in the range $k < \sim 10 \text{ Å}^{-1}$. This is comparable with the contribution of 6 oxygen atoms in the 1st coordination shell! Why then the next-shell interstitial contributions have never been reported and the standard EXAFS, without AXAFS, works well, as proven on many reference compounds? Where is the distant-shell AXAFS?

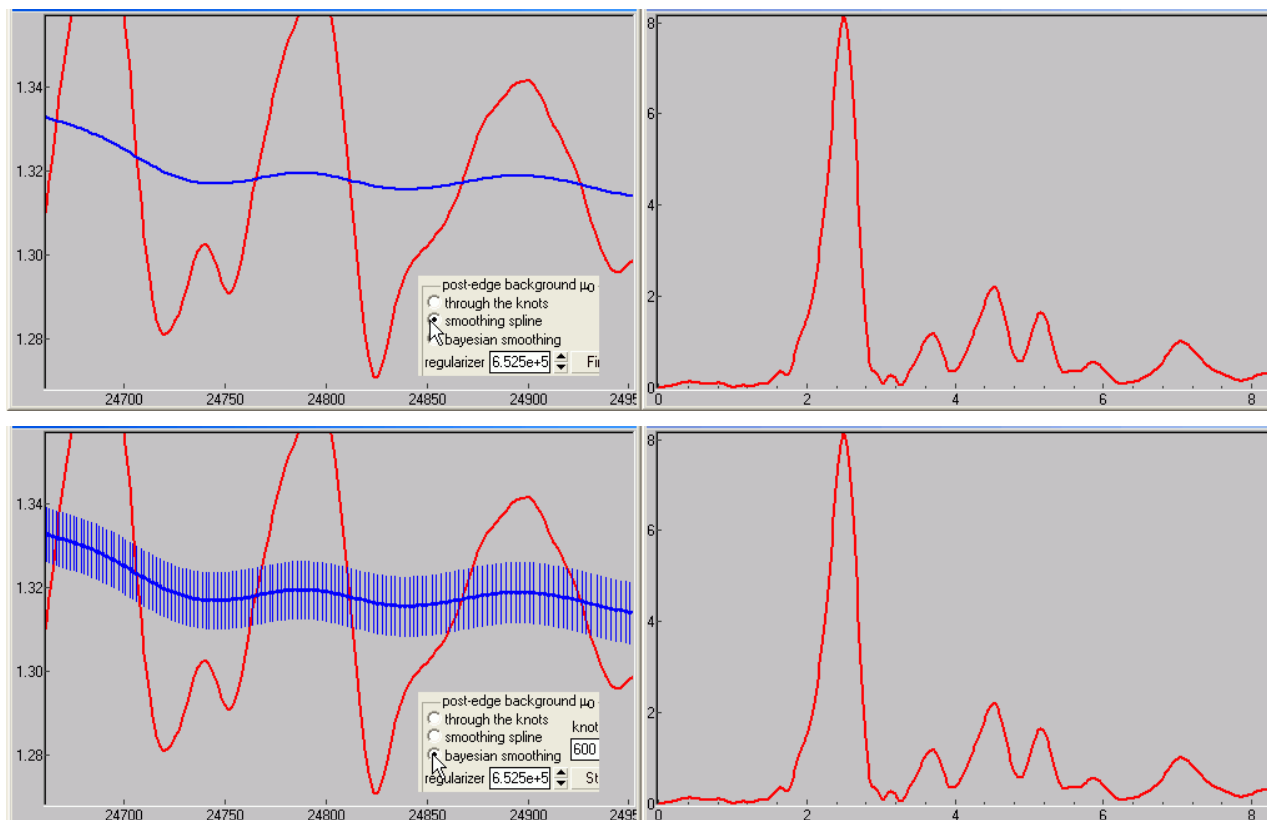
3) The weak-potential inter-atomic regions are large. When treated in a usual EXAFS way, they should have huge distance variance of the order $\sigma^2 \sim 0.1 \text{ Å}^2$. The Debye-Waller factor $\exp(-2\sigma^2 k^2)$ must damp the AXAFS oscillations already at $k > 3 \text{ Å}^{-1}$. For such a strong and far oscillating AXAFS, as given by FEFF, the scattering potential must have a very well localized feature. It is there, the muffin-tin wall. Thus the calculated AXAFS seems to be an artifact due to the muffin-tin jumps rather than a real absorption feature.

4) Why in the big article by Rehr & Albers [2000] there are only two sentences about AXAFS, and nothing about it as a real tool? They write (in Section II.B.1): "However, a definitive theory of atomic XAFS will likely require corrections to the muffin-tin approximation." I read this as a soft version of "Sorry, FEFF cannot (and could not) calculate it". In fact, the whole modern story (after 1995) about AXAFS began from FEFF calculations, and it seems that without FEFF AXAFS has no confirmation.

5) Does any other EXAFS code see AXAFS? To my knowledge, no. But I might be wrong here.

5.5.3 μ_0 as a Bayesian smoothing curve

This procedure is rather slow. For speeding it up, there is an option of the preceding n -point smoothing where the experimental points are grouped into N_{points}/n knots. The Bayesian smoothing curve is very similar to the smoothing spline. When without the preceding smoothing, it is exactly the same as the smoothing spline. However the Bayesian μ_0 has an important addition: as the final values are calculated in terms of a probability function, also the error bars can be simultaneously calculated:



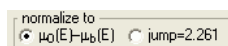
These pictures show, by the way, a case with a bit too little smoothing parameter; as seen with such high magnification, μ_0 follows partially the EXAFS wiggles and the χ obtained has lost some amplitude.

5.6 Normalization of χ

Given μ_0 , the EXAFS function χ is determined as

$$\chi = \frac{\mu - \mu_b - \mu_0}{\mu_0 - \mu_b}$$

The normalization, i.e. the division by the denominator, can be taken either energy dependent or constant:



The former option is the correct and usual one, and the latter is necessary sometimes when experimental troubles make the absorption coefficient bend up or down. The constant normalization may give by $\sim 20\%$ decreased χ at the far end of the spectrum (at ~ 1 keV or ~ 16 Å⁻¹); the corresponding overestimation in σ^2 is $\sim 5 \cdot 10^{-4}$ Å².

5.7 Corrections to μ and χ

$\mu(E)$, $\chi(E)$ or $\chi(k)$ can be corrected by user-defined functions $f(E)$ or $f(k)$ in order to correct, for example, the fluorescence self-absorption. The function $f(E)$ may refer to the theoretical tables of absorption coefficients incorporated into VIPER. There is a specialized self-absorption correction routine that works for both XANES and EXAFS regions.

5.7.1 Self-absorption correction

Many papers have addressed the self-absorption effect. Most of them provided restricted correction. The early papers by [Goulon et al. 1982; Tan et al. 1989; Tröger et al. 1992] were limited only to the EXAFS case. The correction functions there had discontinuity at the edge and thus were not applicable to XANES. Moreover, those works provided corrections only for infinitely thick samples with an exception of [Tan et al. 1989] where also thin samples were considered but only as pure materials (e.g. single element foils).

The first self-absorption correction for the whole absorption spectra (also including XANES) was proposed with two different strategies by Eisebitt et al. [1993] and Iida and Noma [1993]. Eisebitt et al. [1993] estimated the two unknowns μ_{tot} and μ_X (see the notations below) from two independent fluorescence measurements with different positioning of the sample relative to the primary and fluorescence beams. An obvious disadvantage of this method is that it is solely applicable to polarization-independent structures (amorphous or of cubic symmetry). On the other hand, it does not require any theoretical tabulation, which is the case in the method of Iida and Noma [1993], who proposed the background part $\mu_{\text{back}} = \mu_{\text{tot}} - \mu_X$, to be taken as tabulated. The advantage of their approach is its applicability to any sample with only one measurement. Moreover, this method is applicable to samples of general thickness, not only to thick samples as required by the method of Eisebitt et al. [1993]. It is the method of Iida and Noma [1993] which is implemented, with some variations, in VIPER and XANES dactyloscope. The method was re-invented (i.e. published without citing Iida and Noma [1993]) by Pompa et al. [1995], Haskel [1999] and Carboni et al. [2005]. These three works, however, were simplified down to infinitely thick limit.

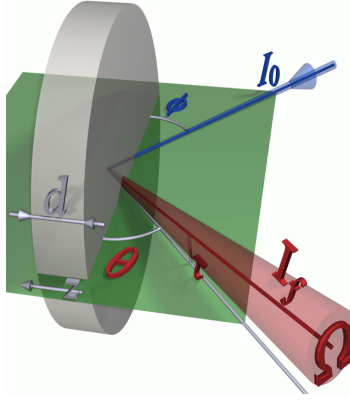
The correction was extended somewhat by considering a variable escape angle in order to account for the finite (not infinitely small) detector area: only in the synchrotron orbit plane, in EXAFS [Brewer et al. 1994] and also out of plane: in EXAFS [Pfalzer et al. 1999] and XANES [Carboni et al. 2005]. All three works operated in the thick limit. To my believe, detector pixels are always small in the sense that the self-absorption effect can be considered as uniform over each single pixel and therefore the correction can be done only for one direction towards the pixel center.

An interesting approach to correcting the self-absorption effect was proposed by Booth and Bridges [2005] who considered another small parameter, not the usual $\exp(-\mu d)$, which allowed simplifying the formulas also beyond the thick limit but the treatment was limited to EXAFS.

Another re-invention of the Iida and Noma method with calling it “new” was presented by Ablett et al. [2005]. The merit of that work was implementing the method without restriction to the thick limit and providing many application examples and literature references.

5.7.2 Description of self-absorption correction

The derivation of the fluorescence intensity can be found, with different notations, in almost all the papers cited above. Here it is repeated because VIPER and XANES dactyloscope add some extra factors. The standard expression for the fluorescence intensity originated from the layer dz at the depth z is given by the trivial sequence of propagation and absorption (with neglected scattering):



$$dI_f(z, E) = \underbrace{I_0}_{\text{primary flux}} \underbrace{e^{-\mu_T(E)z/\sin\phi}}_{\text{primary x-ray transmitted to depth } z} \underbrace{\mu_X(E)}_{\text{absorbed in layer } dz \text{ due to edge of interest}} \underbrace{\frac{dz}{\sin\phi}}_{\text{transformed into fluorescence}} \underbrace{\epsilon_f}_{\text{directed into solid angle } \Omega} \underbrace{\frac{\Omega}{4\pi}}_{\text{fluorescence x-ray transmitted to detector from depth } z} e^{-\mu_T(E_f)z/(\sin\theta\cos\tau)}$$

where μ_T is the total linear absorption coefficient at the primary x-ray energy E or the fluorescence energy E_f , μ_X is the contribution from the edge of interest, ϵ_f is the fluorescence quantum yield – the probability to create a fluorescence photon from an absorbed photon. After integration over z from 0 to d :

$$I_f(E) = C \frac{\mu_X(E)}{\mu_T(E) + \mu_T(E_f) \frac{\sin\phi}{\sin\theta\cos\tau}} (1 - e^{-\mu_T(E)d/\sin\phi} e^{-\mu_T(E_f)d/(\sin\theta\cos\tau)}), \quad (*)$$

where the constant C includes all the energy independent factors and is treated as unknown because the actual solid angle is usually unknown and also because it implicitly includes the detector efficiency.

The total absorption coefficient is decomposed as $\mu_T = \mu_X + \mu_b$, where the background absorption coefficient μ_b is due to all other atoms and other edges of the element of interest. The constant C is found by equalizing all μ 's at a selected energy E_{norm} (“normalization energy”) to the tabulated ones. Now the equation (*) can be solved for μ_X at every energy point E , which is the final goal of the self-absorption correction.

When the sample is thick ($d \rightarrow \infty$), the exponent factors vanish. This “thick limit” approximation allows finding the μ_X by simple inversion of (*), without solving the non-linear equation, and is optional in VIPER and XANES dactyloscope.

5.7.3 Realization in VIPER

5.7.3.1 Extended correction options

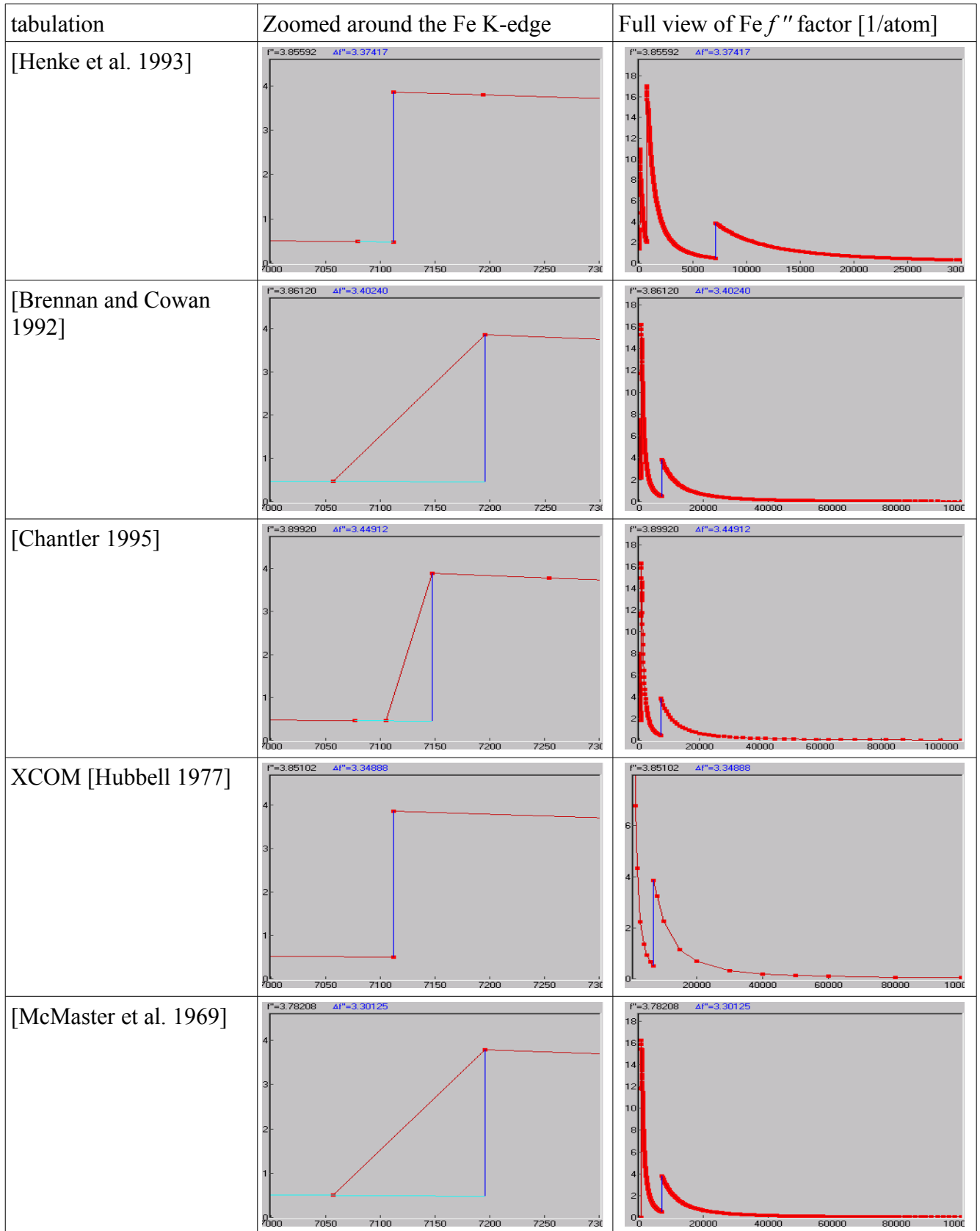
Some of the options offered by VIPER and XANES dactyloscope are non-standard (extended):

- 1) The additional term $\cos\tau$ in (*) is not quite standard; one can also find it in [Carboni et al. 2005] and [Ablett et al. 2005].
- 2) Absorption by air and by Kapton foils in front of the sample can be taken into account (see the examples below). For this, the primary flux is multiplied by $e^{-\mu_{\text{air}}(E)d_{\text{air}}} e^{-\mu_{\text{Kapton}}(E)d_{\text{Kapton}}}$. The similar term at E_f is implicitly included into the constant C .
- 3) μ_b is usually taken to be energy independent. In VIPER it is energy dependent.
- 4) One can select among five different tabulations of absorption coefficients (actually, scattering factors f'') in VIPER.

5.7.3.2 How the tables of scattering factors are used?

In order to use the equation (*), it is prerequisite to know the sample stoichiometry, i.e. the molar weighting factors x_i for each atom type i in the sample. Then the linear absorption coefficient is proportional to the atomic absorption cross section σ_a : $\mu_X \propto x_X \sigma_{aX}$ and $\mu_T \propto \sum_i x_i \sigma_{ai}$. The atomic cross sections, in turn, are calculated from the tabulated scattering factors f'' : $\sigma_a = 2 r_0 ch N_A f'' / E$.

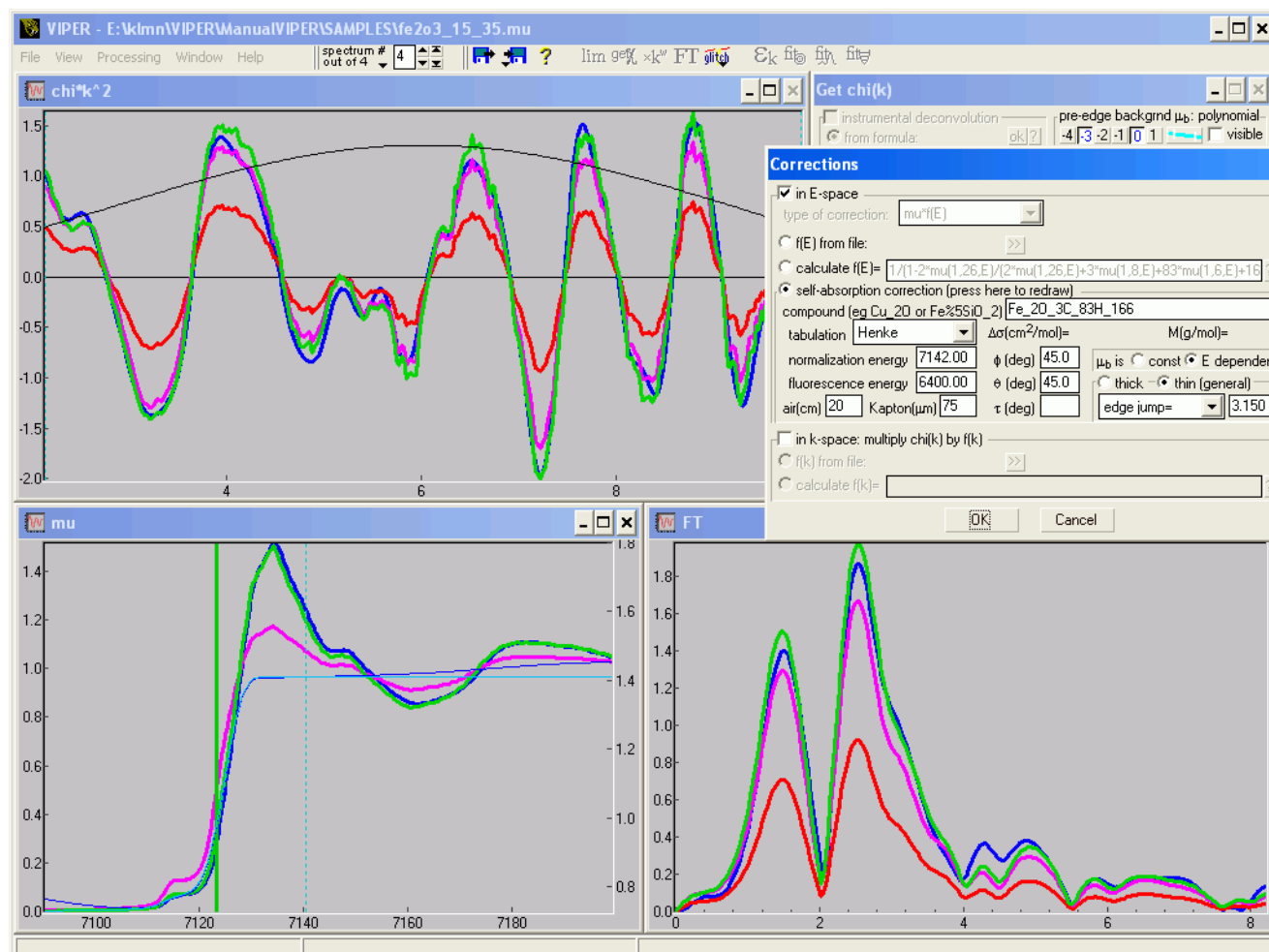
Since all the tabulations do not contain the partial contributions of each absorption edge of an element but only the combined result of all atomic shells, an isolation of μ_X and the pre-edge background is required. In VIPER and XANES dactyloscope this is done by extrapolating the pre-edge region by the Victoreen polynomial. The polynomial coefficients are found over only two pre-edge points, as the tabulations are usually sparse. As illustrated below for each tabulation used, the **edge jump** is the difference between the first post-edge value and the extrapolated **background**:



VIPER searches for an absorption edge (where the derivative is positive) within -250 eV from the specified normalization energy. When an edge is found, the jump in molar cross section is displayed.

5.7.4 Example of self-absorption correction

Load the example project *Samples/fe2o3_tr_fl.vpj*. It has a spectrum of Fe₂O₃ (hematite) measured in transmission and a three times loaded fluorescence spectrum. The sample is a 13-mm-diameter pressed pellet containing 11 mg of hematite mixed with 80 mg of polyethylene (PE) powder. The pellet was wrapped by adhesive Kapton foil.



As seen on the picture, the fluorescence spectrum, colored red (overlapped with the magenta one in the μ window), essentially differs from the transmission one, colored blue. The fluorescence spectrum is corrected by the Tröger's formula (magenta) and by the equation (*), colored green. Notice that the Tröger's approach does not correct the μ curve but χ . For this thick sample the difference between the two corrections is minor in k - and r -space. For thin samples the difference is significant.

The parameters for the self-absorption correction are seen in the screenshot above. It is essential to remember about the sample matrix or the supporting agent (here: PE) and to put its chemical formula as well. Here, the weight '83' of PE (CH₂) was calculated as

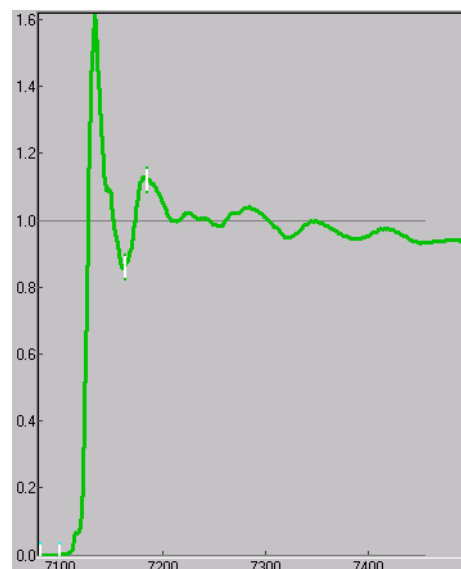
$$m_{\text{PE}}M_{\text{Fe}_2\text{O}_3} / m_{\text{Fe}_2\text{O}_3}M_{\text{PE}} = 80\text{mg}160\text{g/mol} / 11\text{mg}14\text{g/mol} = 83.$$

In order to use equation (*) for thin samples, one must provide the sample thickness. This could be the *physical* thickness; then one would need to know the sample density for calculating the linear absorption coefficient in the exponent. A more direct way is to use the *optical* thickness $\mu_t d$, or just its jump at the edge, which is usually possible directly to measure in transmission spectra (remember, we are speaking here about thin samples, otherwise use the 'thick' option). If the *physical* thickness is known, which is usual for foils, use the program XAFSmass to calculate $\mu_t d$ or $\Delta\mu_x d$ from the sample composition, the thickness and density.

In the example above, the edge jump was found from the transmission spectra times $\sqrt{2}$ because the transmission spectra were measured at normal incidence whereas the fluorescence spectra were taken with the same sample at 45°. [For future versions of the manual: redo this example with simultaneously measured transmission and fluorescence]



Equation (*) is also useful for correcting the high-energy behavior of μ . This correction is especially relevant to samples with low concentration of the element being probed or the samples measured in air or at low energies. In these cases the energy dependence of the background absorption μ_b and air absorption become important.



The left picture differs from the right one by added 20 cm of air.

The energy dependent μ_b and air absorption should always be opted. The option ' μ_b is constant' is meant for illustration and for comparison with other programs.

5.8 k-weighting of χ

The k -weighting power can be selected by the button " $\times k^w$ ". Fly the mouse cursor over it without clicking or with clicking to get more powers.



5.8.1 What k -weighting to use?

Usually people use k , k^2 or k^3 . The rationales people tell for using the weighting functions are (i) to compensate the amplitude decay at high k values, (ii) to shift the sensitivity towards high-Z neighbors and (iii) to reduce the correlations between the EXAFS parameters when doing multiple k -weighting fits. However, *if you just multiply by k^w without increase of sample time or without incorporating the increased noise into the statistical analysis – you will not gain any additional information*. To understand this, notice that the statistical χ^2 function, from which the fitting errors are derived, does not depend on k^w . Indeed, noise standing in its denominator is also multiplied by the weighting function, together with the k^w weighting of the data and the model, so that k^w cancels. The fact that the fitting errors are independent of k^w tells us that we cannot extract an additional information just by using the weighting function. There are many papers (I don't want to cite them here) doing fitting with various k^w -weighting which report some 'usual' errors, same for all the coordination shells, like 0.02 Å for distances and 10% for coordination numbers, regardless of the fitting quality and signal strength. I would be cautious with the conclusions inferred from such k^w -weighting exercises.

The idea of reduction of the correlation between the EXAFS parameters when doing multiple k -weighting fits is very attractive. However, if you get such reduction in an EXAFS analysis program just by simple multiplying, not by using differently long (in time) data sets, then the correlations are calculated false because they do not take into account the k^w -weighting of noise. See Section 9.7 for additional discussion.

The k^w -weighting makes sense if (a) you take into account the k -weighting of noise in all the formulas of the error analysis or (b if you cannot do a) if the resulting $\chi(k) k^w$ has uniform noise. The uniform noise in $\chi(k) k^w$ means the experimental noise in the absorption coefficient must decrease as k^{-w} . Hence, the sample time must grow as k^{2w} . For k^3 weighting this means sample time $\propto k^6$ and hence if you start with $\Delta t = 0.1$ s at $k = 1$ Å⁻¹, you will measure with 500 hours per point at the end of an EXAFS spectrum.

Most frequently I measure with k^2 weighting and then apply $\times k$ in χ . Frequently I also do $\chi(k) k^2$ weighting, but I remember about the proper experimental k -dependence of noise when doing error analysis!

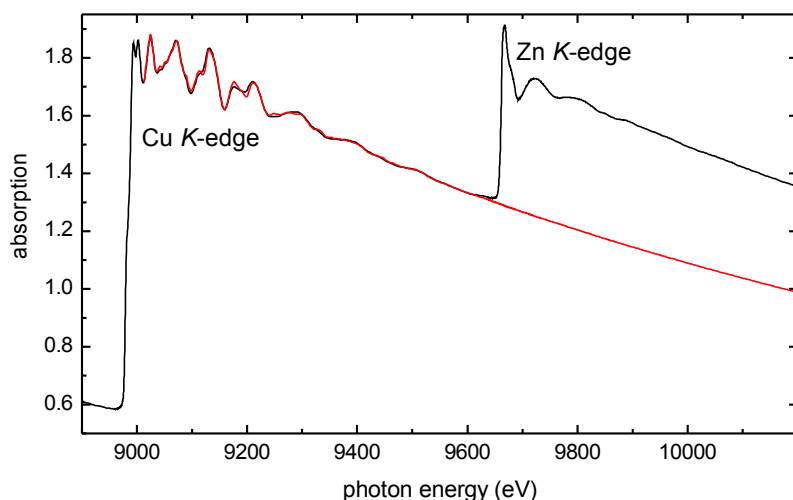
My reasoning in applying k or k^2 weighting in χ is mostly aesthetic: the FT must look good, i.e. all the FT peaks of interest should have nearly equal power.

5.9 Subtracting EXAFS due to another closely situated absorption edge

This procedure was implemented in VIPER, hoping that it should give clearer spectra in the cases of close edges. For example, I have used it in the following cases: K-edge spectra for CuZn catalysts, L_{3-} edge spectra for BaPbBiO superconductor, $L_{2,3}$ -edge spectra of rare earths. By the way, you can try to separate close absorption edges of neighboring elements by high energy resolution fluorescence detection (HERFD) but you cannot do this for $L_{2,3}$ -edge spectra of the same element.

This procedure works indeed but gives only little influence on the second spectrum. Here it is described more for curiosity than for real usage.

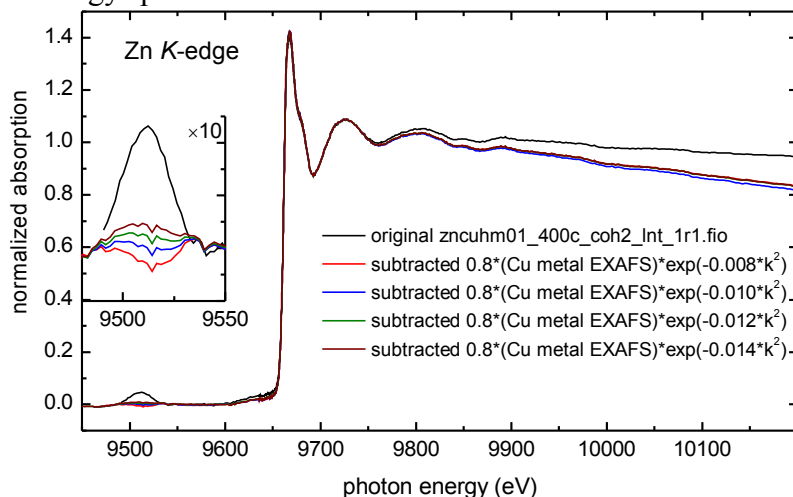
Consider an example of a CuZn catalyst (my previous work with group of Prof. Grünert, Bochum University). To subtract the Cu EXAFS from the Cu+Zn spectra I used a spectrum of Cu foil. The EXAFS of the foil is much stronger so that I had to reduce it in order to match the CuZn data before the Zn edge. Here in red is shown the Cu foil spectrum multiplied by $0.8 \cdot \exp(-0.010 \cdot k^2)$ and superimposed on (extrapolated) μ_0 curve [the multiplication was done in 'Get chi(k)' dialog using 'Corrections...' button and the superimposition was done using 'Subtract...' button]:



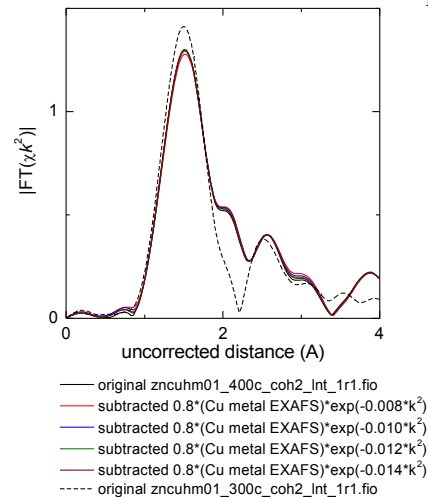
Of course, the most ideal subtraction would be given with the same catalyst but without Zn. The subtracted Cu-edge spectrum must be sufficiently long ($E_{\text{end}} > 10200 \text{ eV}$). Our other Cu spectra were not that long, therefore I subtracted the Cu foil spectrum.

Here is shown the subtraction of differently weighted EXAFS of Cu foil from Zn EXAFS:

In energy space:



... and in the Fourier-transformed space:



The subtraction story was motivated by the question: "Is the difference between the Zn EXAFS spectra for samples '300C' and '400C' (both are seen on the FT picture) due to different Cu:Zn ratio and therefore different influence of the Cu EXAFS onto the Zn EXAFS?" The answer was: "No, the Cu EXAFS is negligible above the Zn edge. It is also seen that the Debye-Waller correction to the subtracted spectrum is not important. The coordination number correction was also tried (not shown here), with similar little importance.

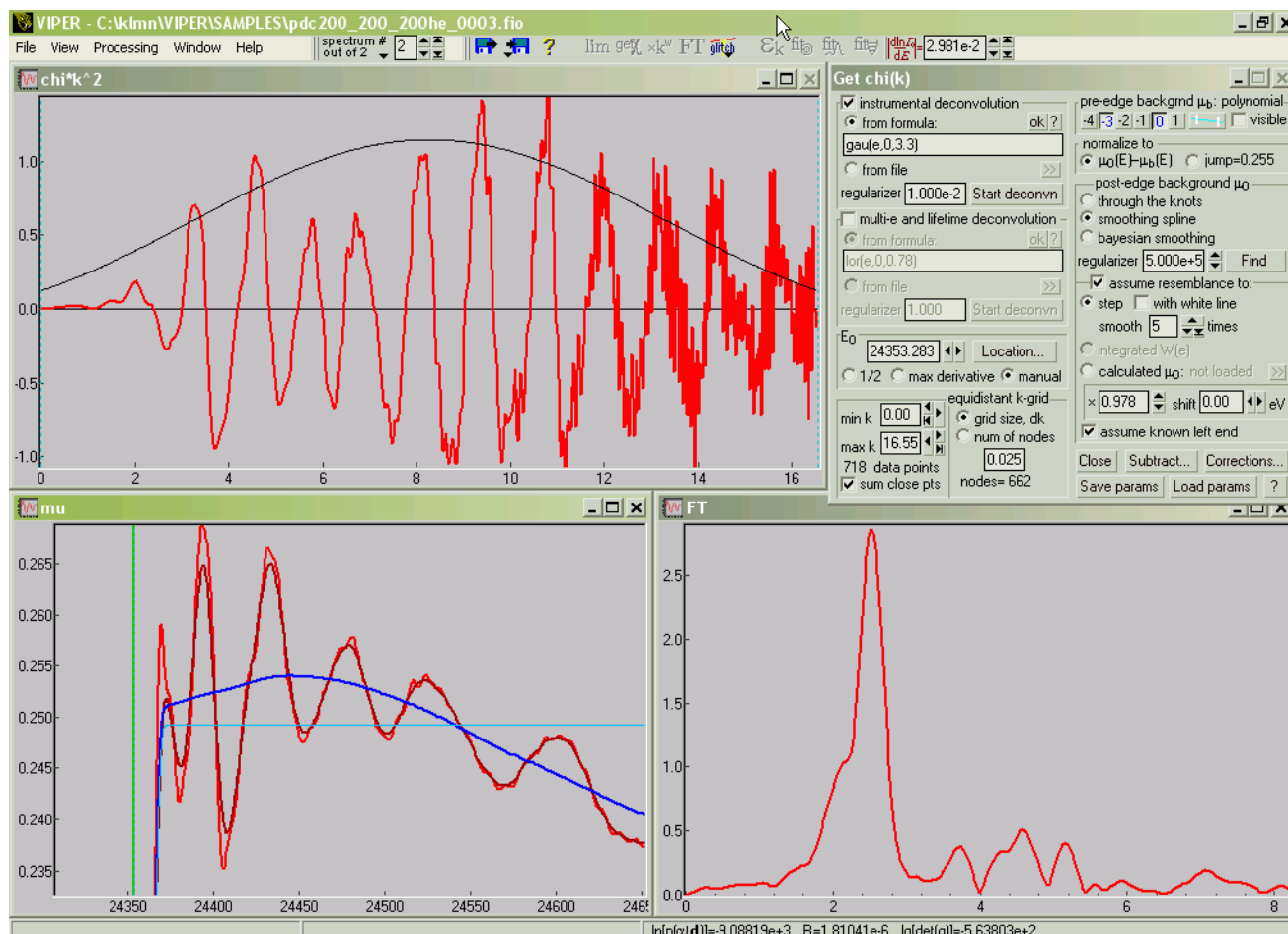
5.10 Deconvolution of life-time and experimental broadening

See [Klementev, 2001b] for the description of the Bayesian deconvolution. This procedure is implemented in VIPER and XANES dactyloscope.

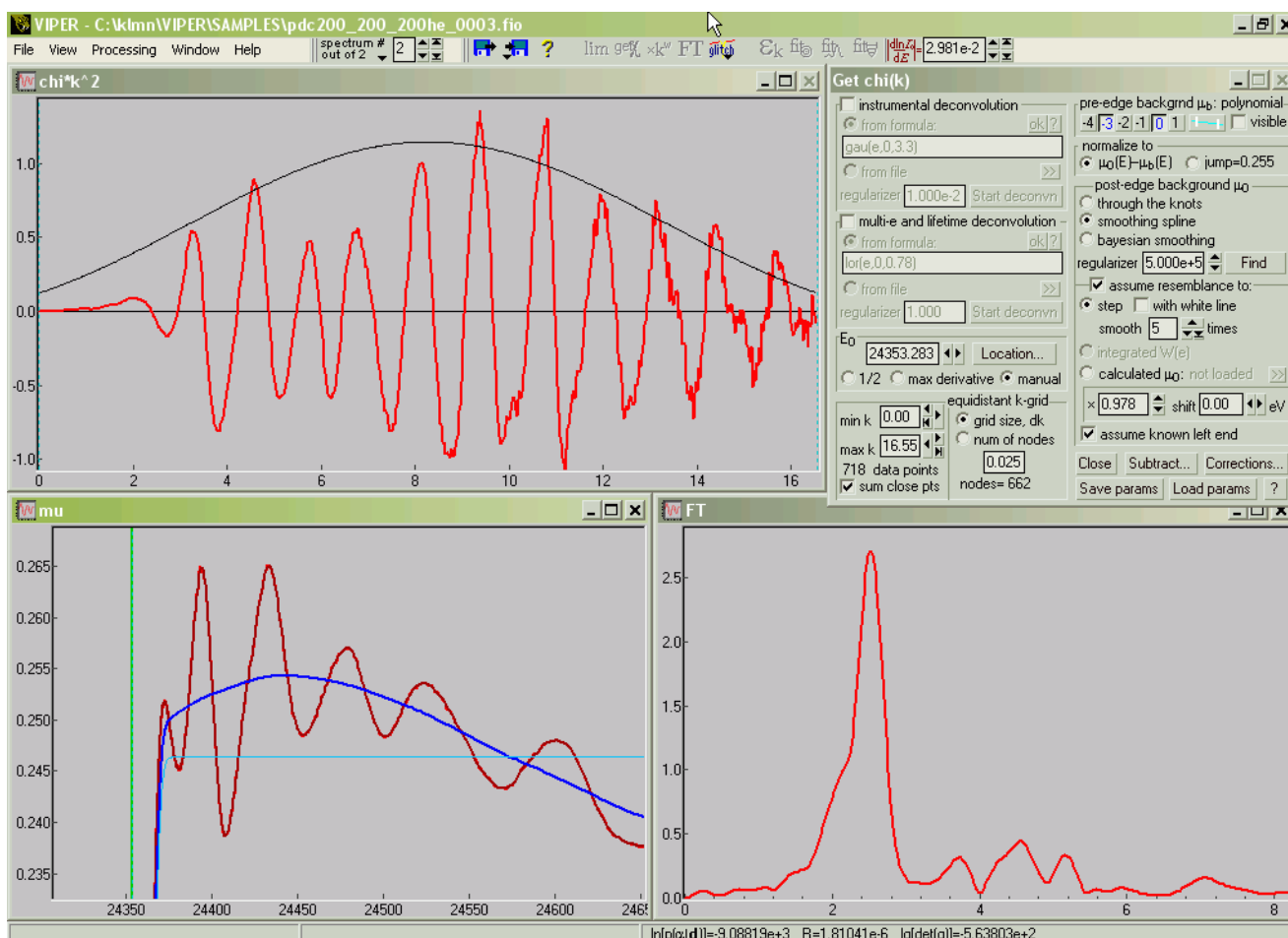
You may notice that it is rather slow. Yes, I must tell that the implementation is outdated. The deconvolution is found by solving the eigenvalue problem for an $N \times N$ matrix (N is the number of the energy mesh points) with one or a few non-zero bands, where the band width is given by the width of the broadening function. Typically, this band is of 10% of N at the absorption edge, where the mesh is dense, and $\sim 1\%$ to the end of the spectrum. Thus the matrix is sparse. The sparse algorithms scale as N^2 whereas the dense algorithms scale as N^3 . At the time of coding the Bayesian deconvolution in year 1999, Internet was at early years and I did not have as easy access to linear algebra packages, I did not have much knowledge on sparse algorithms and I took the standard dense ones. Also the dense algorithms did progress in the last years, see RRR algorithm in LAPACK. Thus the deconvolution can be made much faster than it is in VIPER but I will surely not redo it in the nearest future.

There can be *two* deconvolutions made: one is 'instrumental' and the other is 'lifetime'. The former is typically of Gaussian kernel and applied to the measured signals i_0 and i_1 separately (therefore the solution time is doubled). The latter is typically of Lorentzian kernel and is applied to $\mu(E)$.

There is a way how to check the solution: after the deconvolution has been found, the back convolution is performed by true integration and the resulting deconvolved-convolved μ (I do not know if it is better to say "deconvolved-convolved") is displayed in μ window as dark curve:



If you now unselect the deconvolution made, the initial μ and the solution check must superimpose:



The Bayesian deconvolution depends on a parameter (regularizer), denoted as α . When it is small, the solution has rich fine structure, when it is big, the solution is smooth.

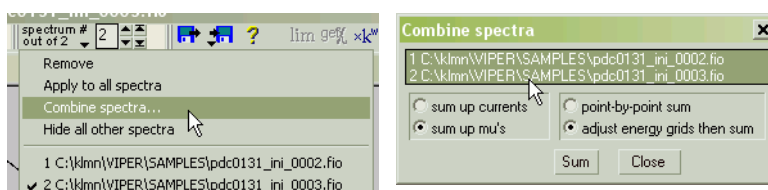
5.10.1 How to select the regularizer?

As found in [Klementev, 2001b] this is not important for EXAFS because the resulting $\chi(r)$ at $r < 8-10 \text{ \AA}$ is independent of the regularizer for a very large range of the latter. You can take it as 10^{-5} or 10 – the first Fourier peaks will be the same. However, XANES and $\chi(k)$ certainly do depend on α . One may try to define an optimal, in some sense, α . In [Klementev, 2001b] I proposed three possible ways for this. Unfortunately, what I did wrong, I did not consider the spectrum length scaling. For a full-length spectrum the optimal α must be the same as for its shorter piece. The third method does not fulfill this. It seems that the second method (the conservation of S/N ratio) is reasonable. It resembles the sharpening tools in modern post-processing of digital photography: good sharpening does not increase the visible noise while enriches the picture with fine details.

The figures of merit introduced in [Klementev, 2001b] are reported in VIPER in its status bar at the bottom. One can utilize them for (non-automatized) search for an optimum α .

5.11 Combining several spectra together

You can combine several repetitions of one spectrum or several channels of your fluorescence detector. For this, use in the Spectra menu the command 'Combine spectra...':



In the appeared dialog select the spectra to combine. If these are fluorescence channels of one EXAFS scan, they can be summed directly: use "sum up currents" and "point-by-point sum". If these are different EXAFS scans, you cannot directly accumulate the measured intensities because of strong correlations between I_0 and I_1 (or I_t). Therefore you have to normalize to I_0 individually for every scan and only then sum the scans. This corresponds to "sum up mu's". The energy grid in different

EXAFS scans may be different. In this case, point-by-point summation is incorrect; you should do "adjust energy grids and then sum".

6 Fourier analysis

6.1 Forward Fourier transform (FT)

$$FT[f(k)](r) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} W(k) f(k) e^{-2ikr - i\delta(k)} \frac{1}{B(k)} dk ,$$

where $\delta(k)$ and $B(k)$, if used, are correcting phase and amplitude. In modern EXAFS analysis, a common practice is to incorporate the phase and amplitude $\delta(k)$ and $B(k)$ into the fitting rather than into the Fourier transform. Note that in EXAFS Fourier transform the exponent is multiplied by 2.

Fast Fourier transform is realized in VIPER on the grid of $N=2^{12}=4096$ points: $dr = \frac{\pi}{(N-1)dk}$.

The following windowing functions $W(k)$ are used:

Gaussian:
$$W(k) = \exp\left(-\frac{\pi}{2} A \frac{(k - \bar{k})^2}{\bar{k}^2}\right),$$

Kaiser-Bessel:
$$W(k) = I_0\left(-\pi A \left|1 - \frac{(k - \bar{k})^2}{\bar{k}^2}\right|^{1/2}\right) / I_0(\pi A),$$

6.2 Back Fourier transform (BFT)

$$BFT[f(r)](k) = \frac{B(k)}{W(k)} \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} W'(r) f(r) e^{2ikr + \delta(k)} dr$$

The following windowing functions $W'(r)$ are used:

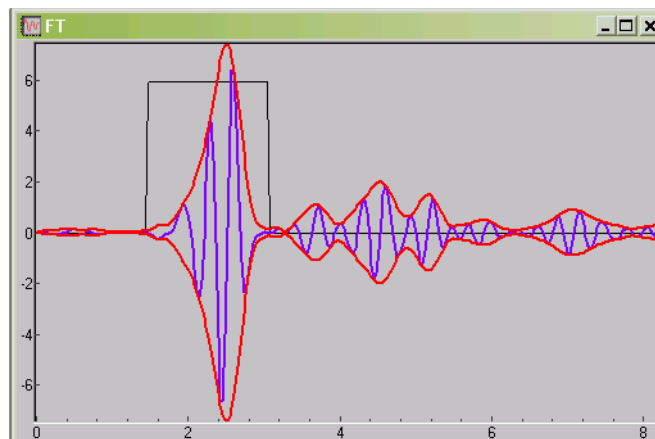
Gaussian:
$$W'(r) = \begin{cases} 1 & r_{\min} + A < r < r_{\max} - A \\ \exp\left(-10 \frac{(r - r_{\min} - A)^2}{A^2}\right) & r < r_{\min} + A \\ \exp\left(-10 \frac{(r_{\max} - r - A)^2}{A^2}\right) & r > r_{\max} - A \end{cases},$$

Hanning:
$$W'(r) = \begin{cases} 1 & r_{\min} + A < r < r_{\max} - A \\ 0.5 \left(1 - \cos\left(\pi \frac{(r - r_{\min})}{A}\right)\right) & r < r_{\min} + A \\ 0.5 \left(1 - \cos\left(\pi \frac{(r - r_{\max})}{A}\right)\right) & r > r_{\max} - A \end{cases},$$

Rectangular:
$$W'(r) = \begin{cases} 1 & r_{\min} < r < r_{\max} \\ 0 & r < r_{\min} \\ 0 & r > r_{\max} \end{cases}.$$

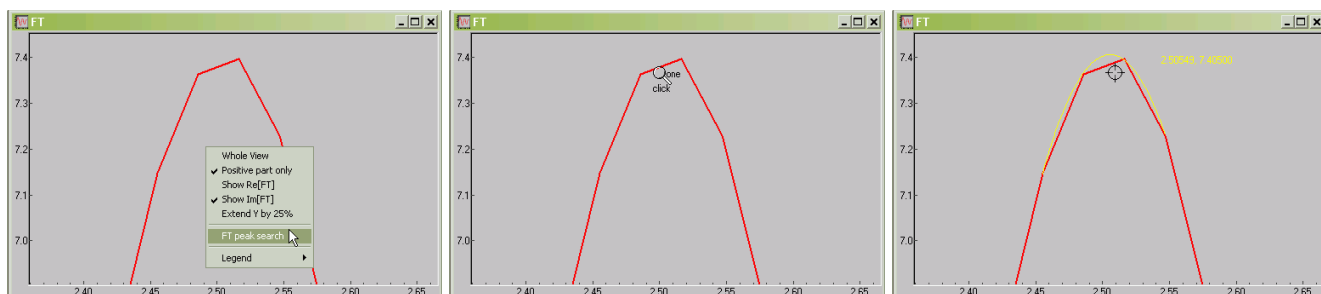
6.3 Notes on the program interface

Use the pop-up menu in the FT window to select a view: only positive part or both positive and negative, as shown here:



and to switch on and off the real and the imaginary FT parts.

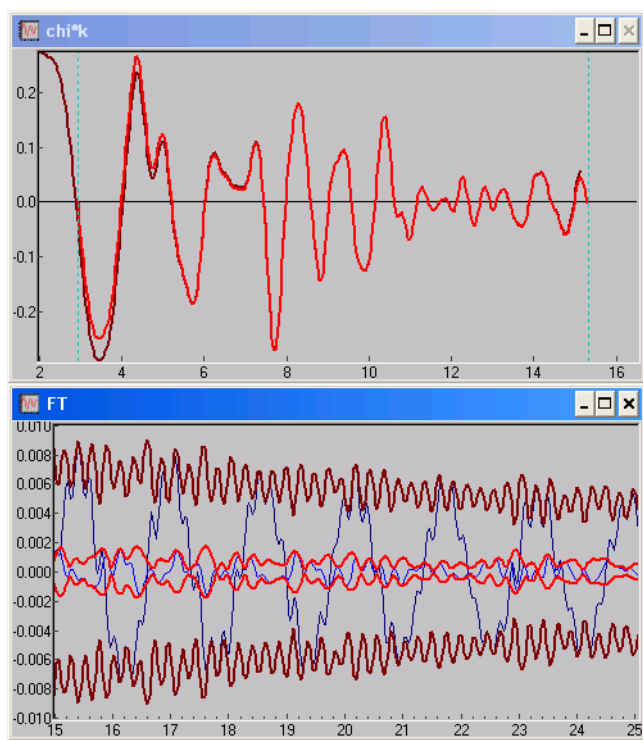
There is a possibility for peak search:



The two x and y values are also inserted into Clipboard.

6.4 Selection of k_{\min} and k_{\max}

Consider an example of the same spectrum $\chi(k) \cdot k$ with two slightly different k limits:

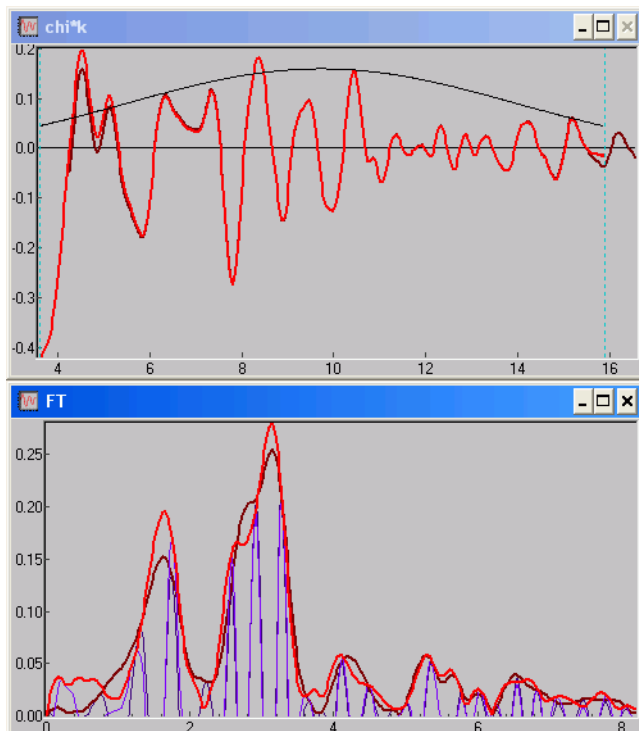


← notice the FT range here : $15 < r < 25 \text{ \AA}$.

As seen, the non-zero ends cause quite a strong FT signal in the far r -region. This region is frequently used for the estimation of noise, see Section 7.1. As seen, it can be easily overestimated by an order of magnitude. The effect of end jumps is especially strong for the k_{\min} end, and therefore is stronger pronounced for low k^w -weighting. For high k^w -weighting the k_{\min} end is effectively damped.

As a matter of exercise, from the oscillation periods in the FT one can estimate the positions in k -space. There are two oscillations in the dark FT curve: with ~ 6.3 periods over 10 \AA and of 42 periods. These correspond to the k -features at $\pi/\Delta r$, i.e. to $\sim 2 \text{ \AA}^{-1}$ and $\sim 13.5 \text{ \AA}^{-1}$. The correspondence to the end positions is not exact because of interference with other frequencies.

Also the low- r FT part is sensitive to k_{\min} and k_{\max} :



Here the two times loaded spectrum of PdO has two EXAFS functions extracted with all the parameters equal but the k -limits different (Δk range is preserved the same). As seen, the FT region $r < 1 \text{ \AA}$ depends strongly on k_{\min} . Keep this in mind when you read papers on EXAFS (see Section 5.5.2.1 above for some notes on EXAFS).

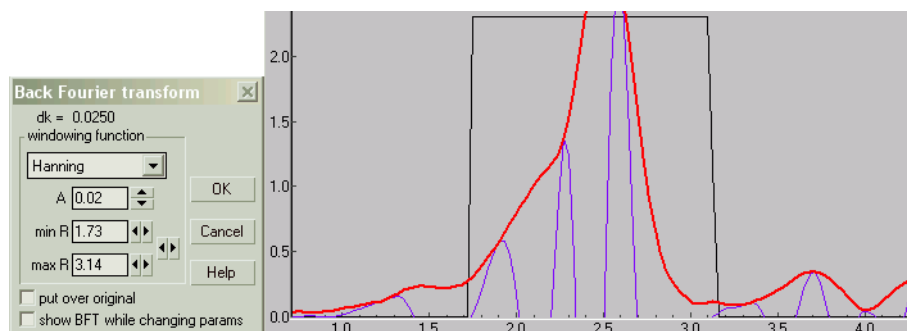
Finally, the k limits is better to chose at zeros of $\chi(k)$, even when you use a high k^w -weighting and even with a windowing FT function.

6.5 Which windowing function to use?

In the forward transform the $\chi(k) k^w$ function is padded with zeroes outside the $[k_{\min}, k_{\max}]$ interval. If at k_{\min} or k_{\max} the function $\chi(k) k^w$ has a non-zero value, the thus introduced jump gives a fringe structure in r -space. One can easily recognize the fringes as weak equidistant peaks on $FT(r)$ with a typical period of 0.2 \AA . In order to damp the fringes, a dome-like window is recommended, e.g. Gaussian or Kaiser-Bessel, constructed in a way that its value at the ends was not lower than 0.1. Lower ends would increase the computation errors in BFT due to the division by the forward window.

In the back transform I simply use the rectangular window, as it is easy to operate by mouse.

There is an opinion I have heard several times that the BFT window must be selected with its borders placed at zeros of the imaginary part:

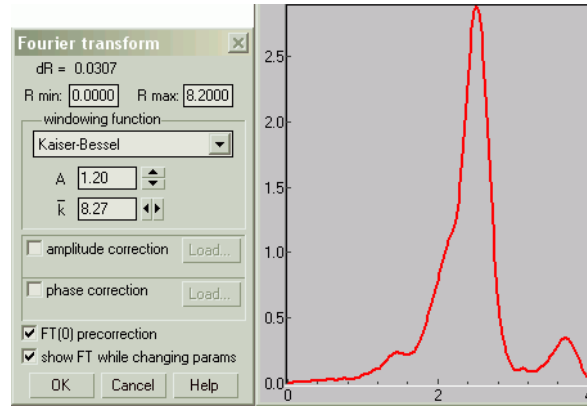
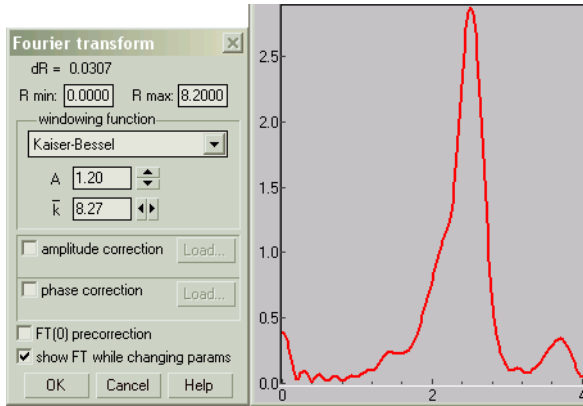


I find this strange: the real and imaginary parts are equal in rights. Why not zeros at the real part? Visualize them both. When the magnitude is zero they are both zero. This is a clear case. But what if the magnitude does not go to zero in between the FT peaks? I would simply select the minimum position, regardless of whether the $\text{Im}(\text{FT})$ is zero there.

6.6 FT pre-correction

$\text{FT}[f(k)](0)$ pre-correction can be done (and is done by default) before the transform, $f(k)$ is vertically shifted so that

$$\text{FT}[f(k)](0) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} W(k) f(k) dk = 0.$$



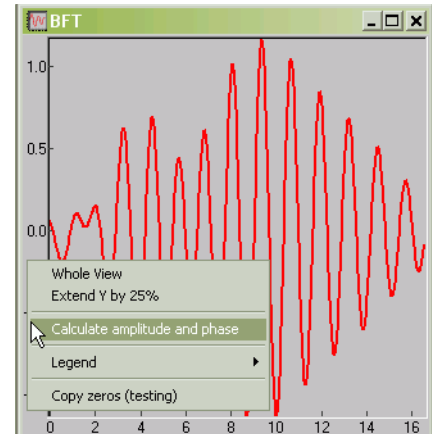
6.7 Extraction of amplitudes and phases

When BFT is done, the amplitude and phase can be calculated:

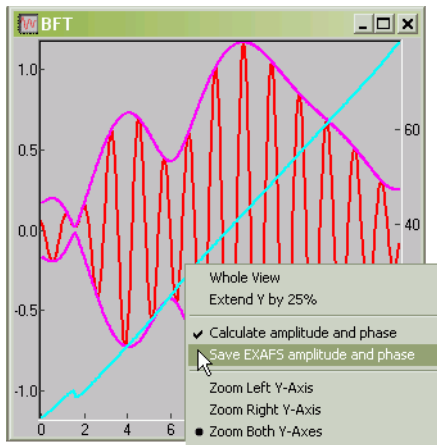
$$\text{Amp}[f(k)] = \{\text{Re}(\text{BFT}[f(r)](k))^2 + \text{Im}(\text{BFT}[f(r)](k))^2\}^{1/2},$$

$$\text{Ph}[f(k)] = \arctan\{\text{Im}(\text{BFT}[f(r)](k)) / \text{Re}(\text{BFT}[f(r)](k))\} + \pi/2.$$

The latter is calculated within a multiple of π . The additive $n\pi$ is calculated to make the phase continuous.



Now the EXAFS amplitude and phase can be saved:



The EXAFS amplitude and phase are related with the FT amplitude and phase as:

$$F(k) = \text{Amp}[\chi(k)k^n] \cdot R_0^2 k^{1-n} \exp(2k^2\sigma^2) / N,$$

$$\phi(k) = \text{Ph}[\chi(k)k^n] - 2kR_0.$$

The values of R_0 , σ^2 and N can be chosen in the Save dialog.

The experimental EXAFS amplitudes and phases are rarely used because the corresponding FT peak (a) must be well separated, which happens not very often and (b) must have no multiple scattering contributions, which can be only for the 1st shell.

7 Experimental errors in EXAFS curve

Many EXAFS people do not calculate fitting errors and in their publications they only report some 'usual' errors, same for all the coordination shells, like 0.02 Å for distances and 10% for coordination numbers, regardless of the fitting quality and signal strength. Therefore I feel that the discussion on which method to take to estimate the errors in $\chi(k)$ is not much interesting. This Section is for those who, like me, find it important. The others, please go to the next Section 8.

The errors in $\chi(k)$ are important in two respects:

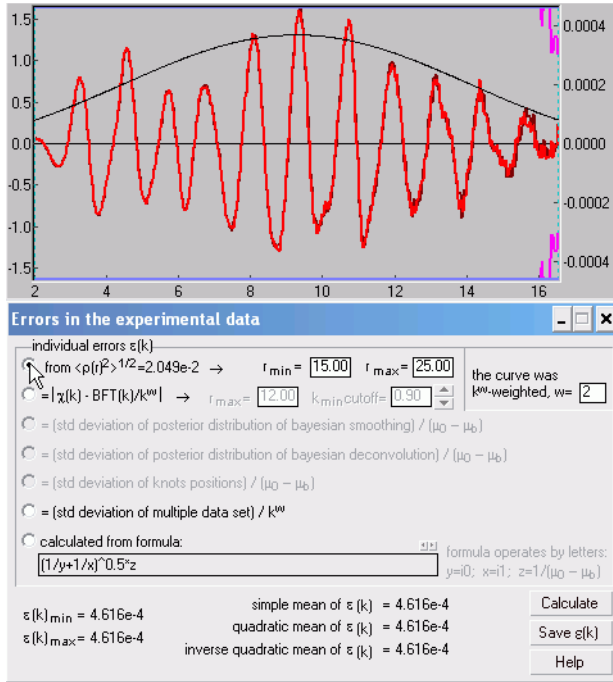
- 1) They are inside of χ^2 statistics, and the latter is a starting point for all estimations of the fitting errors.
- 2) They determine the useful length of EXAFS spectra. This circumstance is most frequently hidden. To clarify, let us consider the following. Assume we have an EXAFS spectrum of length Δk and we analyze an FT range Δr . Then the number of independent parameters N_{ind} is known (see Section 9.1). Now I double the Δk range by padding zeroes or by padding noise (which happens when I measure the added range extremely fast). Will I double N_{ind} ? Surely not... The range Δk is not just the full spectrum regardless of its quality. It is where the signal is stronger than noise.

VIPER offers several ways of how to estimate the errors in $\chi(k)$. Below, each one is described and at the end a general comparison is given.

You can start the dialog of errors by activating the χ window and pressing the ε_k button.

The noise refers to the appeared right Y-axis. In the same scale is shown the amplitude of the *non-weighted* $\chi(k)$, which allows you to visually compare the noise and the signal.

7.1 Using the high- r portion of $\chi(r)$



Use with caution!

This method was proposed by Newville et al. [1999]. The starting point is the assumption that FT of $\chi(k) \cdot k^w$ at long distances does not have any structural signal and is only due to noise ε_r . Then a high- r portion of power spectrum $|\chi(r)|^2$ is replaced by its averaged value σ_r^2 , which is equivalent to saying that the k -noise ε_k is white. ε_k and ε_r are related by Parseval's equality:

$$\int_{k_{\min}}^{k_{\max}} |\varepsilon_k k^w|^2 dk = 2 \int_0^{\pi/2\delta k} |\varepsilon_r|^2 dr,$$

where δk is the k -mesh step. A comment must be added here: if the noise ε_k is white, $\varepsilon_k k^w$ is also white because it remains uncorrelated and with zero mean value. Therefore the averaging of the high- r portion of $|\chi(r)|$ can be done at any k -weighting.

$|\varepsilon_k|^2$ in the lhs integral is not constant but is rapidly oscillating around its rms value $|\sigma_k|^2$. Then $|\varepsilon_k k^w|^2$ in every local averaging gives $|\sigma_k k^w|^2$. Finally,

$$\sigma_k^2 = \sigma_r^2 \frac{\pi(2w+1)}{\delta k (k_{\max}^{2w+1} - k_{\min}^{2w+1})}.$$

Newville et al. [1999] recommend averaging in the range 15 Å < r < 25 Å.

As seen from the derivation, the resulting σ_k should not depend on the k -weighting. But it frequently does, and σ_k obtained for k^3 -weighting can be an order of magnitude, or even more, smaller than for k^1 -weighting. The reason for this is the violation of the starting assumption: the high- r FT may have a very long tailed signal due to the jumps at the ends k_{\min} and k_{\max} , see Section 6.4.

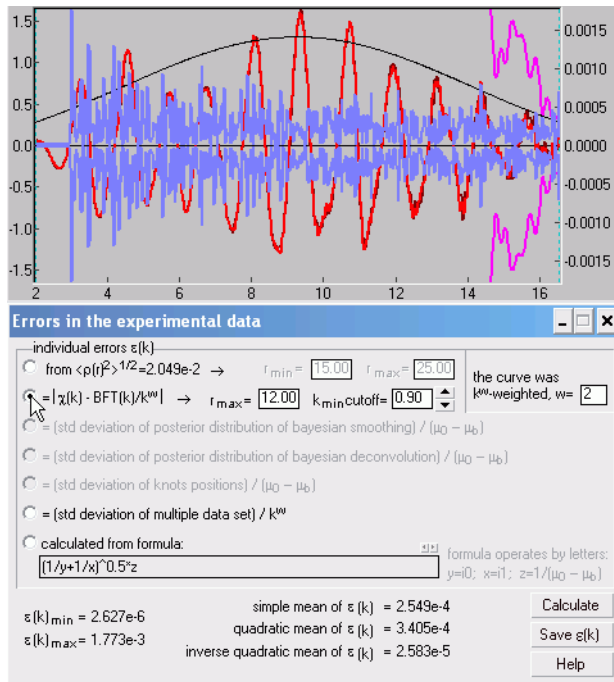
If you use this method, it is recommended to visually check that $|\chi(r)|^2$ in the high- r region varies close to zero. Alternatively, there must be no visible low-frequency oscillations in $\text{Re}(\chi(r))$ or $\text{Im}(\chi(r))$ – visualize them. Additionally, you should check that different k -weightings give close results for σ_k .

Note that the k -noise obtained by this method is *displayed* in VIPER as constant but it is *not* constant. This is a zero-mean white noise with the given variance σ_k^2 . The actual distribution law for its amplitude is unknown; it can be normal, Poissonian or another. Therefore it is displayed as constant σ_k .

[A note for future development] The high- r portion of $\chi(r)$ can be used for alarming that the limits k_{\min} and k_{\max} are bad because they introduce FT fringes: the variation of ε_r^2 around its middle σ_r^2 should not be (much) smaller than σ_r^2 .

Notice the reported value of σ_r (shown in the screenshot as $\langle \rho(r)^2 \rangle^{1/2} = \text{'value'}$). This can be used for error analysis in r -space fitting.

7.2 Using FT filtering

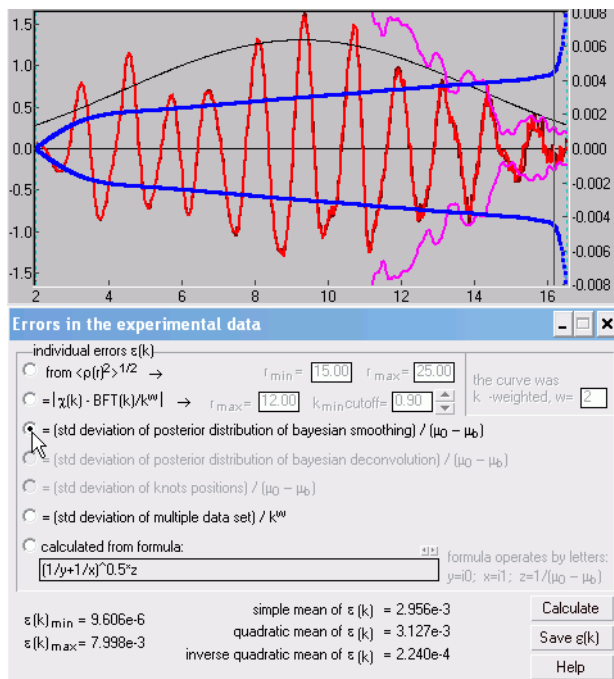


Use with caution!

This method is similar in idea that the high- r portion of $\text{FT}[\chi(k) \cdot k^w]$ is due to noise. The noise is obtained as a difference between the non-weighted $\chi(k)$ and BFT divided by k^w . The back FT has errors near the ends; because of the division by k^w the k_{\min} end has an increased level of noise and should be cut-off.

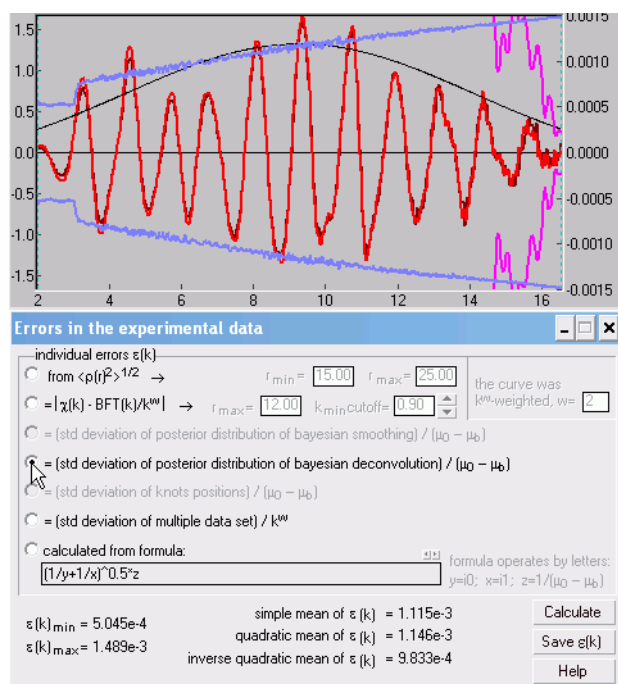
Wrong settings for k_{\min} and k_{\max} can affect in the obtained noise in the same way as described above. You should check that different k -weightings give close results for σ_k .

7.3 Using μ_0 obtained by Bayesian smoothing



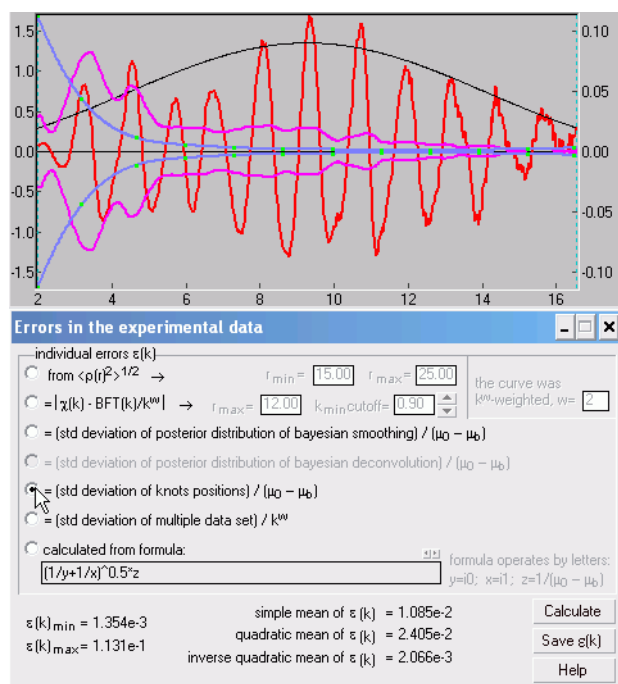
This option is enabled if the Bayesian smoothing curve has been calculated. The single idea of using the Bayesian smoothing method is the capability of getting the uncertainties. Try to do the Bayesian smoothing with and without the a priori information on the behavior of μ_0 : "assume resemblance to step" and "assume known left end". Watch how the uncertainties behave at the left end of $\chi(k)$.

7.4 Using Bayesian deconvolution



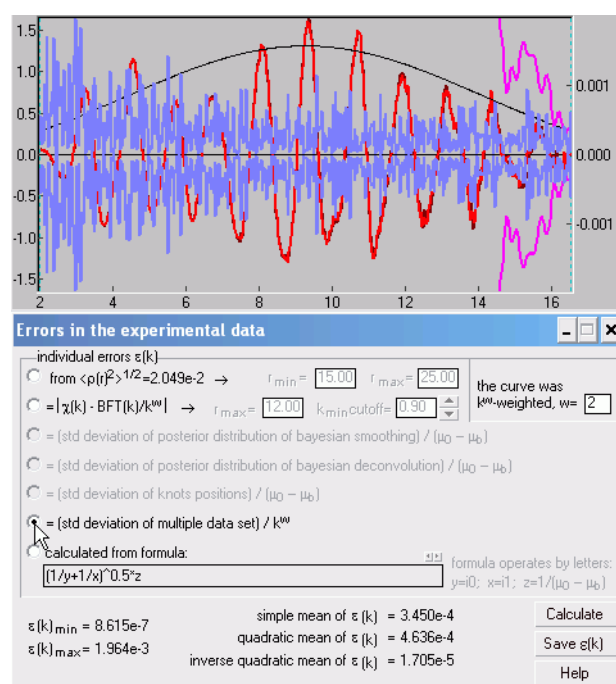
This option is enabled if the Bayesian deconvolution has been calculated. This method is a lateral usage of the implemented deconvolution routines. The jump at $\sim 3 \text{ \AA}^{-1}$ in this particular spectrum is due to a sudden increase of the energy steps specified in the spectra acquisition program; there the constant dE regime was switched to the constant dk regime. Thereafter, the linearly increased uncertainty is owing to the steadily increasing energy steps. The same increase is seen in the previous method, where the 3 \AA^{-1} jump is smoothed out because of the much stronger requirements for smoothness in that procedure.

7.5 Using μ_0 drawn through varied knots



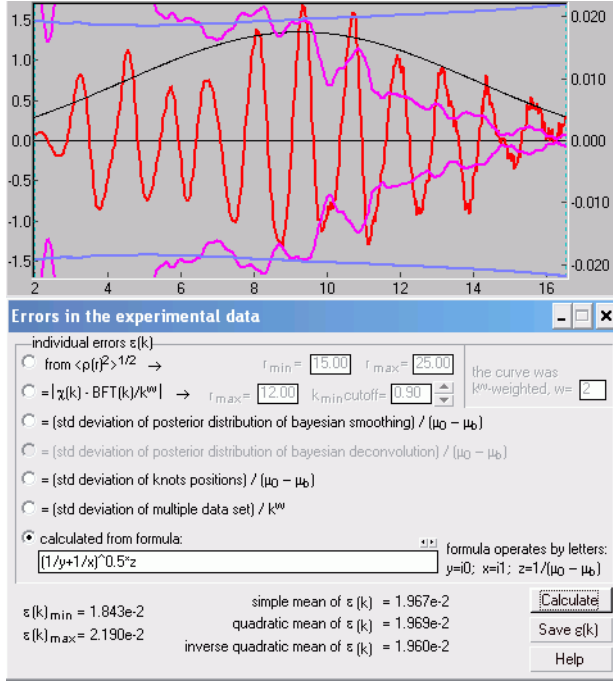
This option is enabled if the knots have been optimized and their errors have been calculated (using 'Statistics' dialog). The uncertainties of the first and the last knots usually exceed the EXAFS amplitude.

7.6 Using standard deviation of multiple data



This option is enabled when there are several spectra loaded.

7.7 Calculated from a user formula



Noise can be estimated based on Bayesian considerations applied to the detection statistics. Let the probability of a single count to occur within the time interval dt be given by $P(1|\lambda)=\lambda dt$. It can be shown [Jaynes, 1990] that merely from this assumption it follows that the counts obey the Poisson distribution law:

$$P(N|\lambda, T) = \frac{(\lambda T)^N \exp(-\lambda T)}{N!},$$

where T is the sampling time. The problem is to find the intensity λ and its variance. Using Bayes theorem and introducing prior probabilities $P(N)=1/N$ and $P(\lambda)=1/\lambda$ [Jeffreys, 1939], one obtains:

$$P(\lambda|N, T) = \frac{P(N|\lambda, T)P(\lambda)}{P(N)} = \frac{T(\lambda T)^{N-1} \exp(-\lambda T)}{(N-1)!}$$

that is after measurement the variate $2T\lambda$ follows the χ^2 -distribution with $2N$ degrees of freedom. Hence $\langle\lambda\rangle=N/T$, $\langle\lambda^2\rangle=N(N+1)/T^2$, and $\delta\lambda=N^{1/2}/T$.

Denote the counts from the detectors measuring i_0 and i_1 as I_0 and I_1 . By definition, the variate

$$\xi = \frac{i_1 / 2I_1}{i_0 / 2I_0}$$

follows the Fisher-Snedecor F-distribution with $(2I_1, 2I_0)$ degrees of freedom. Its expected value and variance are: $\langle\xi\rangle = I_0/(I_0-1)$, $\delta^2\xi=I_0^2(I_0+I_1-1)/[(I_0-1)^2(I_0-2)I_1]$, from where for the absorption in the fluorescence mode ($\mu x=i_1/i_0$):

$$\langle i_1 / i_0 \rangle = \frac{I_1}{I_0 - 1}, \quad \delta^2(i_1 / i_0) = \frac{I_1(I_0 + I_1 - 1)}{(I_0 - 1)^2(I_0 - 2)}.$$

Further, the variate $\eta=\frac{1}{2}\ln\xi$ follows the Fisher's z-distribution with $(2I_1, 2I_0)$ degrees of freedom. Its expected value and variance are: $\langle\eta\rangle=0$, $\delta^2\eta=\frac{1}{4}(I_0+I_1)/(I_0I_1)$, from where for the absorption in the transmission mode ($\mu x=\ln(i_0/i_1)$):

$$\langle \ln(i_0 / i_1) \rangle = \ln \frac{I_0}{I_1}, \quad \delta^2 \ln(i_0 / i_1) = \frac{1}{I_0} + \frac{1}{I_1}$$

The noise of the transmission EXAFS-function is

$$\epsilon_k = \frac{\delta\mu}{\mu_0 - \mu_b} = \left(\frac{1}{I_0} + \frac{1}{I_1} \right)^{1/2} \frac{1}{\mu_0 - \mu_b}$$

This formula can be used in VIPER as shown in the screenshot.

Note again that all the above formulas are solely based on the assumption of the probability of a single count $P(1|\lambda)=\lambda dt$. In practice this condition is realized as: $P(c|\lambda)=\lambda dt$, that is the amplification path works in such a way that one photon gives birth to c measured counts. This circumstance changes the degrees of freedom in the above distributions from $2N$ to $2N/c$; the measured counts need substitution: $I_{0,1}$ to $I_{0,1}/c$.

[For future versions of the manual: give an example of how to estimate c]

7.8 Comparison of different estimations of experimental errors

There are two major contributions to the experimental errors in $\chi(k)$: those due to (i) measurements and (ii) uncertainty in μ_0 . The former one does not only include noise at every individual point but also must depend on the *density* of the measured points. Indeed, if the density is growing, there are more points falling into a single dk interval, thus the $\chi(k)$ points must get less noise. The methods which cal-

culate the errors by taking the data from the equidistant k -space are obviously insensitive to the density of points in E -space. Such methods are the two Fourier methods 7.1 and 7.2. The multiple data averaging 7.6 is done point-by-point and also only sensitive to amplitude noise.

The standard deviation of the knot positions 7.5, in contrast, only represents errors in μ_0 but not the errors due to measurement statistics.

The Bayesian methods 7.3 and 7.4 are sensitive to the data density. They also can directly incorporate the individual noise at every point.

[A note for future development] The Bayesian methods (smoothing and deconvolution) are now taking measurement noise as a global (unknown but most probable) parameter. These methods should be re-coded to give options for selecting noise as (a) known, (b) proportional to a selected function, at least given by Poissonian statistics, with the most probable coefficient of proportionality and (c) the most probable global noise, as it is now.

The Bayesian smoothing 7.3 also gives errors in μ_0 .

It appears that all the listed methods can either estimate measurement noise or uncertainty in μ_0 (the Bayesian smoothing 7.3 can, in principle, be sensitive to both but not in the present version of VIPER). But are these two errors equally important? As seen in the above screenshots, μ_0 errors are about one order of magnitude bigger. When these errors are used to determine the fitting χ^2 statistics, the latter takes usual good values. In contrary, a pure statistical noise gives much too high χ^2 values and feeds the popular in the EXAFS community discussions about magic 'systematic' errors.

Finally, I would recommend using the Bayesian smoothing once for every new set of data. After the errors have been estimated and the k -range for fitting has been settled (this is where the signal is stronger than noise), I switch to standard smoothing spline with the same smoothing parameter.

8 Fitting EXAFS

Scattering amplitudes and phases must be calculated by other programs (FEFF format is recognized) or can be extracted by VIPER from a reference spectrum. If the photoelectron momentum l is even (absorption edges K, L₁, M₁, M₄, M₅ etc.), π must be added to the loaded phase. **Important:** You must not add π for the phases generated by FEFF because it is already there.

If the amplitudes and phases are loaded as feffNNNN.dat files, these are calculated as:

$$f(k) = \text{column}(\text{'mag[feff]'}) \cdot \exp(-2R / \text{column}(\text{'lambda'})) \cdot \text{column}(\text{'red_factor'})$$

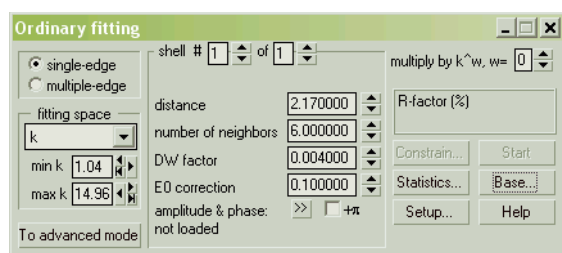
$$\phi(k) = \text{column}(\text{'real[2phc]'}) + \text{column}(\text{'phase[feff]'})$$

The global loss factor S_0^2 is assumed constant and can be set in 'Options...' of every fitting procedure. See Section 8.7 for discussion on S_0^2 .

8.1 Fitting by ordinary EXAFS formula

EXAFS-function is treated as a sum over coordination spheres (shells):

$$\chi(k) = \sum_j \chi_j(k) = S_0^2 \sum_j \frac{N_j}{kR_j^2} |f_j(k)| \sin(2kR_j + \phi_j(k)) \exp(-2\sigma_j^2 k^2)$$



The varied parameters are: N_j – coordination number, R_j – radius of the j -th sphere (Å), σ_j^2 – distance variance (Å²), $(\Delta E_0)_j$ – energy shift of E_0 (eV). Scattering amplitude $f_j(k)$ and phase $\phi_j(k)$ must be loaded from an outer file.

Parameters of a coordination sphere can be fixed or related by equalities or inequalities with the same parameters (n with n , r with r etc.) of the other spheres. For example:

$r_2 > 4.3 - r_1$ distances;
 $s_2 = s_1$ distance variances;
 $n_3 = 12 - n_1 - n_2$ coordination numbers;
 $n_1 = 6$ coordination number;
 $s_1 \ln 0.001..0.02$ distance variance, means $0.001 < s_1 < 0.02$
 $e_2 \ln(e_1 - 1) .. (e_1 + 1)$ energy shifts, means $|e_2 - e_1| < 1$

Each expression can use (almost) any function of variables: $r_1, n_1, s_1, e_1, r_2, \dots$. Put one expression per line.

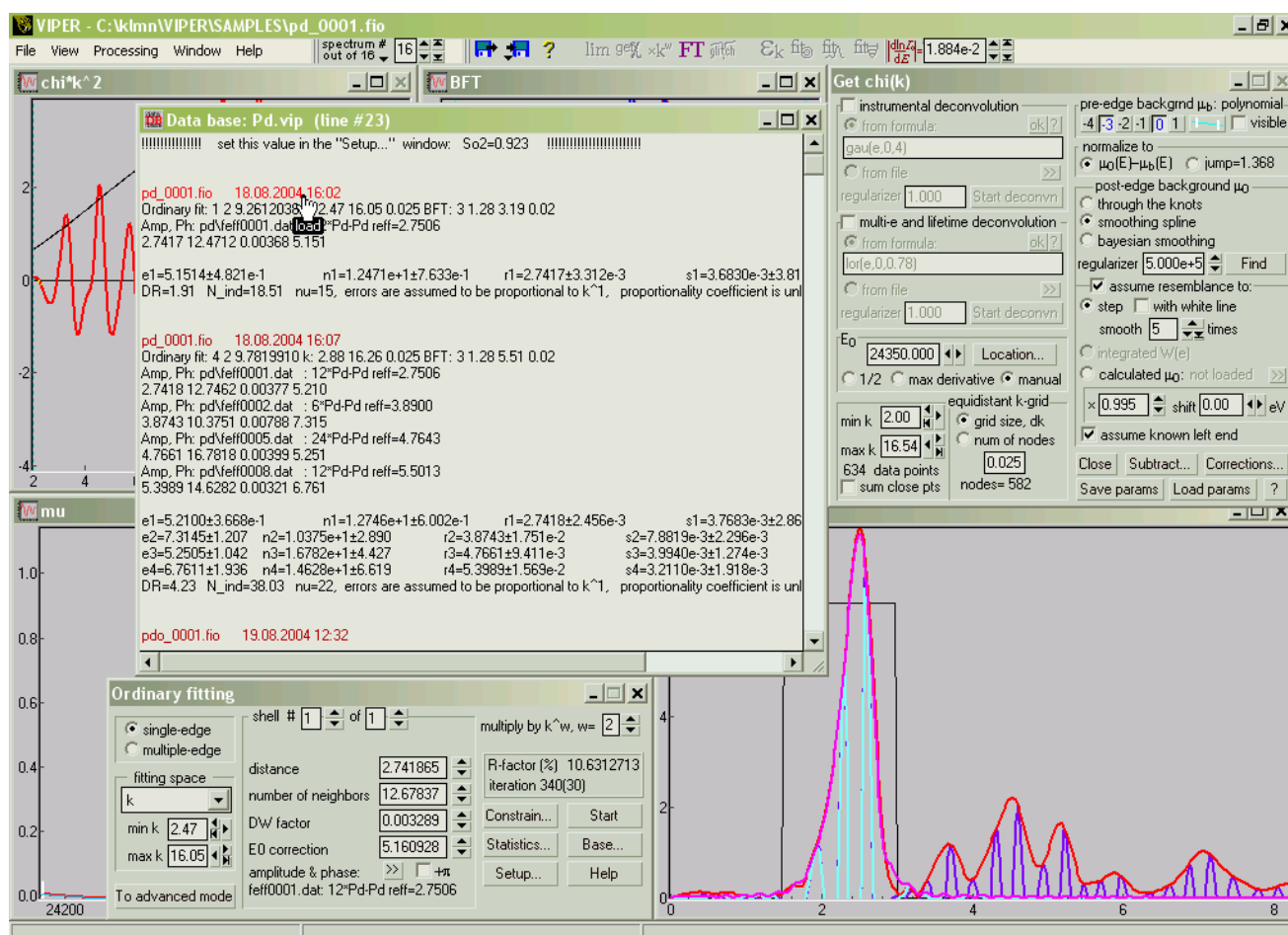
8.1.1 How about multiple scattering (MS) fitting? Can VIPER do it?

Yes, it can. FEFF calculates the effective amplitudes in the way that $\chi(k)$ is expressed in terms of the same standard EXAFS formula. If you use FEFF amplitudes, it does not matter whether a scattering path is of single or multiple scattering type, you just specify the corresponding feffNNNN.dat file.

8.1.2 Why the energy shift is not global but is different for different shells?

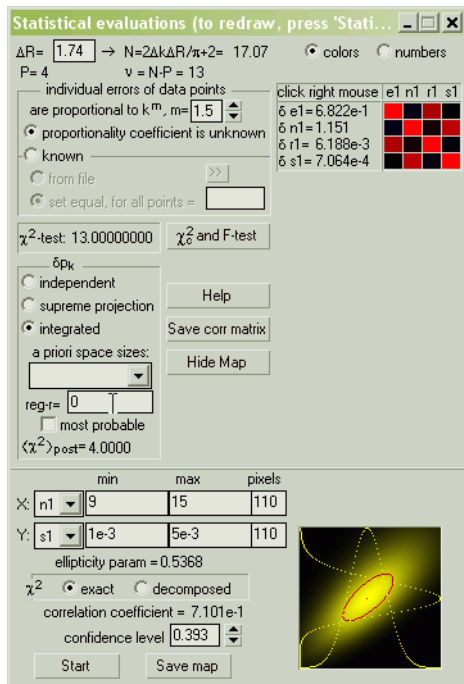
Indeed, uncertainty in the E_0 positioning (see Section 5.3) is global and is common for all the coordination shells. However, the *calculated* scattering amplitudes and phases may have shifts in Fermi energy. Moreover, the shifts may be different for different types of atoms. However, this difference should not be big: it is typically about 1 eV with self-consistent calculations and about 3 eV with overlapped atom potentials [FEFF8.10 manual]. Make sure that the E_0 shifts obtained are not much different from each other. The energy shifts for the same atoms in different coordination shells should be constrained as equal (like 'e2=e1').

8.1.3 How to load a model and save the fitting results?



Load the project 'Samples/PdCOLd.vpj'. Select in the Spectra menu the command 'Hide all other spectra'. Set focus to the BFT window (click on it) and press 'fit' button. Press 'Base...' button in the fitting window and open the 'Samples/Pd.vip' data base. Now click to the header of the first entry. Under 'Setup...' button, set the global S_0^2 parameter equal to 0.923, as given by FEFF for Pd. Play with the

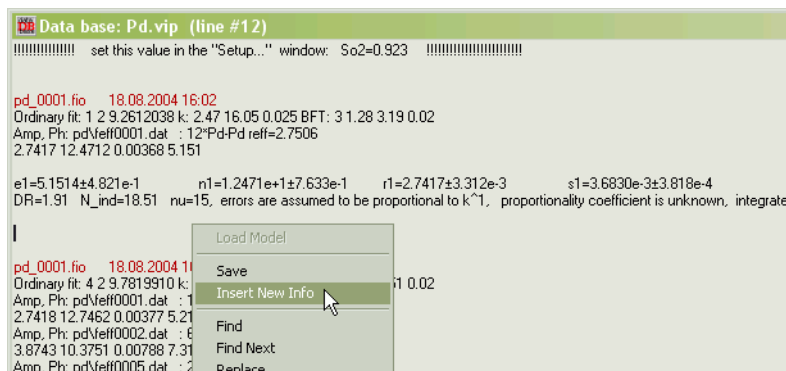
BFT windowing function and with the fitting parameters. Start the automatic optimization. Stop/Start it several times; in this way it goes faster. When you see that the R -factor decreases very slowly, stop it. Invoke 'Statistics' dialog. This dialog and the underlying procedures will be described in Section 9.8.



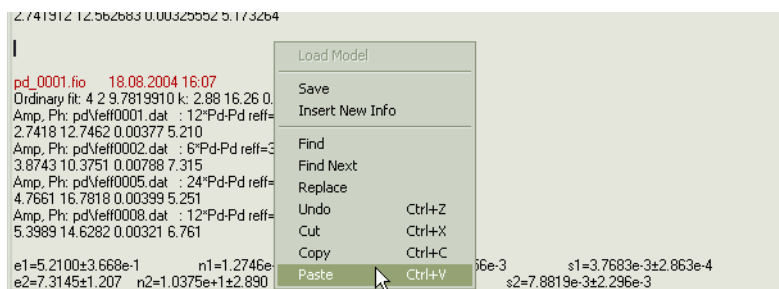
Now set $m=1.5$ for the noise k^m -weighting (why? – will be explained in Section 9.7). Right-click on the panel with the correlation coefficients and copy the resulting fitting errors.



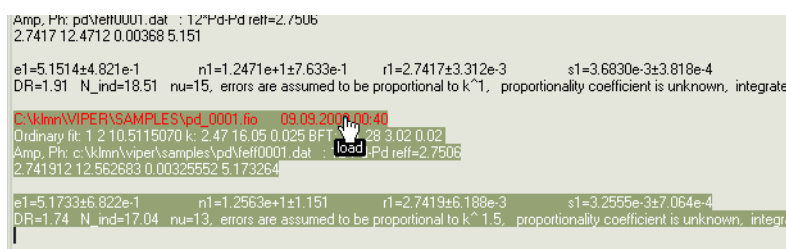
Close the Statistics dialog and open the data base again. If you want to create a new data base, close the opened one and in the 'Open Base' dialog give a new name. Prepare a place where you want to insert a new model description (insert empty lines and put the cursor there) and by right click do 'Insert New Info':



Then paste the copied fitting errors:



Try then to load the inserted model:



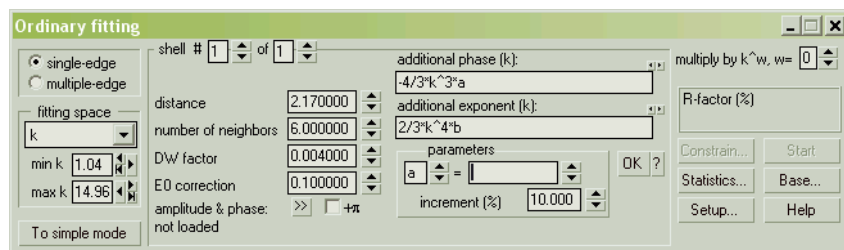
Save the file via right-click menu. The data base is a text file. You can edit it by any common editor. The supplied examples have local path references to the amplitude&phase files in the sub-directories of Sample directory. A new data base entry has full path references. If you copy the data base together with the amplitude&phase files to another computer, you should manually change the paths accordingly.

8.1.4 How to create a new model?

Specify amplitude & phase for the first shell. As soon as this is done, the model becomes visible. Now one can add the next shell. When a model has more than one shell, one has the possibility to leave only one (current) shell visible in order to see how strongly it contributes.

8.2 Fitting by a user-expanded EXAFS formula (cumulant expansion)

Toggle to the advanced mode by pressing the button 'To advanced mode/To simple mode'.



The cumulants [Rehr & Albers, 2000, Section IV.E] can be added by user-defined formulas, up to a desired order. The added variables can be any except r , n , s , e (which are already in use) and k (independent variable).

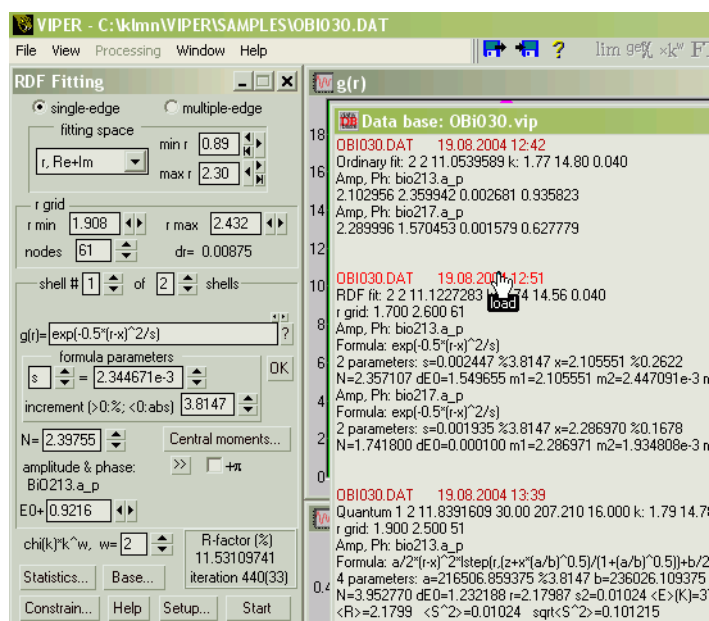
8.3 Fitting using radial distribution function specified by user-defined formula

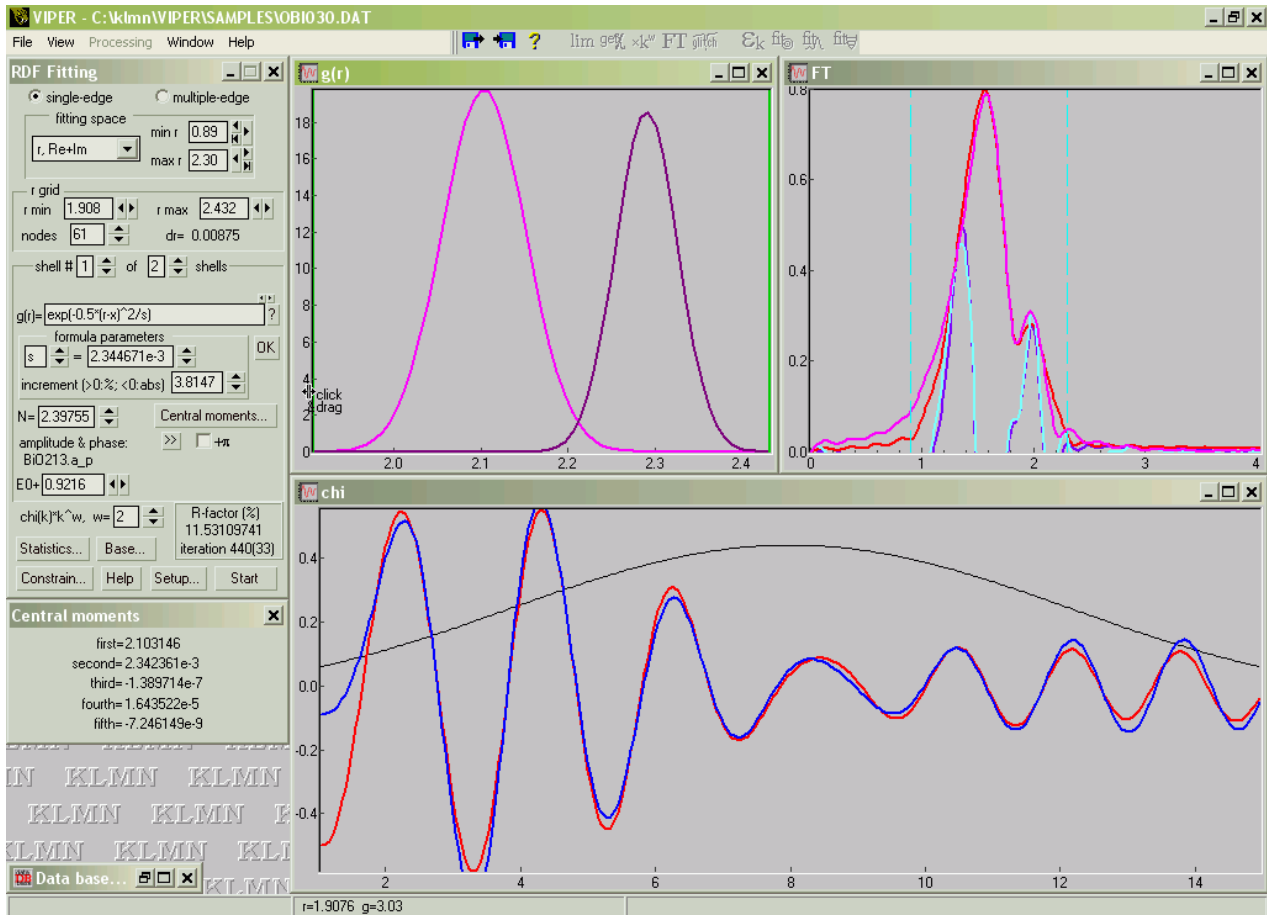
EXAFS-function is treated as a sum over coordination spheres (shells):

$$\chi(k) = \sum_j \chi_j(k) = S_0 \sum_j \frac{|f_j(k)|}{k} \int_{r_{\min}}^{r_{\max}} g_j(r) \sin(2kr + \phi_j(k)) / r^2 dr, \quad \int_{r_{\min}}^{r_{\max}} g_j(r) dr = N_j$$

$g_j(r)$ may be of arbitrary shape and is given by a user-defined formula. The character parameters of the formula from "a" to "z" are varied (r , n , and e are reserved). Besides these parameters, the varied parameters are: N – coordination number, ΔE – energy shift of E_0 (eV). Scattering amplitude $f_j(k)$ and phase $\phi_j(k)$ must be loaded from an outer file.

Close the previous example: press F8 and close the fitting dialog. Load the project file 'Samples/fitting.vpj' and press the second 'fit' button. Open the data base ('Base...' button) 'Samples/OBi030.vip' and load the second entry:





8.4 Fitting using oscillatory potential $U(r)$ of the absorber-scatterer pair

EXAFS-function is treated as:

$$\chi(k) = S_0^2 \frac{|f(k)|}{k} \int_{r_{\min}}^{r_{\max}} g(r) \sin(2kr + \phi(k)) / r^2 dr,$$

where the atomic radial distribution function $g(r)$ is calculated as:

$$g(r) = N \frac{\sum_n |\Psi_n(r)|^2 e^{-\varepsilon_n / kT}}{\sum_n e^{-\varepsilon_n / kT}} \quad (\text{in quantum approach}),$$

$$g(r) = N \frac{e^{-U(r)/kT}}{\int e^{-U(r')/kT} dr'} \quad (\text{in classical approach}),$$

where N is coordination number.

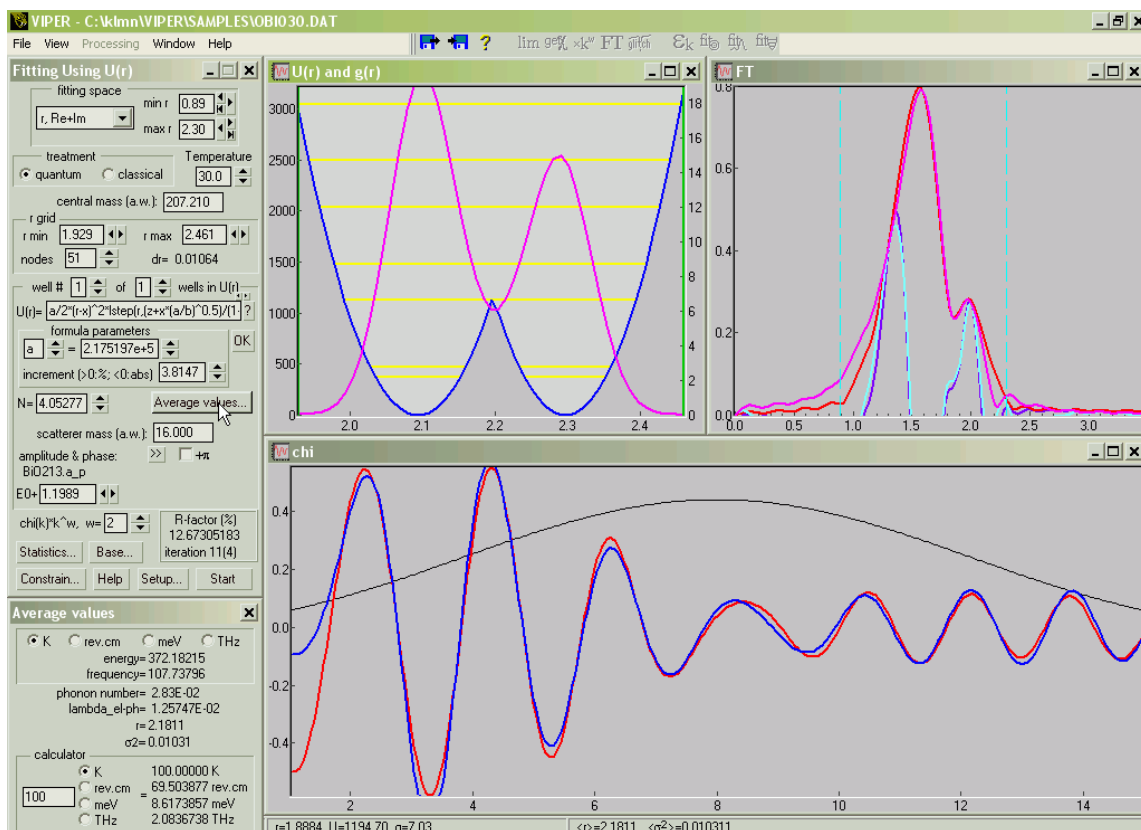
Potential $U(r)$ depending on the distance r between absorber and scatterer atoms in which the particle of the reduced mass of these atoms oscillates may be of arbitrary shape and is given by a user-defined formula. The character parameters of the formula from "a" to "z" are varied, "r" is independent variable, "n" and "e" are reserved.

Examples:

$U(r) = a/2 * (r-x)^2 * \text{step}(r, (z+x*(a/b)^{0.5}) / (1+(a/b)^{0.5})) +$
 $b/2 * (r-z)^2 * \text{rstep}(r, (z+x*(a/b)^{0.5}) / (1+(a/b)^{0.5}))$ – double-well parabolic potential,
 $U(r) = a * (1 - \exp(-b * (r-z)))^2$ – Morse potential.

All energetic parameters are measured in Kelvins, distances in angstroms. It is possible to set several independent wells. Besides these parameters, the varied parameters are: N – coordination number, ΔE – energy shift of E_0 (eV).

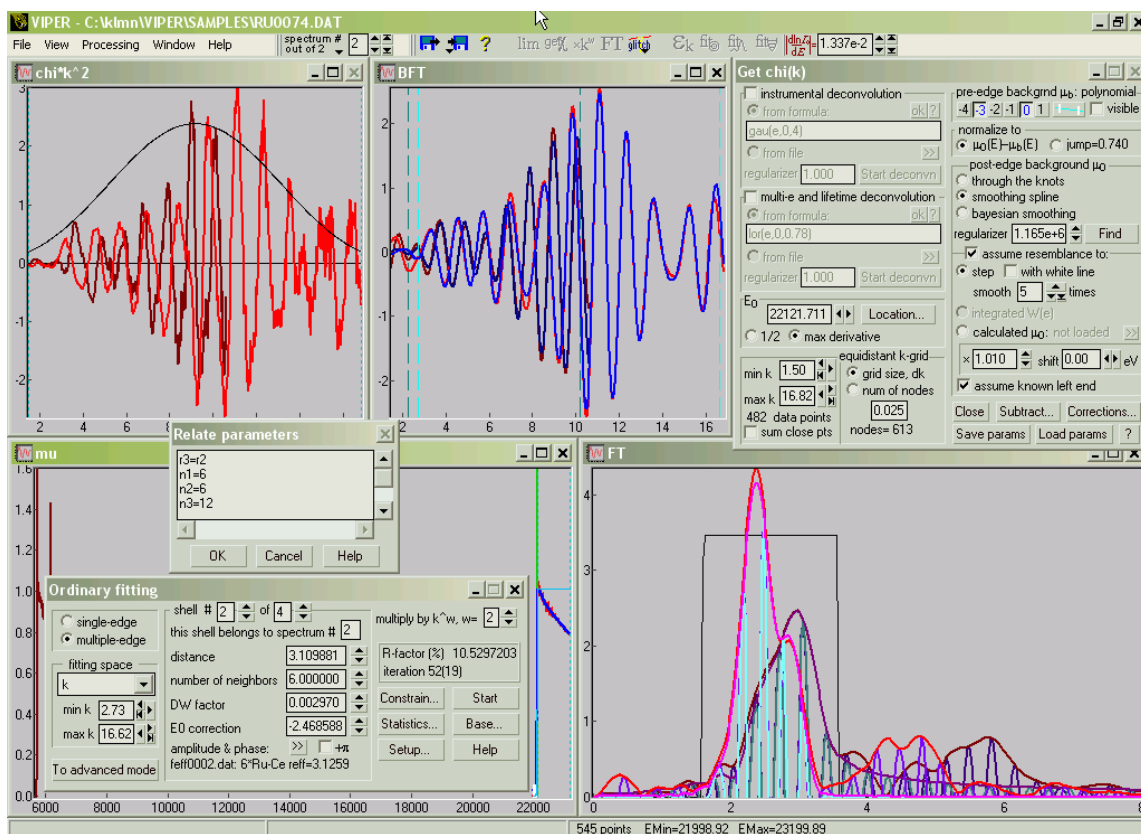
Close the previous dialog and open the 3rd 'fit' dialog. In the opened already data base load the third entry:



8.5 Multi-edge fitting

If it is possible to measure EXAFS for the same sample at several absorption edges then the distances and Debye-Waller factors as seen from the several origins must be equal. Therefore one can constrain these parameters to be equal in a multi-edge fitting procedure and thus obtain reduced fitting errors.

The example project Sample/CeRu2.vpj has two spectra, at L_3 -Ce and K-Ru, for the same sample CeRu₂. In the multiple-edge fitting, one can equalize the Ce-Ru and Ru-Ce distances:



Multi-edge fitting is possible in all the above fitting procedures except the last one.

8.6 Fitting in k - and r -space

All the fitting procedures in VIPER are possible to perform in k - or r -space. What is better? Fitting in " k -space" here means in "Fourier filtered k -space" because to include all the scattering atoms up to very distant ones that still contribute stronger than noise is impossible. This would require too many fitting parameters while this number is limited, see Section 9.1.

If you superimpose BFT over its original χ (use the corresponding option in the BFT dialog), you can see that the ends of the BFT (typically 0.1–0.4-Å⁻¹-long) are somewhat distorted. Therefore if you fit in the filtered k -space, you must cut off the ends and you lose some available data.

There are two major disadvantages in doing fitting in r -space. (i) At every iteration you must do FT of your model χ . This slows down the fitting a lot. (ii) Whereas one can estimate uncertainties of μ_0 in k -space, to do so in r -space is difficult. However, one can easily estimate measurement noise in r -space as constant and taken from a far- r region (see Section 7.1). As seen in Section 7.8, it is uncertainties of μ_0 , not the measurement noise, that mostly contribute to uncertainties of χ . Therefore, to determine the χ^2 statistics, and hence to find the true fitting errors, in r -space is more difficult.

8.7 Is it possible to fit S_0^2 ?

Some people determine the many-body factor S_0^2 experimentally from a reference spectrum. For this they fix the coordination numbers to the known values and let the factor S_0^2 vary. Then they use the obtained S_0^2 for the other spectra at the same absorption edge. Is this way good?

There are several factors which contribute to the EXAFS amplitude:

- 1) coordination numbers,
- 2) S_0^2 : intrinsic losses (core-hole lifetime broadening).
- 3) extrinsic losses (photoelectron free path). In FEFF, these are calculated from the self-energy based on a simplistic electron gas model and thus the free path is not well accurate [Rehr & Albers, 2000].
- 4) experimental broadening due to finite energy resolution.
- 5) other experimental factors, like non-linearity of fluorescence detector, the presence of pin-holes or high-order harmonics. All these can significantly lower the measured EXAFS amplitude.
- 6) some amplitude losses when μ_0 was constructed by a smoothing spline.

Even if the contributions '5' and '6' are carefully eliminated, the "experimental" (fitted) S_0^2 contains at least the effect of the contribution '4'. The fitted S_0^2 is therefore not purely due to intrinsic losses and is not of much physical interest.

There is another way of how to tune the global amplitude: to put the reduction factors onto the calculated scattering amplitude.

The latter way is preferable because it accounts for the experimental and core-hole lifetime broadening more correctly due to the following. The EXAFS wiggles are approximately of constant width in k -space and thus of growing width in E -space as one goes away from the absorption edge. Therefore any broadening of a convolution type is less important far from the absorption edge: the broadening function there is more δ -function-like in comparison with the wide EXAFS wiggles, hence the convolution integral leaves the absorption coefficient there almost unchanged. As seen, the amplitude reduction due to the broadening effects is not constant along the spectrum. This is the way how the broadening appears in FEFF, whereas the "experimental" (fitted) S_0^2 assumes all the amplitude factors to be constant.

Finally, if you want "experimental" S_0^2 , do the following. Set $S_0^2 = 1$ in the fitting options and run the fitting for your reference spectrum. Divide the obtained coordination numbers by their true values. Use then the thus obtained S_0^2 for the spectra of interest.

If you include the amplitude corrections into the amplitude calculated by FEFF (as I do), do the following. Use the cards S02 or EDGE or HOLE and set the value of $S_0^2 < 0.1$. This will make FEFF cal-

culate it. Use the obtained value in the VIPER fitting options. Use EXCHANGE card of FEFF with a positive imaginary part. Check that the coordination numbers obtained in VIPER with the calculated amplitudes are correct within the fitting errors. If not, put another value into the EXCHANGE card and run FEFF and fitting again. Usually one needs just a few such steps to achieve good coordination numbers with discrepancies much smaller than the fitting errors.

8.8 Details of the fitting algorithm

The minimized function of the fitting parameters is

$$R = 100 \left[\frac{\sum_{i=1}^{N_{pts}} (y_{i \text{ exp}} - y_{i \text{ mod}})^2}{\sum_{i=1}^{N_{pts}} (y_{i \text{ exp}})^2} \right]^{1/2},$$

where y_i denotes k^w -weighted $\chi(k_i)$ if the fitting is done in k -space or $\chi(r_i)$ if the fitting is done in r -space.

The minimization algorithm is the simplex downhill method which is shortly described here. Starting from the initial parameter set, each of the P parameters, one by one, gets an increment. There are now $P+1$ points, counting also the initial one, which form the simplest geometrical figure in the P -dimensional space (whence 'simplex'). In every apex of the simplex the figure of merit R is calculated. The highest apex (where R is largest) is reflected relative to the center of the opposite simplex face. At the new point the R value is calculated. The number of reflections is shown in VIPER fitting dialogs as the number of iterations (small iterations). The reflections continue until the highest apex remains the highest also after the reflection. Then the lowest apex is found and a new smaller simplex is constructed around it. The number of simplex constructions is shown as the number of iterations in parentheses (big iterations).

Termination criteria are the following: (1) the difference in R -factor at the highest and at the lowest apex $< 10^{-6}$ or (2) the number of big iterations $> 10 \cdot P$. These criteria are arguable, of course, but according to my experience are reasonable. Another value based on my experience is the shrinkage factor at the big iteration. I have found the fastest convergence with the factor equal to e (the base of natural logarithm).

Earlier versions of VIPER had the possibility of initial simplex deformations such that all the simplex edges had (almost) equal differentials of the function minimized. Sometimes this gave faster convergence. However, if the model had a bad parameter, i.e. a parameter to which the model was weakly sensitive, such deformations led to extremely long convergence. At present, I suggest setting the initial increments by hand (do this in the Options dialog). In this way you may visually check how responsive is your model to your parameters when you change them by spin buttons. For the standard fitting procedure the default increments are reasonably good.

9 Statistical evaluations in fitting

The ultimate goal of these methods is to find the confidence limits for the found fitting parameters.

9.1 χ^2 statistics

Assume for the experimental curve \mathbf{d} defined on the mesh $x_1, \dots, x_{N_{pts}}$ there exists a model \mathbf{m} that depends on P -dimensional parameter vector \mathbf{p} . In XAFS fitting problems as \mathbf{d} may serve both $\chi(k)$ and $\chi(r)$. The problem is to find the parameter vector \mathbf{p}_0 that gives the best coincidence of the experimental and the model curves. Introduce the figure of merit: the χ^2 -statistics (do not confuse it with the symbol of EXAFS function) as

$$\chi^2 = \sum_i^{N_{pts}} \frac{(d_i - m_i)^2}{\varepsilon_i^2}, \quad (9.1)$$

where ε_i is the error of d_i . The variate χ^2 obeys the χ^2 -distribution law with $N_{pts}-P$ degrees of freedom.

Often a preliminary processing (before fitting) is needed: smoothing, filtration etc. During the pre-processing some part of the experimental information is lost, and on the variates $\xi_i = (d_i - m_i)/\varepsilon_i$ additional

dependencies are imposed (before, they were bound solely by the model \mathbf{m}). It is necessary to determine the number of *independent* experimental points N_{ind} . For the commonly used in EXAFS spectroscopy Fourier filtering technique, the number of independent points is given by [Stern, 1993]:

$$N_{ind} = 2\Delta k \Delta r / \pi + 2, \quad (9.2)$$

where $\Delta k = k_{max} - k_{min}$ and $\Delta r = r_{max} - r_{min}$ are the ranges in k - and r -spaces used for the analysis, and $r_{min} > 0$. If $r_{min} = 0$ then

$$N_{ind} = 2\Delta k \Delta r / \pi + 1. \quad (9.3)$$

Instead of keeping in the sum (9.1) only N_{ind} items which are equidistantly spaced on the grid $x_1, \dots, x_{N_{pts}}$, it is more convenient to introduce the scaling factor N_{ind}/N_{pts} :

$$\chi^2 = \frac{N_{ind}}{N_{pts}} \sum_i^{N_{pts}} \frac{(d_i - m_i)^2}{\varepsilon_i^2}. \quad (9.4)$$

Now the variate χ^2 follows the χ^2 -distribution with $\nu = N_{ind} - P$ degrees of freedom. It can be easily verified that with the use of all the available data ($r_{min} = 0$ and $r_{max} = \pi/2dk$) the definition (9.4) turns into (9.1).

Important: ν must be positive, i.e. $P < N_{ind}$. Otherwise the statistical properties of the variate χ^2 are undefined. In practice this leads to very unstable fitting: a small variation even in a single parameter may lead to an essentially different fit.

9.2 Posterior distribution

Let us now derive the posterior distribution for an arbitrary fitting parameter p_j :

$$P(p_j | \mathbf{d}) = \int d\mathbf{p}_{\neq j} P(\mathbf{p} | \mathbf{d}) \quad (9.5)$$

where $P(\mathbf{p} | \mathbf{d})$ is the joint probability density function for all model parameters \mathbf{p} , and the integration is done over all $p_{i \neq j}$. According to Bayes theorem,

$$P(\mathbf{p} | \mathbf{d}) = \frac{P(\mathbf{d} | \mathbf{p}) P_{prior}(\mathbf{p})}{P(\mathbf{d})}, \quad (9.6)$$

with $P_{prior}(\mathbf{p})$ being the joint prior probability for all parameters, $P(\mathbf{d})$ a normalization constant. Assuming that N_{ind} values in \mathbf{d} are independent and normally distributed with zero expected values and the standard deviations ε_i , the probability $P(\mathbf{d} | \mathbf{p})$, so-called likelihood function, is given by

$$P(\mathbf{d} | \mathbf{p}) \propto \exp(-\chi^2/2) \quad (9.7)$$

where χ^2 was defined above by (9.4). Its expansion in \mathbf{p} near the minimum χ_0^2 (where $\nabla_p \chi^2 = 0$) which is reached at $\mathbf{p} = \mathbf{p}_0$ yields:

$$P(\mathbf{d} | \mathbf{p}) \propto \exp(-\chi_0^2/2) \exp\left(-\frac{1}{4} \sum_{k,l}^P H_{kl} \Delta p_k \Delta p_l\right) \quad (9.8)$$

where $\Delta p_k = p_k - p_{0k}$, and the Hessian H components (the second derivatives $H_{kl} \equiv \partial^2 \chi^2 / \partial p_k \partial p_l$) are calculated in the fitting program at the minimum of χ^2 . The sufficient conditions for the minimum are:

$$H_{kk} > 0 \quad \text{and} \quad H_{kk}H_{ll} - H_{kl}^2 > 0 \quad \text{for any } k, l. \quad (9.9)$$

Hence, the surfaces of constant level of $P(\mathbf{d} | \mathbf{p})$ are ellipsoids.

9.3 Simplest cases: partial correlations

If one ignores the prior probability then the posterior probability density function $P(\mathbf{p} | \mathbf{d})$ coincides with the likelihood $P(\mathbf{d} | \mathbf{p})$. Let us consider here two widely used approaches.

(a) *Parameters are perfectly uncorrelated*

In this case the Hessian is diagonal and

$$P(p_j | \mathbf{d}) \propto \exp\left(-\frac{1}{4} H_{jj} \Delta p_j^2\right)$$

The standard deviation of p_j is just

$$\delta^{(a)} p_j = (2/H_{jj})^{1/2}. \quad (9.10)$$

(b) *Parameter p_j essentially correlates solely with p_i*

In this case

$$\begin{aligned} P(p_j | \mathbf{d}) &= \int dp_i P(p_i p_j | \mathbf{d}) \propto \int dp_i \exp\left(-\frac{1}{4} H_{jj} (\Delta p_j)^2 - \frac{1}{2} H_{ij} \Delta p_i \Delta p_j - \frac{1}{4} H_{ii} (\Delta p_i)^2\right) \\ &\propto \exp\left(-\frac{1}{4} [H_{jj} - H_{ij}^2 / H_{ii}] (\Delta p_j)^2\right) \end{aligned}$$

from where one finds the mean-square deviation

$$\delta^{(b)} p_j = \left(\frac{2H_{ii}}{H_{jj}H_{ii} - H_{ij}^2} \right)^{1/2}. \quad (9.11)$$

In practice, to find the strongly correlated pairs of parameters one finds the pair-correlation coefficients:

$$r_{ij} = -\frac{H_{ij}}{\sqrt{H_{ii}H_{jj}}}. \quad (9.12)$$

taking the values from -1 to 1 . Two parameters are uncorrelated if their correlation coefficient is close to zero. Through the correlation coefficient, the mean-square deviations found for the cases (a) and (b) are simply related:

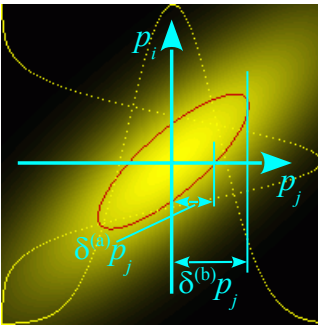
$$\delta^{(a)} p_j = \delta^{(b)} p_j \sqrt{1 - r_{ij}^2} \quad (9.13)$$

This relation can be also represented graphically. Consider the joint probability function

$$P(p_i p_j | \mathbf{d}) \propto \exp\left(-\frac{1}{4} H_{jj} (\Delta p_j)^2 - \frac{1}{2} H_{ij} \Delta p_i \Delta p_j - \frac{1}{4} H_{ii} (\Delta p_i)^2\right)$$

shown in the figure below as a color map. The ellipse of standard deviation (shown by red) is described by

$$\frac{1}{2} H_{jj} (\Delta p_j)^2 + H_{ij} \Delta p_i \Delta p_j + \frac{1}{2} H_{ii} (\Delta p_i)^2 = 1 \quad (9.14)$$



For this ellipse, the point of intersection with the line $\Delta p_i = 0$ and the point of maximum distance from the line $\Delta p_j = 0$ give the standard mean-square deviations $\delta^{(a)} p_j$ and $\delta^{(b)} p_j$.

9.4 General case: total correlations and *a priori* information

Now, we should define the prior probability. Let the parameter p_k be known to be within the range of the size S_k . Then the prior probability can be expressed as:

$$P_{prior}(\mathbf{p} | \alpha) \propto \alpha^{P/2} \exp\left(-\frac{\alpha}{2} \sum_k^P (\Delta p_k / S_k)^2\right). \quad (9.15)$$

The rationale of this prior is that it maximizes the information theory entropy $-\int P_{prior} \ln P_{prior} d\mathbf{p}$ under the constraints $\int P_{prior} d\mathbf{p} = 1$ and $\langle p_k p_l \rangle_{prior} = \delta_{kl} S_k^2$ ($\langle \dots \rangle_{prior}$ means averaging over the prior probability P_{prior}). In other words, this prior introduces minimum information in addition to the approximate know-

ledge of the sizes S_k . The regularization parameter α specifies the relative weight of the prior probability; at $\alpha = 0$ there is no prior information, at $\alpha \rightarrow \infty$ the fitting procedure gives nothing and the posterior distribution coincides with the prior one. In the expression (9.15) α appears as a known value. In reality, α is yet to be determined. This problem will be considered below.

Finally, for the joint posterior probability density function we have:

$$P(\mathbf{p} | \mathbf{d}, \alpha) \propto \alpha^{P/2} \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right), \quad (9.16)$$

where $A_{kl} = H_{kl}/2 + \alpha \cdot \delta_{kl} S_k^{-2}$.

Now, if α was known, the standard errors of the fitting parameters could be readily obtained:

$$(\delta^{(c)} p_j)^2 = \frac{\int \Delta p_j^2 P(p_j | \mathbf{d}, \alpha) d\mathbf{p}}{\int P(p_j | \mathbf{d}, \alpha) d\mathbf{p}} = \frac{\int \Delta p_j^2 \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}}{\int \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}}. \quad (9.17)$$

Further, we find the eigenvalues λ_i and corresponding eigenvectors \mathbf{e}^i of the matrix A and change the variables:

$$b_i = \sqrt{\lambda_i} \sum_k^P \Delta p_k e_k^i, \quad \Delta p_k = \sum_i^P \frac{b_i e_k^i}{\sqrt{\lambda_i}},$$

where e_k^i is the k -th component of the eigenvector \mathbf{e}^i .

Using the properties of eigenvectors:

$$\sum_k^P A_{lk} e_k^i = \lambda_i e_l^i, \quad \sum_k^P e_k^l e_k^i = \delta_{li},$$

one gets the matrix A diagonalized and:

$$(\delta^{(c)} p_j)^2 = \frac{\int \left(\sum_i^P b_i e_j^i / \sqrt{\lambda_i}\right)^2 \exp\left(-\frac{1}{2} \sum_i^P b_i^2\right) d\mathbf{b}}{\int \exp\left(-\frac{1}{2} \sum_i^P b_i^2\right) d\mathbf{b}} = \sum_i^P \frac{(e_j^i)^2}{\lambda_i}, \quad (9.18)$$

which is the fitting error taking into account *all* pair correlations.

9.5 Most probable *a priori* weight

The only problem that remains to be solved is to determine the parameter α . On one hand, α specifies the relative weight of the prior information (put $\alpha = 0$ if you do not want it), on the other hand, it makes the matrix A be positively defined. In general, the matrix H is not positive, especially when P , the number of fitting parameters, is big and several ill-conditioned directions appear in the parameter space or when the model and experimental curves differ significantly. Thus, the regularization with a sufficiently big α guarantees that all λ 's in Eq. (9.18) are positive and essentially not zeros.

In the modern Bayesian methods, α itself is determined by Bayesian arguments that maximize the posterior probability of α given the data [Turchin & Nozik, 1969]:

$$P(\alpha | \mathbf{d}) = \int d\mathbf{p} P(\alpha, \mathbf{p} | \mathbf{d}) = \int d\mathbf{p} P(\alpha) P(\mathbf{p} | \mathbf{d}, \alpha) \quad (9.19)$$

Using a prior $P(\alpha) = 1/\alpha$ (so-called Jeffreys prior [Jeffreys, 1939]), one obtains the posterior distribution:

$$P(\alpha | \mathbf{d}) \propto (\lambda_1 \cdots \lambda_P)^{-1/2} \alpha^{P/2-1}. \quad (9.20)$$

Having found the maximum of this distribution, one obtains for α its most probable value α_{mp} and the corresponding matrix A . Then by Eq. (9.18) one finds the Bayesian errors of fitting parameters. I have found (the proof is in Appendix 9.10) that at $\alpha = \alpha_{mp}$:

$$\left\langle \chi^2 - \chi_0^2 \right\rangle_{post} \Big|_{\alpha=\alpha_{mp}} = 2, \quad (9.21)$$

where $\langle \dots \rangle_{post}$ means averaging over the posterior probability $P(\mathbf{p}|\mathbf{d}, \alpha)$. The difference $\langle \chi^2 - \chi_0^2 \rangle_{post}$ taken at $\alpha = \alpha_{mp}$ appears to be independent of the number of fitting parameters P and ε_i (and any pre-factor in the definition of χ^2) although χ^2 itself does depend on P and ε_i . Eq. (9.21) can be considered as an equivalent equation for the maximization of $P(\alpha|\mathbf{d})$.

9.6 What if the experimental errors ε_i were determined incorrectly?

Notice that the above formulas need χ^2 which, in turn, needs the experimental errors ε_i . The latter were determined in Section 7 by various ways and with different results. How important is to know the errors ε_i correctly?

If in the definition of χ^2 (9.4) the errors ε_i are taken by a factor γ smaller than the correct ones then χ^2 is bigger by γ^2 , the Hessian H with all its components and the eigenvalues is also bigger by γ^2 . By expressions (9.10), (9.11) and (9.18), the fitting errors are then by γ smaller than the correct ones.

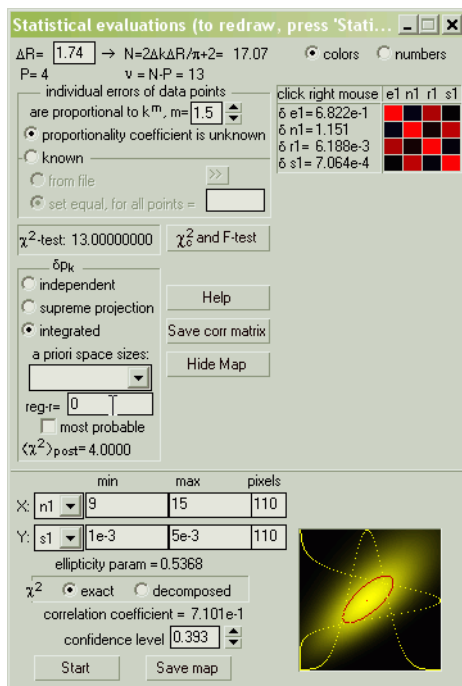
Finally, underestimated errors ε_i give proportionally underestimated fitting parameters.

9.7 What if the experimental errors ε_i are unknown?

Even in this case it is possible to estimate the fitting errors. We assume that χ^2 follows the χ^2 -distribution with $\nu = N - P$ degrees of freedom. Its most probable value is ν . Thus we can determine the errors ε_i from the equality for the best fit χ^2 value: $\chi_0^2 = \nu$. For the experimental errors ε_i in k -space this approach can be further extended. We can set the spectrum acquisition program such that $\varepsilon_i \propto k^m$, and the proportionality constant is again found from $\chi_0^2 = \nu$. If the experimental EXAFS curve was measured with sampling time $\propto k^{2m}$ then the measurement noise is $\propto k^{-m}$ and the noise ε_i of $\chi \cdot k^w$ is $\propto k^{w-m}$. For the example in Section 8.1.3 the spectrum was measured with sampling time $\propto k$ and then the k^2 weighting was applied to $\chi(k)$. The resulting noise should have the behavior $\varepsilon_i \propto k^{1.5}$. If you measure with constant sampling time, your experimental noise is simply weighted by k^w weighting of $\chi(k)$.

Note that if you forget to weight the noise by k^w , the resulting fitting errors are *underestimated* (play with k^m weighting in the 'Statistical evaluations' dialog and watch how the fitting errors are changing).

9.8 'Statistical evaluations' dialog



← (shown here with 'Mapping...' activated and calculated)

$\Delta k = k_{\max} - k_{\min}$ — the range in k -space, calculated automatically.

$\Delta R = r_{\max} - r_{\min}$ — the range in r -space, calculated automatically for fits in r - or Fourier-filtered k -space, i.e. when the FT window exists. Otherwise, must be set manually.

N_{ind} — the number of independent points (9.2), here denoted by N .

P — the number of fitting parameters.

$\nu = N_{ind} - P$ — degree of freedom.

data errors — if the experimental errors are known from Section 7, load them as a curve or set as a value.

k^m — weighting of noise: see Section 9.7.

χ^2 -test — if the data errors are set unknown, the χ^2 value is forced to be ν (displayed in black). If the errors are given explicitly, the χ^2 -test is performed: the χ^2 value smaller than the critical value $(\chi^2)_c$ for the given significance level c (see Section 9.9.1) is displayed in **green**, otherwise in **red**.

fitting errors δp_k — the three options correspond to the errors $\delta^{(a)}p_k$ (9.10), $\delta^{(b)}p_k$ (9.11) and $\delta^{(c)}p_k$ (9.18). The errors themselves are displayed at the left of the colored correlation table. The last option of the

three is the most correct one. The other two are primarily of educational merit, for showing how smaller the fitting errors can be in neglecting the correlations.

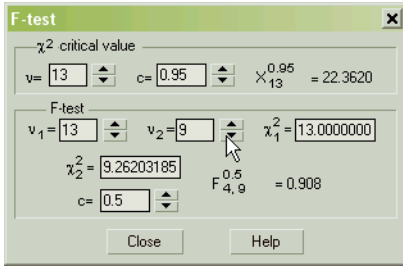
A priori space sizes, regularizer — If you want to use the *a priori* part, set reasonably big limits S_k in the drop-down list and select 'most probable' for the regularizer α .

If you do not want the *a priori* part (normally I do not use it), set the regularizer $\alpha=0$ and unselect 'most probable'.

The colored matrix at the top right part shows the correlation coefficients (red = close to 1, blue = close to -1, black = close to 0). The column at its left represents the main result of the statistical evaluations – the fitting errors. It may happen that some correlation coefficients are displayed white and/or some fitting errors are set to zero. This happens when the conditions (9.9) are not fulfilled because the minimum was not reached or when the Hessian cannot be inverted because of strong correlations.

The 2D mapping is a graphical tool visualizing the likelihood function $\exp(-\chi^2)$ dependent on two selected fitting parameters. This visualization does not have much practical value and only serves to demonstrate the presence or absence of correlations.

9.9 Statistical tests



9.9.1 χ^2 -test

A good EXAFS model should give the variate χ^2 defined by (9.4) that follows the χ^2 distribution law, that is this variate should not fall within the tail of this distribution. In other words, the χ^2 value should be smaller than the critical value $(\chi^2)_c$ for the given significance level c . Typically, the significance level for this test is selected to be 0.95.

Of course, this test strongly depends on the estimation for the experimental errors ε_i .

9.9.2 *F*-test

Let there be a possibility to choose between two EXAFS models depending on different numbers of parameters P_1 and P_2 . Which one of them is more statistically important? For instance one wishes to decide whether a single coordination sphere is split into two.

Let χ_1^2 and χ_2^2 follow the χ^2 -distribution law with $v_1 = N_{ind} - P_1$ and $v_2 = N_{ind} - P_2$ degrees of freedom, correspondingly. From the linear regression problem (near the minimum of χ^2 , the likelihood function is expressed by (9.8) and is identical in form to that of the linear regression problem) it is known that the value

$$f = \frac{(\chi_1^2 - \chi_2^2)/(v_1 - v_2)}{\chi_2^2/v_2}$$

obeys the Fisher-Snedecor *F*-distribution law with $(v_1 - v_2, v_2)$ degrees of freedom if exactly $r = v_1 - v_2$ parameters in the second model are linearly dependent. In order the linear restrictions on the second model be absent, the value f should *not* follow the *F*-distribution, that is it should be greater than the critical value $[F_{(v_1-v_2, v_2)}]_c$ for the specified significance level c : $f > [F_{(v_1-v_2, v_2)}]_c$, or

$$\chi_2^2 < \chi_1^2 \left((F_{v_1-v_2, v_2})_c \frac{v_1 - v_2}{v_2} + 1 \right)^{-1} \quad (9.22)$$

Notice, that the expression (9.22) means the absence of exactly r linear restrictions on the second model parameters. Even if (9.22) is realized, less number of linear dependencies are possible. If, for instance, the splitting of a single coordination sphere into two does not contradict to the F-test (9.22), some of the parameters of these two spheres may be dependent, but not all. This justifies the introduction of a new sphere into the model EXAFS function.

Thus, having specified the significance level c , one can answer the question "what decrease of χ^2 must be achieved to increase the number of parameters from P_1 to P_2 ?" or, inside out, "what is the probability that the model 2 is better than the model 1 at the specified (P_1, χ_1^2) and (P_2, χ_2^2) ?"

Notice, that since in the definition for f the ratio χ_1^2/χ_2^2 appears, the actual values of ε_i become not important for the F -test (only if they all are taken equal to a single value).

Consider an example with the following numbers: ($v_1 = 7, \chi_1^2 = 16.8$) and ($v_2 = 4, \chi_2^2 = 5.3$). For these values, $f = 2.89 = (F_{3, 4})_{0.84}$, from where we can assert that with the probability of 84% the model 2 is better than the model 1.

9.10 Appendix

Proof of Eq. (9.21) [I have never seen this equation presented anywhere whereas it looks nice and is easy to use in practice. In order to arrange my old notes, I am putting it here]

1) Rewrite the posterior probability (9.16) as:

$$P(\mathbf{p} | \mathbf{d}, \alpha) \propto \alpha^{P/2} \exp\left[-\frac{1}{2}(\chi^2 - \chi_0^2 + \alpha\chi_{prior}^2)\right].$$

2) Denote

$$c \equiv \int d\mathbf{p} P(\mathbf{p} | \mathbf{d}, \alpha), \quad \langle \dots \rangle_{post} \equiv \frac{1}{c} \int d\mathbf{p} (\dots) P(\mathbf{p} | \mathbf{d}, \alpha).$$

3) Notice

$$\langle \chi^2 - \chi_0^2 + \alpha\chi_{prior}^2 \rangle_{post} = \frac{\int \left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}}{\int \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}} = \frac{\int \left(\sum_i b_i^2\right) \exp\left(-\frac{1}{2} \sum_i b_i^2\right) d\mathbf{b}}{\int \exp\left(-\frac{1}{2} \sum_i b_i^2\right) d\mathbf{b}} = P,$$

$$\frac{\partial c}{\partial \alpha} = -\frac{1}{2} \langle \chi_{prior}^2 \rangle_{post} \cdot c + \frac{P}{2} \frac{c}{\alpha},$$

$$P(\alpha | \mathbf{d}) = c / \alpha.$$

4) Maximize $P(\alpha | \mathbf{d})$:

$$\frac{\partial}{\partial \alpha} P(\alpha | \mathbf{d}) = 0 \Rightarrow \frac{\partial c}{\partial \alpha} = \frac{c}{\alpha} \quad (\text{at } \alpha = \alpha_{mp}).$$

5) Finally, at $\alpha = \alpha_{mp}$

$$\langle \chi^2 - \chi_0^2 \rangle_{post} = \langle \chi^2 - \chi_0^2 + \alpha\chi_{prior}^2 \rangle_{post} - \langle \alpha\chi_{prior}^2 \rangle_{post} = P - (P - 2) = 2. \quad \square$$

In a more general case of the prior $P(\alpha) = \alpha^{-\beta}$, in the rhs of (9.21) would stay 2β .

The posterior average of $\chi^2 - \chi_0^2$ is calculated in practice as

$$\langle \chi^2 - \chi_0^2 \rangle_{post} = \sum_{ikl} \frac{H_{kl}}{2} \frac{e_k^i e_l^i}{\lambda_i}.$$

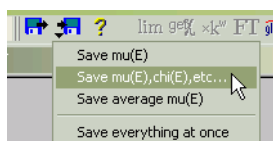
This follows from the decomposition

$$\langle \chi^2 - \chi_0^2 \rangle_{post} = \frac{\int \left(\frac{1}{2} \sum_{kl} H_{kl} \Delta p_k \Delta p_l\right) \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}}{\int \exp\left(-\frac{1}{2} \sum_{kl} A_{kl} \Delta p_k \Delta p_l\right) d\mathbf{p}},$$

diagonalization of the exponent power and noticing that the non-diagonal terms in the pre-exponent factor are odd and thus cancel in integration.

10 Exporting data and saving project file

The curves visible in any window (currents, μ , χ , FT and BFT) can be exported to column files. For this, activate the window of interest and use the top 'Save' button or the main menu 'File'.

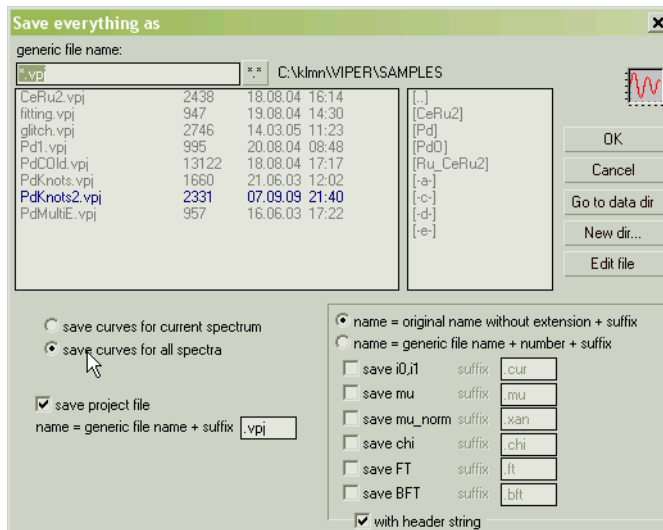


The actual menu configuration depends on the current program status.

One can export several curves at one go. Use the menu command 'Save everything at once':

This way is especially useful with many loaded spectra. When I use this dialog, I normally keep the original names of spectra and add different extensions. [Note: although the generic file name is not used in this case, the wildcards '*' and '?' must not be in it as these prevent the dialog from closing; give any dummy name without wildcards.]

Another very useful function of this dialog is to save project files. A project file has description of data files, energy calibration, parameters of χ extraction and Fourier analysis etc. It does not save deglitching or fitting parameters. The latter should be saved directly from a fitting dialog. The project files Samples/*.*vpj* have been saved in this dialog.



Important: Project files are text file. You can edit them by any common editor. A newly created project file has full path references to the data files. If you move the data files, you should change the paths accordingly. I normally keep project files in the same directory with data. Then I keep only the file names in a project file and manually delete the directory paths.

References

- Ablett J M, Woicik J C and Kao C C (2005) *International Centre for Diffraction Data, Advances in X-ray Analysis* **48**, 266.
- Booth C H and Bridges F (2005) *Physica Scripta* **T115**, 202.
- Brennan S and Cowan P L (1992) *Rev. Sci. Instrum.* **63**, 850
<http://www.bmsc.washington.edu/scatter/periodic-table.html>
ftp://ftpa.aps.anl.gov/pub/cross-section_codes/
- Brewe D L, Pease D M and Budnick J I (1994) *Phys. Rev. B* **50**, 9025.
- Carboni R, Giovannini S, Antonioli G and Boscherini F (2005) *Physica Scripta* **T115**, 986.
- Chantler C T (1995) *J. Phys. Chem. Ref. Data* **24**, 71
<http://physics.nist.gov/PhysRefData/FFast/Text/cover.html>
<http://physics.nist.gov/PhysRefData/FFast/html/form.html>
- Eisebitt S, Böske T, Rubensson J-E and Eberhardt W (1993) *Phys. Rev. B* **47**, 14103.
- Goulon J, Goulon-Ginet C, Cortes R and Dubois J M (1982) *J. Physique* **43**, 539.
- Haskel D (1999) Computer program *FLUO: Correcting XANES for self absorption in fluorescence data*, <http://www.aps.anl.gov/xfd/people/haskel/fluio.html>.
- Henke B L, Gullikson E M and Davis J C (1993) *Atomic Data and Nuclear Data Tables* **54**, 181.
http://www-cxro.lbl.gov/optical_constants/

- Hubbell J H (1969) *Natl. Stand. Ref. Data Ser.* **29**; Hubbell J H, *Radiat. Res.* **70** (1977) 58-81.
<http://physics.nist.gov/PhysRefData/Xcom/Text/XCOM.html>
- Iida A and Noma T (1993) *Jpn. J. Appl. Phys.* **32**, 2899.
- Jaynes E T (1990) *Probability theory as logic*, Maximum-Entropy and Bayesian Methods, edited by Paul F. Fougere (Kluwer Academic Publishers, Dordrecht, Holland), Vol. Proceedings, (Revised, corrected, and extended version is available from bayes.wustl.edu).
- Jeffreys H (1939), *Theory of Probability* (Oxford University Press, London), later editions: 1948, 1961, 1983.
- Kissel L, Zhou B, Roy S C, Sen Gupta S K and Pratt R H (1995) *Acta Crystallographica A* **51**, 271; Pratt R H, Kissel L and Bergstrom Jr. P M, *New Relativistic S-Matrix Results for Scattering - Beyond the Usual Anomalous Factors/ Beyond Impulse Approximation*, in *Resonant Anomalous X-Ray Scattering*, edited by G. Materlik, C. J. Sparks and K. Fischer (North-Holland: Amsterdam, 1994); Kane P P, Kissel L, Pratt R H and Roy S C *Physics Reports* **140**, 75-159 (1986); Kissel L and Pratt R H, *Rayleigh Scattering - Elastic Photon Scattering by Bound Electrons*, in *Atomic Inner-Shell Physics*, edited by Bernd Crasemann (Plenum Publishing: New York, 1985).
<http://www-phys.llnl.gov/Research/scattering/index.html>
- Klementev K V (2001a) *J. Phys. D: Appl. Phys.* **34**, 209.
- Klementev K V (2001b) *J. Phys. D: Appl. Phys.* **34**, 2241.
- McMaster W H, Kerr Del Grande N, Mallett J H and Hubbell J H (1969) *Compilation of X-Ray Cross Sections* Lawrence Livermore National Laboratory Report UCRL-50174 Section II Revision I available from National Technical Information Services L-3, U.S. Dept. of Commerce
<http://ixs.csrri.iit.edu/database/programs/mcmaster.html>
<http://cars9.uchicago.edu/~newville/mcbook/>
- Newville M, Boyanov B I & Sayers D E (1999) *J. Synchrotron Rad.* **6**, 264 (Proc. of Int. Conf. XAFS X).
- Pfalzer P, Urbach J-P, Klemm M, Horn S, denBoer M L, Frenkel A I and Kirkland J P (1999) *Phys. Rev. B* **60**, 9335.
- Pompa M, Flank A-M, Delaunay R, Bianconi A and Lagarde P (1995) *Physica B* **208&209**, 143.
- Press W H, Flannery B P, Teukolsky S A & Vetterling W T (1992-2007) *Numerical Recipes, The Art of Scientific Computing*, 2nd or 3rd Edition, Cambridge University Press.
- Ramaker D E, van Dorssen G E, Mojet B L and Koningsberger D C (2000), *Top. Catal.* **10**, 157.
- Rehr J J, Zabinsky S I, Ankudinov A and Albers R C (1995) *Physica B* **208&209**, 23.
- Rehr J J and Albers R C (2000) *Rev. Mod. Phys.* **72**, 621.
- Stern E A (1993) *Phys. Rev. B* **48**, 9825.
- Tan Z, Budnick J I and Heald S M (1989) *Rev. Sci. Instrum.* **60**, 1021.
- Tröger L, Arvanitis D, Baberschke K, Michaelis H, Grimm U and Zschech E, (1992) *Phys. Rev. B* **46**, 3283.
- Turchin V F and Nozik V Z (1969). *Izv. Atmospheric and Oceanic Physics* **5**, 29
- Wende H and Baberschke K (1999), Proc. VUV-XII San Francisco Aug. 1998, *J. Electr. Spectr. Relat. Phenom.* **101-103**, 821.

Citation: You are requested to cite VIPER as: "K. V. Klementiev, VIPER for Windows, freeware; K. V. Klementev, *J. Phys. D: Appl. Phys.* **34**, 209-17 (2001)" *[There is no typo in the names! Put it as written. Both versions refer to the same me :) The latter is an older version of the Cyrillic-to-Latin transcription of my surname. I can give you six more versions if you want :)]*