2D SIFT processing

The user's manual

Developed by
Yoh Matsuki, Osaka University
Judith Herzfeld, Brandeis University

Last update: January 26, 2016

0. Preparation using nmrPipe

"SIFT in Matlab" program reads NMR data in the nmrPipe format. So, the first thing you do is as usual the conversion of the data. For Bruker data, for example, the script will look like:

```
#!/bin/csh
bruk2pipe -in ./ser \
  -bad 0.0 -noaswap -DMX -decim 16 -dspfvs 12 -grpdly -1  \
  -xN 1024          -yN 128                \
  -xT  512          -yT 64                 \
  -xMODE DQD        -yMODE Echo-Antiecho   \
  -xSW 8389.3       -ySW 1946.7            \
  -xOBS 600.333     -yOBS 60.838           \
  -xCAR 4.771       -yCAR 118.083          \
  -xLAB HN          -yLAB 15N              \
  -ndim 2           -aq2D States           \
  -out ./pre.fid -ov
```

Make sure that –yT is set to the number of collected NUS points in the F1 dimension, 64 here. The number of collected NUS points is the number of lines in the NUS schedule file you used in the experiment.

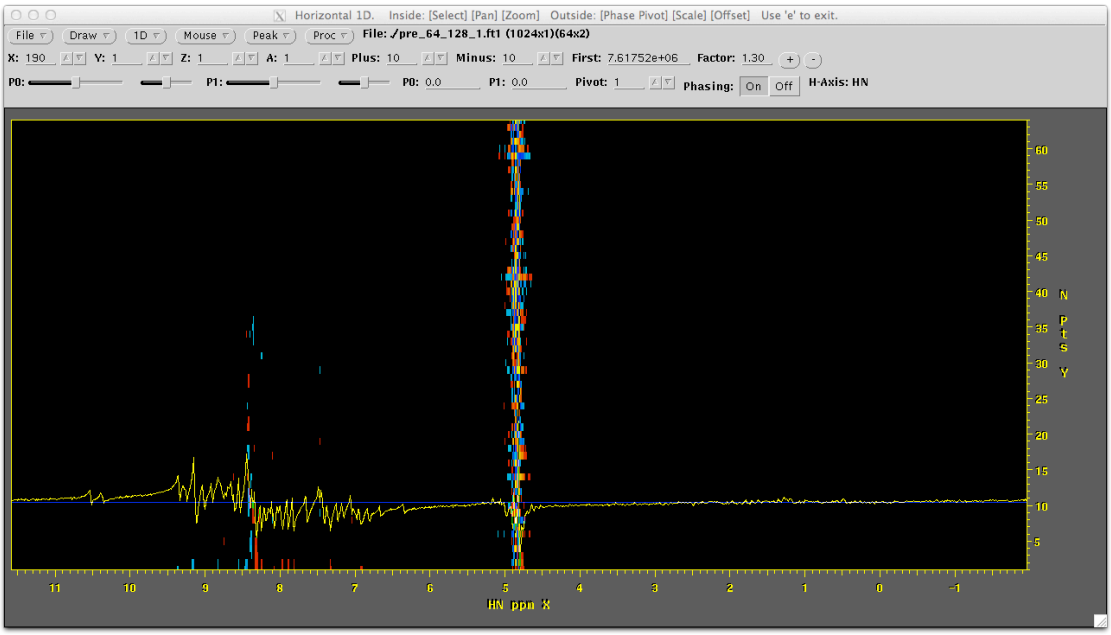
Next, determine the phase in the direct acquisition dimension. Use a script similar to this:

```
nmrPipe -in ./pre.fid                              \
| nmrPipe -fn SOL                                  \
| nmrPipe -fn ZF -auto -size 2048                  \
| nmrPipe -fn FT                                   \
| nmrPipe -fn PS -p0 -0 -p1 0 -di                  \
| nmrPipe -fn POLY -auto -ord 0                    \
| nmrPipe -fn EXT -x1 6.5ppm -xn 10ppm -sw         \
| nmrPipe -out ./pre.ft1 -ov
```

The phase is still set to all zero. Then, draw the spectrum:

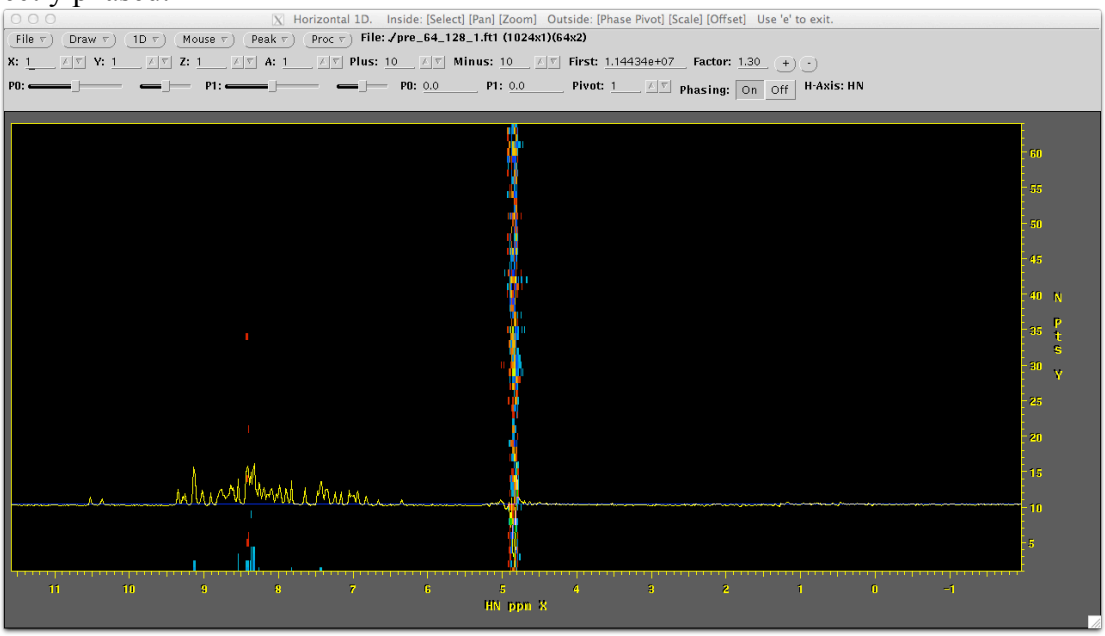
```
nmrDraw ./pre.ft1 &
```

Issue an nmrPipe shortcut "SHIFT+T", and click the "Draw" button to plot the 2D area with higher contour levels. This helps to see the slices more clearly. Hit "H" to enable the horizontal slice mode, and see the 1D slice in the first row by pointing the lowest part of the spectral area, or setting "Y:" equals to 1 in the tool area of nmrDraw. Then you will see something like this:



Looking at the 1D slice, phase it as usual. It is important to take note of this phase correction found here, `[ph0,ph1]=[-99.5, 16.8]`, for later use.

Now set the phase values in the script, and execute the script again. Check if the data is correctly phased:



If the data has been digitally filtered, the number of points in the x dimension at this stage may be changed from the one you collected in the experiment. Confirm the actual number of x points for later use, using the `showhdr` commands:

```
showhdr ./pre.fid
```

1. SIFT processing in Matlab

To start SIFT processing in Matlab, download the 2D SIFT archive from the website (http://people.brandeis.edu/~herzfeld/SIFT/), and unpack it. Do not mix the scripts with

older versions. The archive includes the run script `sift2d_v2_4.m`, a folder for the SIFT subroutines (`SIFT_suite_2D`), and `example` folder involving demo data (`gb1_2d_hsqc.bruker`).

A stepwise instruction is provided directly in the Matlab run script, but here are some tips and recommendations:

Section 1-2 _____

The NUS table is a text file that consists of a column of integers starting from 1, rather than 0, specifying recorded NUS points.

Section 1-3 _____

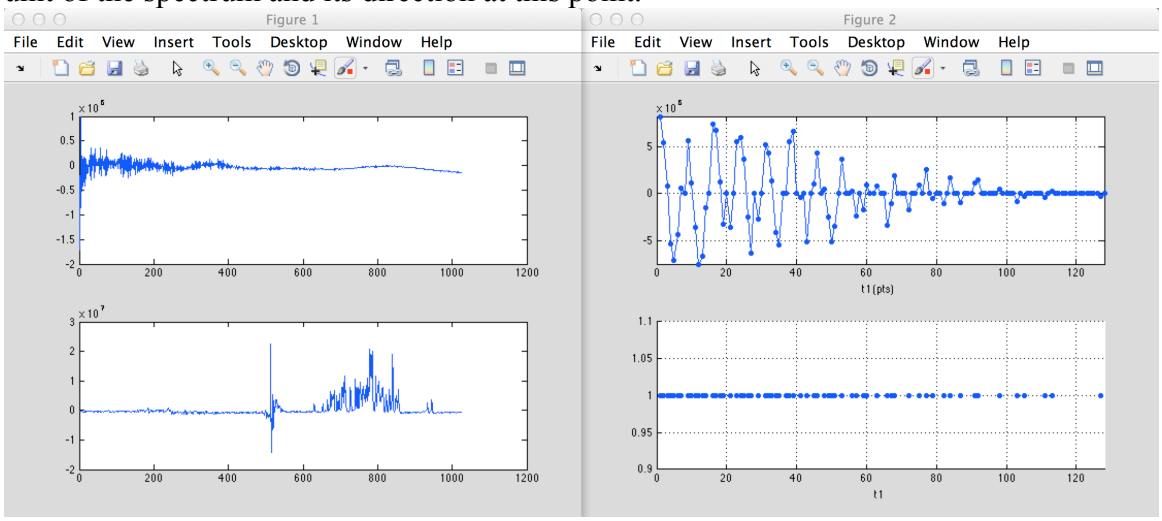
The number of total t1 grids (`nt1`) refers to the number of full grids after the SIFT reconstruction. The number of points in the direct acquisition dimension `nt2` should be the actual number of points after the treatment for the digitally filtered data. The number of points after zero filling (`nt1z, nt2z`) is normally set to the power of two. The phase for the direct acquisition dimension, `ph2`, is the set of values found during the nmPipe processing above.

Section 2-1 _____

For reading the example data ~/SIFT2DinMATLAB/example/gb1_2d_hsqc.bruker, the section 2-1(b) is to be used since the data has no 'zeros' inserted at skipped t1 points.

Section 2-2 _____

Executing this section will plot the initial FID, and its transform. Also shown is the representative interferogram along t1, and the scatter plot of your NUS schedule. Confirm that the imported data is not nonsense, and is properly reshuffled according to the NUS schedule, i.e the data have zeros only at the time grids skipped recording. The "data cursor" tool in the tool bar of each figure will be useful here. Do not worry about the axis unit of the spectrum and its direction at this point.

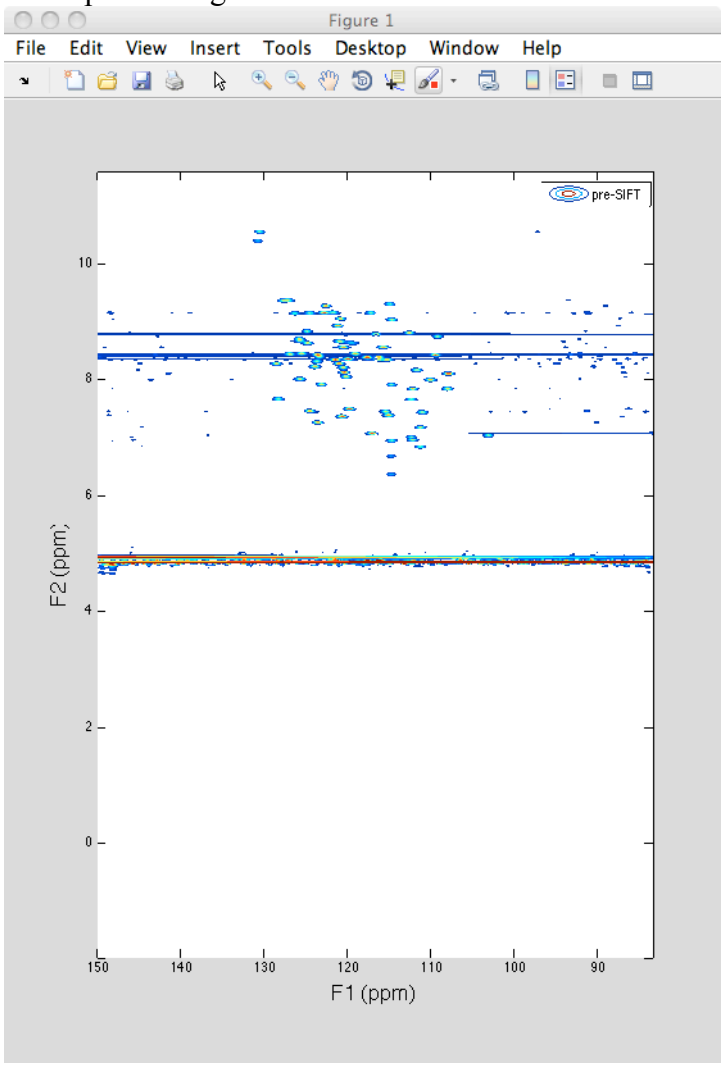


Section 2-3 _____

You can skip this section in the beginning. If you see the axis swap or inversion of the spectrum in the later part of the processing, come back to this section and change the flag.

Section 3-1 _____

Upon execution of this section, a 2D contour plot such as below will pop up; use it to scout for the bright spectral regions in the following section. You can zoom in/out or drag using the tools in the top of the figure.



Section 3-2 _____

Method A:

If you have run the NUS experiments with a slight oversampling as demonstrated in Y. Matsuki et al. *JACS* **131**, 4648- (2009) and *Angew. Chem. Int. Ed.* **49**, 9215- (2010), it is easy to find spectral limits by looking at the 2D spectrum. You could also rely on the BMRB statistics for the expected bright region in your spectrum. For example, a single bright region covering all peaks in a ^{15}N HSQC will typically have the spectral limits in ppm:

```
brightLimits_F1=[100 132]; (15N dimension)
brightLimits_F2=[6 11];    (1H dimension)
```

To define more regions, just add another row in both `brightLimits_F1` and `bright_Limits_F2` as:

```
brightLimits_F1=[100 132;
                 125 135];
brightLimits_F2=[6 9.5;
                 10 11];
```

Each corresponding rows define together a single 2D bright region. You can define indefinite number of bright regions.

Section 3-3 _____

Method B:

You can set the bright regions using the graphical input tool. Specify peak positions by clicking the pre-SIFT spectrum plotted in section 3-1. A set of bright regions centered to the clicked position, and with the width defined by user will be generated.

A recommended protocol is to generate bright regions conservatively in the first round of the SIFT processing. You may come back to this section with the post-SIFT data for scouting additional bright regions and for a second round processing by SIFT.

For a serial NUS data, the reference spectrum can be used for picking peaks as demonstrated in Y. Matsuki et al., *J. Phys. Chem. B* **115**, 13740 (2011).

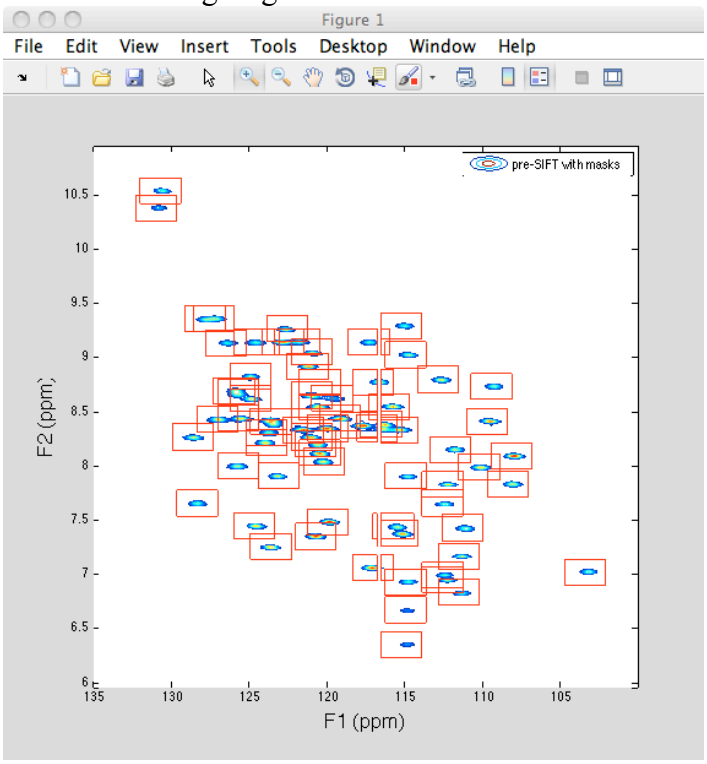
Section 3-4 _____

Method C:

The same thing can be achieved by copying and pasting the Sparky's peak list to `peakList`. For outputting a spectrum into a Sparky file for peak picking, see the section 5-3.

Section 3-5 _____

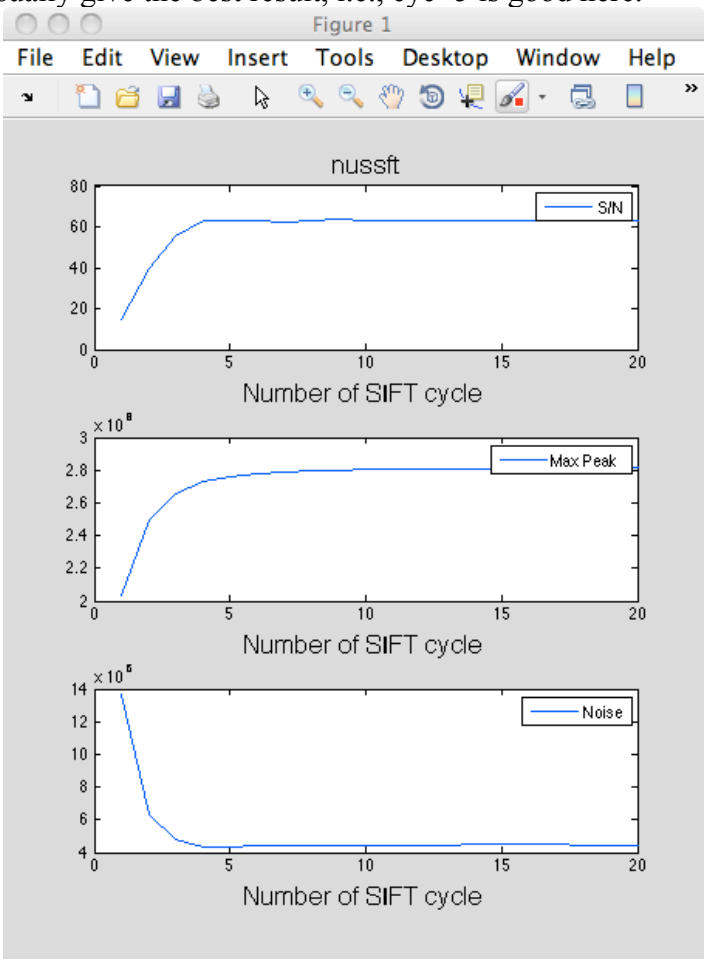
The bright regions generated can be plotted onto the spectrum. The figure below shows the bright regions with red rectangles generated in the Method C.



Section 4-1 _____

A recommended protocol is to do an exploratory SIFT processing with relatively large number of cycles, say cyc=20-50, look at how does the S/N ratio improve along the cycle. Then re-run the final SIFT processing with the best number of cycle.

Once the exploratory run is completed, a figure showing the noise level, peak intensity, and S/N ratio along the SIFT cycle will pop up. The shortest cycle after the S/N ratio leveled off will usually give the best result, i.e., cyc=5 is good here.



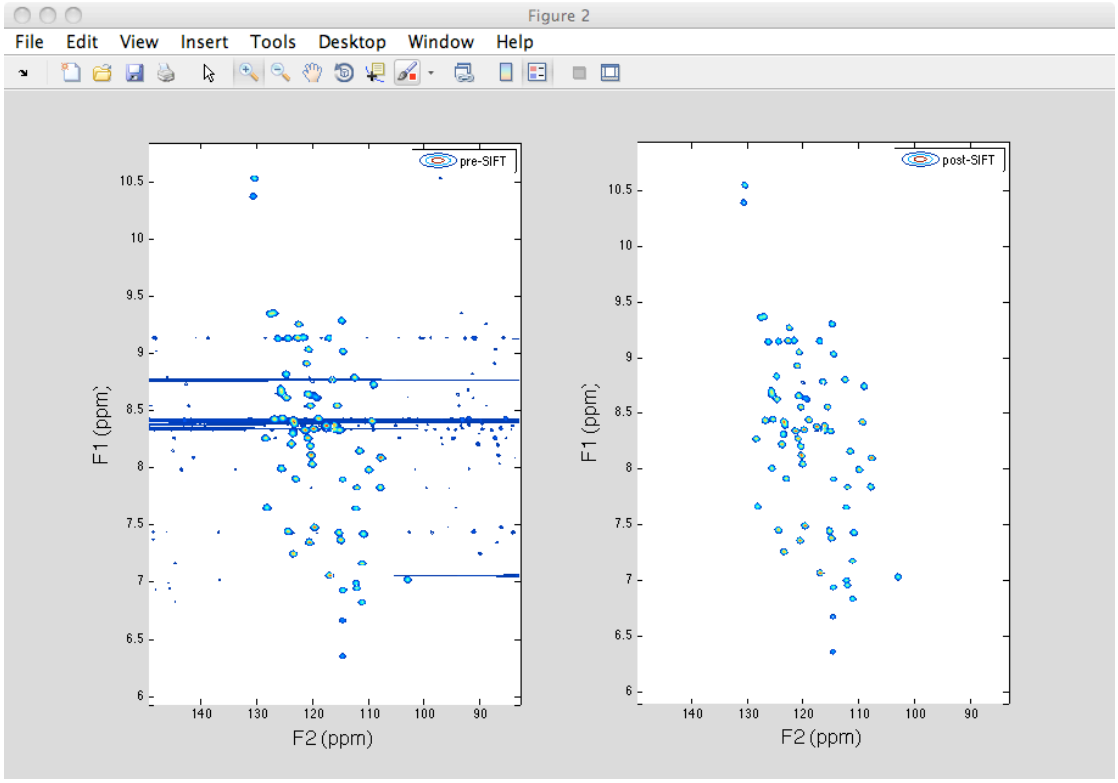
Section 5-1 _____

SIFT process takes data in the time domain, and outputs the SIFTed data in the time domain. Executing this section will output the data into Bruker format for subsequent standard 2D NMR processing.

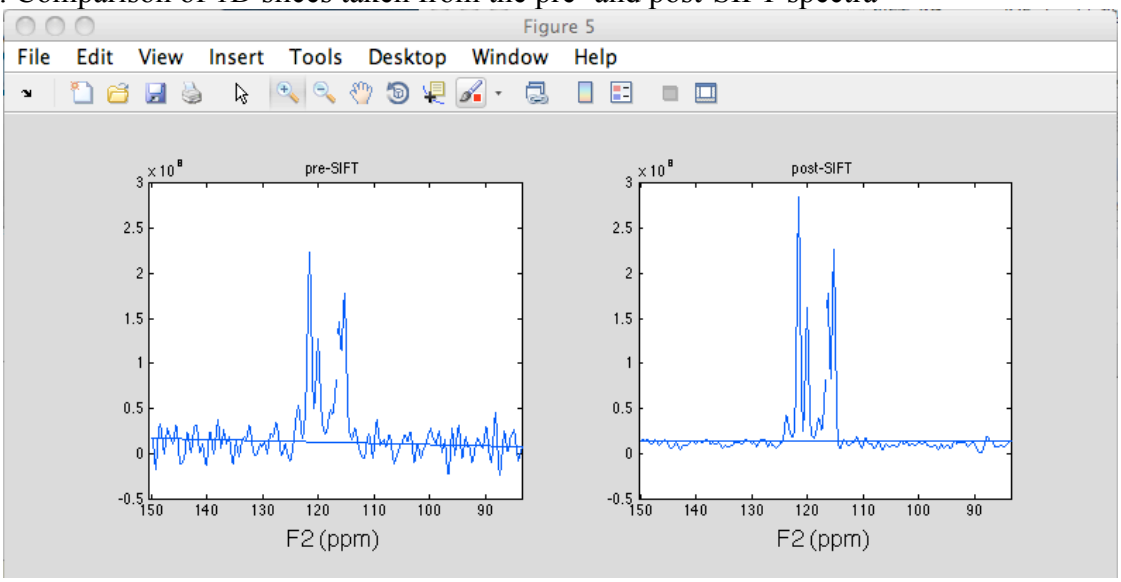
Section 5-4 _____

Data processed in Matlab can be plotted in Matlab. By default, the run script as is provided allows you to plot the following spectra:

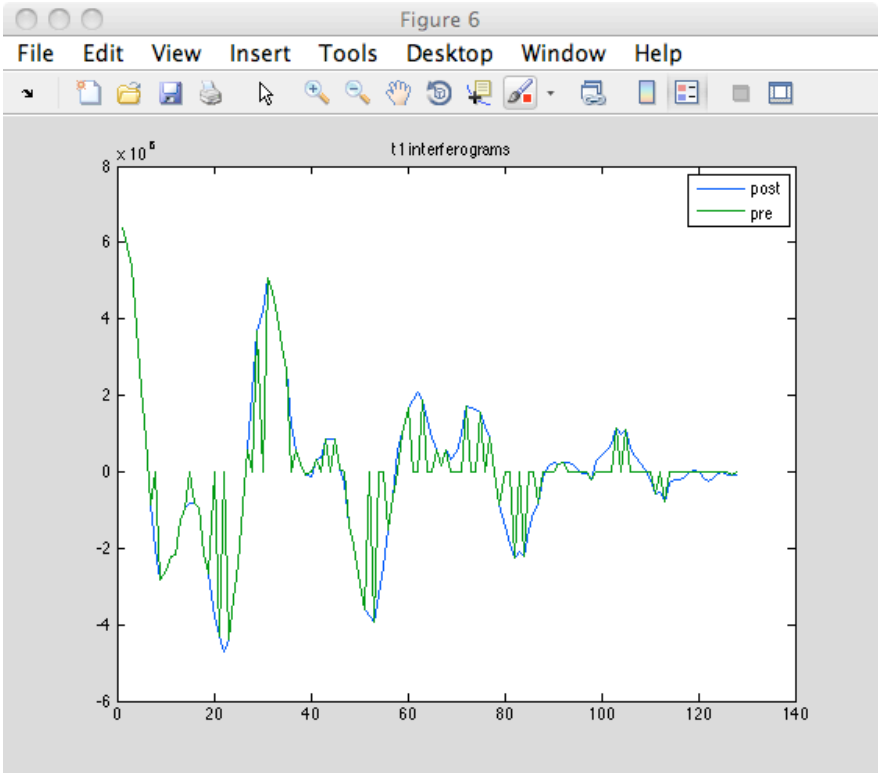
a. 2D pre- and post-SIFT spectra side by side.



b. Comparison of 1D slices taken from the pre- and post-SIFT spectra



c. Comparison of t1 interferograms for the pre- and post-SIFT spectra at a user-defined F2 position. Note that the interferograms are generated and stored as one of the output arguments of the `proc2d` command in the section 5-2.



2. Processing of post-SIFT data

The section 6-1 outputs post-SIFT data in Bruker format. Convert it to the nmrPipe format using a script similar to the following. Now, the number of data points in t1, `-yT`, should be the number of full grids rather than the number of collected NUS points. Also, the acquisition mode should be all "Complex".

```
bruk2pipe -in ./post.bruker \
  -xN 1024          -yN 256                \
  -xT  512          -yT 128                \
  -xMODE Complex    -yMODE Complex         \
  -xSW 8389.3       -ySW 1946.7            \
  -xOBS 600.333     -yOBS 60.838           \
  -xCAR 4.771       -yCAR 118.083          \
  -xLAB HN          -yLAB 15N              \
  -ndim 2           -aq2D States           \
  -out ./pre.fid -ov
```

Subsequent nmrPipe processing of the post SIFT data will just be all conventional 2D processing. Any standard procedures can be performed, including the linear back/forward prediction, windowing, zero-filling, FT/iFT, baseline correction etc.