User Manual
07 March 2013

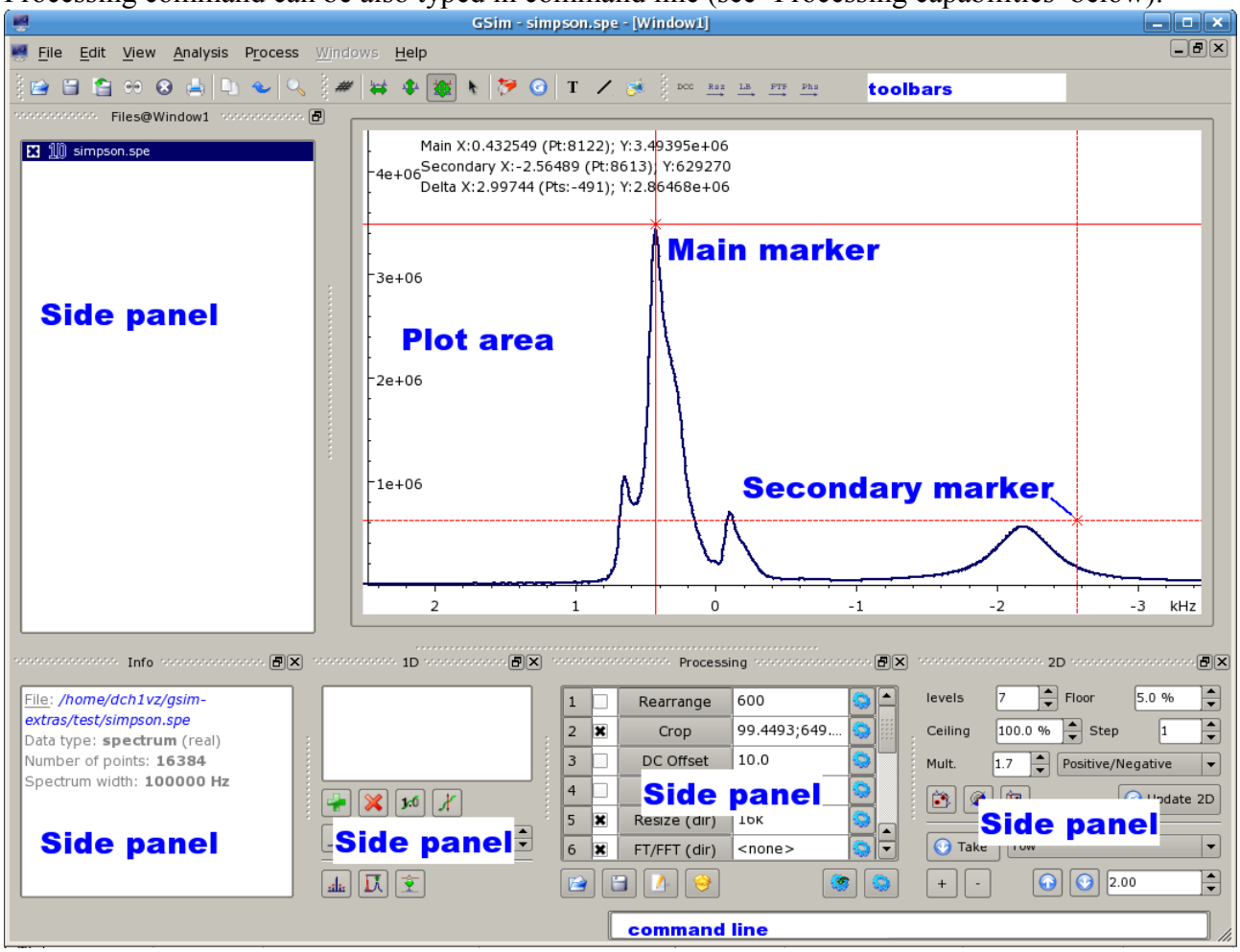
If you have found these programs useful, please, make a references to this web-page in your publications:
Vadim E. Zorin, "*GSim - visualisation and processing tool for NMR experiments and simulations*", URL http://gsim.sourceforge.net

1. Disclaimer

Since 2008 the author of this program is employed by Agilent Technologies UK Ltd. (former Varian Ltd.). Nevertheless GSim is a personal project of the author maintained in his spare time. GSim is NOT associated, distributed or supported by Agilent or Varian. The program DOES NOT use any Agilent/Varian proprietary code, algorithms or methods apart from those described in publicly available documentations or other publications. Any questions related to GSim should be discussed directly with the author of the program.

2. Interface:

The interface consists of the main plot area and several movable side panels. Below the side panels will be referred by their names: ‘Files’ , ‘Browser’, ‘Info’, ‘Processing’, ‘1D’, ‘2D’. All side panels can be moved around by simple mouse dragging. Panels can be combined in a single tabulated panel if you drop the one of the panels on the top of another. Most of the panels can be closed. If you need to open it again use ‘View->Tools’ or click the right mouse button on the tool bars or on the top part of any side panel. The ‘Files’ panel is the only panel that you cannot close. Closing this panel means you lose the control over the active dataset which should be avoided.
The main plot window can display either 1D or 2D NMR data, depending on what kind of data is currently ‘active’ (see below for the ‘active’ dataset definition).
In principle, several plot windows can be opened simultaneously (see, for example, ‘File->New window’ description). Data can be moved between windows using “File->Move to command”. Side panel “Files” contains only files, opened in the current plot window.
Processing command can be also typed in command line (see ‘Processing capabilities’ below).



Key definitions:

All datasets with the same dimensionality (either 1D or 2D) are displayed in the same plot.
There are three types of selection for the datasets in GSim:
Active data: dataset, currently activated in the ‘Files’ panel. Usually it is the last clicked dataset in the ‘Files’ panel.
Selected data: data, which is highlighted (selected) in the ‘Files’ panel. The main difference between an active and selected datasets is that multiple datasets can be ‘selected’ whereas only one dataset can be ‘activated’. Multiple selection can be achieved, say, by mouse click keeping ‘Ctrl’

button pressed. Most processing functions can be applied for the selected data in a sort of a 'batch' mode whereas other functions is applicable only for the active dataset. For example, 'Info' panel contains information only about active dataset. In contrast, invoking "Process" function will transform all datasets selected in 'Files' panel. Of course, the active dataset is always one of the selected datasets. In case only one dataset is highlighted in the 'Files' panel this set then both active and selected.
Visible data: you can set some dataset to be invisible by unchecking the check boxes for the corresponding file in the 'Files' panel. Dataset is not removed – it just becomes invisible. This is completely independent of the previous types of selection.

Markers: the red crosses, specified a certain position in the active spectrum. Depending on the particular operation they could specify a point on the spectrum, a row or a column. There are two markers: main (solid red cross) and secondary (dashed red cross). They can be set by clicking on the left and right mouse buttons respectively.

Mouse Action on the plot area:

Response to the mouse action depends which manipulation mode is selected in the plot toolbar:


In "Move/shift spectral window" mode:

Right mouse button and drag – move spectra in the direction of the drag;
Left mouse button and drag – change visible area (scaling);


In "Vertical shift/scaling" mode
Right mouse button and drag – shift the selected spectra vertically;
Left mouse button and drag – scale the selected spectra vertically;

Keeping 'shift' pressed can modifies the behaviour as described in '1D' panel Take->stack section

In "Horizontal shift/scaling" mode

Right mouse button and drag – shift the selected spectra horizontally;
Left mouse button and drag – scale the selected spectra horizontally;[1]

Keeping 'shift' pressed can modifies the behaviour as described in '1D' panel Take->stack section

[1] *Why different mode for vertical and horizontal shift/scalings ?. Shifting/scaling horizontally and vertically, in fact, are every different operations. If vertical scaling for 1D NMR means intensity scaling (data actually is not affected too much), the horizontal scaling/shift has very dramatic effect: it is equal to the changes of the spectral width or referencing alternation. So it is more safe to have separate functions.*

In “Select/Edit graphic objects” mode
User can manipulate graphic objects, associated with the selected datasets in the way, similar to vector graphics editors. There are three graphic object types supported by GSim: text, straight line and bitmap image (see below how to insert them to the plot). In the graphic mode user can change the position of the object (dragging it by mouse), change text (double mouse click on the text object), scale images and change line appearance. Some graphic object properties as colour, font and text orientation can be changed in context menu invoked by right mouse click when the cursor is over a graphical object. Multiple selection is supported either by area selection or by clicking on each object keeping “Shift” key pressed.

Removes markers from the current plot.

Set window range to be equal to the global range for all visible spectra. Removes markers. It also removes all scaling/shift alterations. In order to keep these alterations activate ‘Edit->fix shift/scaling’ function or press ‘F’ button on the keyboard.

Simple left button click – set the main marker (solid red lines)
Simple right button click – set the secondary marker (dotted red lines)

Some keyboards shortcuts:
“Y” - switching on/off *y*-axis
“X” - switching on/off *x*-axis
“M” - removing markers
“B” - draw the active 1D dataset in bold/normal
“C” - switch on/off colour representation of the 2D contour plots.

Processing capabilities

GSim contains several processing functions that should cover basic needs for typical NMR experiments. An original object-oriented approach has been implemented for invoking of the processing functionality. Any of the available processing functions can be called by one of the following way, depending on user's personal preference:

1. All processing functions are available *via* ‘Process’ menu. At first, user will be interactively asked about required processing parameters and immediately thereafter the processing function will be applied.
2. Any of the processing functions can be placed on the ‘processing toolbar’ as a toolbar button. Customisation of the processing toolbar can be done by clicking ‘Edit’ button on ‘Processing’ side panel.
3. Each processing function has its text command. Text command of the processing function is displayed in squared brackets in the ‘Process’ menu. For example, ‘DC Offset [dc]’ entry means that the function ‘DC Offset’ has a text command ‘`dc`’. For those, who likes typing their commands (special ‘*Hi!*’ for VNMR, Xwin-NMR users) they can type these text commands in a command line at the bottom of the main window. You can specify options for the command directly after it. Syntax should be the same as for the table in the

‘Processing’ side panel (see description below). For example a command ‘`ph 90.0;0`’ will invoke 90 degree zero-order phase correction in direct dimension. If no options specified a user will be asked to enter them interactively. You also can combine several commands in the command line. An example:

```
dc 10 + lb gauss;10.0 + size 8k + ft +ph
```

will apply consequently DC Offset correction using 10% of FID, gaussian linebroadening of 10 Hz, zero-filling up to 8192 points, Fourier transformation and then interactively asks about manual phase correction.

Command line supports autocompletion for most of commands. You can also re-call previously applied command using ‘arrow up’ and ‘arrow down’ keys on your keyboard.

If you type '`table:`' before your actual command sequence, the sequence will substitute the current contents of the table in the 'Processing' panel (see below) without execution.

4. Processing functions can be combined in the ‘Processing’ side panel. Probably, it is the most sophisticated processing facility in GSim. It allows to do complete processing by single button click, store and reload processing list, etc. See its description in the corresponding section below.

3. Menu actions:

File->New window

Opens a new empty plot window.

File->Open:

Allow opening several types of NMR-related files. For the description of each particular file filter see below.
If "Any known file format" is selected the program tries to guess the file format (usually from the directory structure).
Files also can be opened using file browser ('Browser' panel). Double click on the file or drag&drop it to the main plot area.
Drag&Drop to the plot area from a standard file manager application (as *Windows Explorer* on *Windows*, *Konques*t on *KDE* or *Nautilus* on *GNOME*) should also work.
When second or further dataset is about to open, the "Open file" dialog has an option to place the data in a separate window rather then on the same plot.
GSim never alters the original files. All operations are performed on data, loaded in memory – it should be safe to open files directly from the spectrometer.

MATLAB file format

Matlab file format is a container format that is able to accommodate different data types arranged in structures. It is used by proprietary mathematical MATLAB software (http://www.mathworks.com) for storing its information. GSim is expecting to find the '.mat' extension for this file format. Due to flexibility of the format, GSim is able to save practically all important information including shift/scaling, graphics objects, arrays, external parameters, integrals, etc. By this reason I am suggesting to use this file format as a 'native' GSim file format. And of course, you can load this information in external programs as MATLAB itself or an open-source analogue - Octave (http://www.octave.org).

Data, saved from GSim has a structure, described below. The string in brackets '<>' are replaced by actual names as described further. A GSim dataset is presented as a top-level structure with the name, normally identical to the filename. Because the symbol '.' is reserved in MATLAB it simply replaced by '_' in the filename.

In principle, single MATLAB file can contain several datasets and GSim should be able to load them all, however, GSim is not stored data in this way due to some technical difficulties (for instance, what 'reload' function should do for multiple datasets?).

GSim MATLAB file structure

spectrum1 *'spectrum1' is the name of the MATLAB variable where everything is stored*
- data *spectrum stored as a complex matrix (just 1 row if the spectrum is1D)*
- sw *spectral width (for each dimension if 2D)*
- type *type: 0-FID, 1-spectrum (for each dimension if 2D)*
- sfrq *spectrometer frequency (for each dimension if 2D)*
- ref *reference shift (for each dimension if 2D)*
- ph0 *zero-order phase correction (for each dimension if 2D)*
- ph1 *first-order phase correction (for each dimension if 2D)*
- scale *horizontal and vertical visual scaling applied for the dataset*
- shift *horizontal and vertical visual shift applied for the dataset*
- title *a string with dataset title*
- pulprog *the name of the pulse program used*
- is_real *1- if the real part should be visualised and 0 – if imaginary*
- is_visible *0- dataset is invisible, 1 -visible*
- has_peaks *1 if GSim should visualise 'PeakList' array as a peak marks*
- arrays *structure containing all arrays*
  - <array1> *each array has its name and a list with data*
  - ... *several arrays can be stored in the structure 'arrays'*
- integrals *structure containing all integrals*
  - integrals_scale *coefficient to convert the real integral value to the number on the screen*
  - integrals_vscale *"shrinking" factor for the visual representatio of the integral curves*
  - integrals_vshift *position of the baseline for integral curves in real coordinates*
  - integral0 *2-row matrix containing an integral curve (first row – x-values in current units, 2 – y-values)*
  - ... *several integrals can be stored*
- params *non-interpreted parameters taken from original file format*
  - <originalParameterFilename> *filename of the original parameter file*
    - <parameter0> *each parameter presented as a string*
  - ...
- graphics_object0 *graphics object (could be several)*
  - type *(0 – text, 1 – line, 2-pixmap)*
  - scale *scaling factor (used for pixmaps only)*
  - pos *position relatively of the spectrum in the spectrum coordinates (two numbers for pixmap and text, four – for line)*
  - colour *array containing red, green, blue and 'alpha-channel' componenets*
  - font *string with font description (only for text objects)*
  - maindata *main data – either text or binary data for pixmaps (PNG format)*

Chemagnetic/Varian Spinsight file format

In order to open Spinsight data you should select "data" file in the original directory. The file filter requires "acq" and "proc" file being in the same directory as the "data" file. The most useful parameters should be read by this routine, as well as the associated arrays (for example, 'tau' array for spin-echo), if they have been used in the experiment.[2]

[2] *NB: spectrometer frequency could be misread, especially for 2D spectra: you can always correct it in "Spectral parameter" dialog.*

Bruker Xwin-NMR / Top-Spin

For Bruker Xwin-NMR / Top-Spin file format (**FID**s) select ‘fid’ or ‘ser’ file. Spectral parameters are taken from ‘acqu’ and ‘acqu2’ files. Title and processing parameters are taken from ‘pdata’ directory. If the Bruker dataset contains more then one processing directory user will be prompted to select one of the sets.
For Bruker Xwin-NMR / Top-Spin file format (**1D processed data**) select ‘1r’ file. Spectral parameters will be taken from ‘proc’ file.

Varian VNMR

For VNMR a ‘fid’ file should be selected. The filter is also capable to open a raw FID file (fid) from the Varian VNMR(J) current working directory (called exp*N* where *N*- the workspace number). It also has an experimental support for processed data from files 'data' and 'phasefile'.

Simplot

This format doesn’t store a type of data in indirect dimension. Correct it, if required, in “Spectral parameter” dialog.
Paul Hodgkinson’s ‘pNMRsim’ simulation program (see at http://www.dur.ac.uk/paul.hodgkinson/software/pNMRsim) is able to store initial input file in SIMPLOT-type output if ‘save -source’ command has been used. GSim reads this data and stores it as a parameter ‘source’ which can be retrieved via ‘Edit->Spectral parameters’’ menu entry.

Spinevolution

SPINEVOLUTION output is a simple text file which doesn’t preserve any of the spectral characteristics. It is advised to use “Spectral parameters” dialog after opening SPINEVOLUTION ‘.dat’ file and correct missing parameters. If Spinevolution output contains two parts: real and imaginary, they are usually saved in files <basename>_re.dat and <basename>_im.dat. If you are trying to open a file which name ends with ‘_re.dat’, GSim will try to find a part with ‘_im.dat’ and construct a united single dataset with complex points.

Castep

Castep is *ab initio* DFT computer code which uses plane-wave basis set and pseudopotentials: see "*First principles methods using CASTEP*" Zeitschrift fur Krystallographie **220**(5-6) pp. 567-570 (2005). S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, M. C. Payne. (www.castep.org) for details and availability.

Recently an NMR module has been added to CASTEP which allows computing a number of NMR parameters for crystalline solids. GSim filter for CASTEP is able to load isotropic chemical shifts from CASTEP-generated ‘.magres’ files. It is not a real NMR format, it just contains a list of NMR parameters as isotropic chemical shifts or CSA. Compatibility with different versions of CASTEP therefore is not guaranteed. User will be asked several NMR-related questions before a spectrum can be generated:

Choose nucleus: a list of nuclei found in CASTEP file will be shown. Select nucleus you want to see.

Choose NMR frequency: enter the value of NMR frequency for the selected nucleus. This parameter is used to convert values from ppm to Hz-scale.

Enter number of points: total number of points in the generated spectrum.

Estimated shift of the standard: CASTEP calculates absolute chemical shielding. To compare it directly with experiment, it should be converted to the chemical shifts using an appropriate reference shielding. The isotropic chemical shift, $\delta_{iso}$is given by

$$\delta_{iso} = -\sigma_{iso} - \sigma_{ref}$$

where $\sigma_{ref}$is the nuclear shielding of referencing system. For instance, in the case of ^{13}C, a full simulation of liquid tetramethylsilane is required. Of, course there are a lot of different ways to avoid these time-consuming calculations. For instance, in PCCP, 2006, 8, 137-143 the value of $\sigma_{ref} \approx 168 - 169ppm$has been found for ^{13}C from linear interpolation of computed values *vs.* experimental ones.

Linewidth: the linewidth of the each generated line. Lorentzian lineshape is used.

Spectral width and the position of the spectral centre are automatically adjusted to include all lines found in the file.

Average shifts for: is an optional string which can contain a list of atoms which chemical shifts should be averaged. As an example all proton shifts in CH3 groups can be averaged because of the motion. The list should contain atoms with indexes (as appears in .magres file) separated by comma. List can contain more than one group of atoms – in this case the groups should be separated by ";". For example: a string "H1,H2,H3;H20,H21,H22" means that the chemical shifts of atoms H1, H2 and H3 will be assigned to an average value as well as the shift for H20,H21,H22 atoms.

File->Save

Save active spectrum. Supports Matlab, Simplot and Spinsight formats.
Simplot format doesn't store information about a type of data in indirect dimension (see File->Open description).
Spinsight require a creation of a new directory. Directory name should be a completely new, rewriting is not supported so far.

File->Reload

Reloads selected datasets from the disk (if they exist there).

File->Watch

This function regularly checks whether the file which corresponds to the active dataset has been modified on the disk. In case of the modification dataset is being reloaded. The interval between checks should be provided by the user when 'Watch' command has been activated. The background of the plot containing the watched dataset turn to yellow during 'Watch' operation. This command should be useful for control, say, of the acquisition in progress. For instance, Chemagnetic Spinsight sofware stores intermediate data in files 'lastnd' and 'lastfid' which can be watched by GSim. Obviously, the usability of the 'Watch' strongly depends on the acquisition software behaviour.

File->Move to

Moves selected dataset to another plot window. The plot window should already exist. Use "File->New window" to create an empty one. This function is doing nothing if only one window is present.

File->Close

Delete selected spectra.

File->Export to PDF

Export current picture from the active plot to the PDF file.

File->Export to vector graphics

Very often one needs a vector format file “for editing” purposes, e.g. preparing a publication. This function allows to save the current screen as a file in Scalable Vector Graphics (SVG), Encapsulated Postscript (EPS) or Enhanced Metafile (EMF) formats. SVG output is a little bit more correct at the moment then EPS or EMF (at least, it can stores bitmap pictures correctly).

I can advice to use open-source and free (as freedom and as beer) ‘Inkscape’ editor (www.inkscape.org) to edit SVG pictures.

Adobe Illustrator should be able to open GSim-generated EPS files. Illustrator 10 and later has a good SVG support too. As far as I know recent versions of Corel Draw and MS Visio have SVG support too. EMF format is a standard vector graphic format for MS Windows and can be opened by almost every Windows program with graphical capabilities (including MS Office suite).

Because this export function has been created ‘for editing’ purposes, its output isn’t so nice as ‘for printing’ and requires extra manual ‘clean-up’.

File->Print

Print a copy of the screen.

File->Print Preview

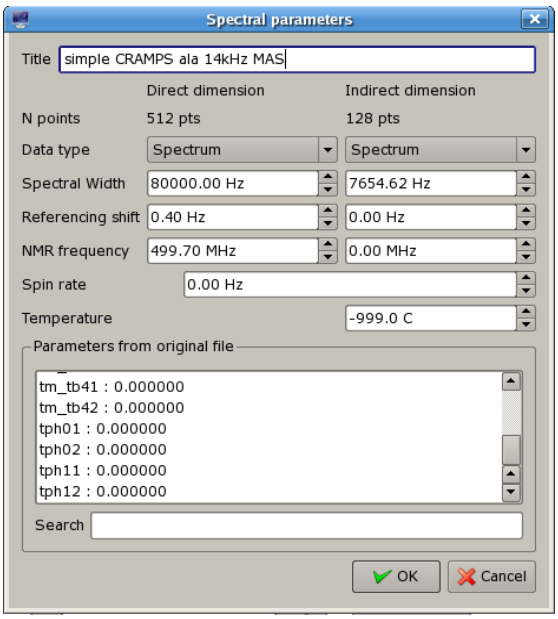
Opens print preview dialog

File->Exit

Exit the program.

Edit->Spectral Parameters

Allows user to edit spectral parameters for the selected spectra, such as spectral width, *etc*. GSim also loads spectral parameters from SPINSIGHT, VNMR and Bruker files. It can be viewed and searched by user.



Edit->Undo

Undo the last action. Only single “undo” step can be done.

Edit->Copy screen

Copy active plot into system clipboard. Picture is stored as a bitmap with the screen resolution.

Edit->Take inset

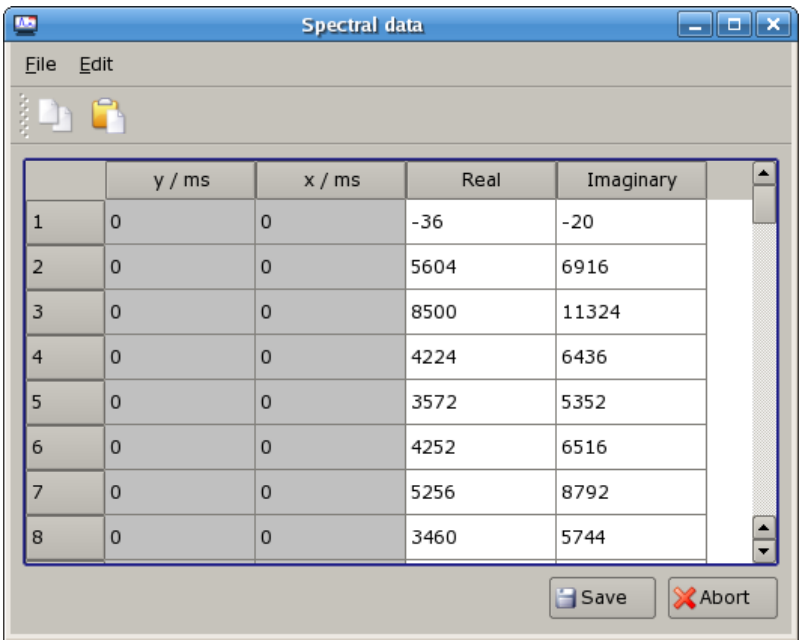
GSim uses the most lazy way for creating insets. Zoom the spectrum as you wish to appear in the inset and activate ‘Edit->takeInset’ function. It will take the screenshot and set it as a bitmap image, associated with the active dataset. The inset can be altered in ‘Select/Edit graphic object’ mode. Resolution of the inset is equal to the the screen resolution, although it appears to be visually scaled down.

Edit->Paste Image

Paste a bitmap image form the system clipboard as a graphic object associated with the active dataset. You also can simply darg&drop an image from the file manager or your graphical application on the plot area.

Edit->Data

Opens the window of data viewer for the active spectrum. In this window you can directly edit any data point. Nice tool for cheaters :-). When you selecting any point in the data table, the main marker in the plot window is automatically placed at the position of the selected point. Ctrl-C copies selected table region into clipboard for export to spreadsheet editors as MS Excel or OpenOffice Calc. You can paste data from clipboard pressing Ctrl+V, or selecting Edit->Paste menu of the data viewer.



Edit->Fix shifts/scalings

If spectra are scaled/shifted by mouse operations, this command is trying to keep it as permanent parameters. Actually it is doing following things:

For 1D NMR: multiplying each data point by vertical scaling and adding vertical shift to them. Multiplying spectrum width by horizontal scaling and adding horizontal shift to reference parameter. As an FID always starts from zero, the horizontal shift for FID cannot be fixed by this way – it's simply ignored.

For 2D NMR: Vertical shift scaling is treated as horizontal (see above) but for indirect dimension.

This function could be used, say, for interactive referencing. Simply move desired line to the position where it supposed to be (for instance, ^{13}C TMS line to 0 Hz/ppm) and activate 'Edit->Fix shifts/scaling's. Alternative and more "classical" way for referencing is 'Edit->Referencing' menu entry (see below).

This function is applied for all open spectra.

View->Real / View->Imaginary

Switch between real and imaginary part.

View->Referencing

This is a "classical" way of spectrum referencing. Switch to the wanted (Hz/ppm) ruler. Simply put the main marker at the position with known chemical shift or frequency and call this function. The desired referencing value(s) will be asked (2D case: for each axis; 1D case: only once for horizontal axis). This function can be applied for all selected datasets simultaneously. However, all selected datatsets should have the same dimensionality (either 1D or 2D).

View->Set Range

Interactively asks about viewing boundaries and set them according to the user input. Alternatively, the boundaries can be taken from the markers position.

View->Ruler->ppm/Hz

Changes the ruler for the active plot area between ‘Hz’ and ‘ppm’ (1D or 2D). Switching to ppm requires all spectrometer frequencies to be set.
There is a more simple ways to do this. You can either change the rulers in View->Options->Axis menu or simply click on the label directly in the plot. If you have a problem to switch to ‘ppm’ scale that means one or more opened dataset has no spectrometer frequency set properly. Change it in ‘View->Options->Axis’ menu.

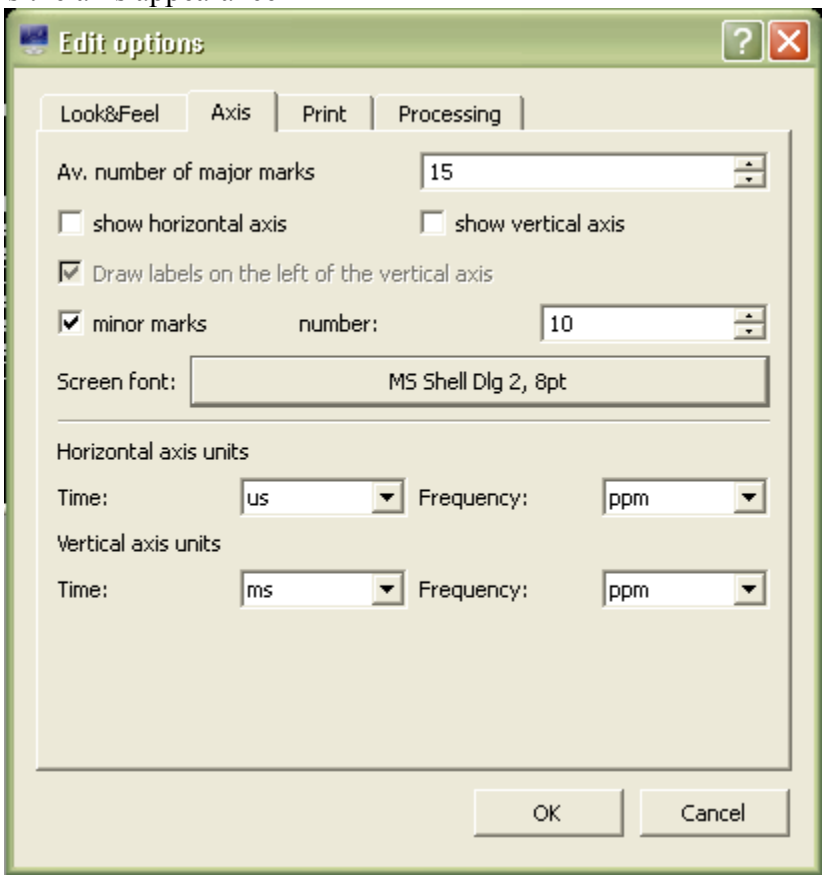
View->Options

Edit program options

Tab ‘Look&Feel’ contains options:
‘Style’ - changes the GSim style. Uses styles provided by Trolltech Qt library and depends on the computer system used.
‘Colour scheme’ specifies a colour scheme for the plot area.
‘Antialiasing’ - set antialiasing (smoothing of the edge of graphics) for the plots. Can be slow on low-end systems.
‘Smooth image transformation’ - use smooth transformation for bitmap graphic objects. For an example, insets doesn’t look very good if this option is not set. From other hand this operation can be slow on low-end machines.
“Show active 1D spectrum in bold” - if checked then the active 1D spectrum will be bolder then others. This function is on by default. It helps to keep a visual track which dataset is active, however, it could slightly obscure fine features as peak multiplicity in solution-state NMR.
“Superimpose FID and spectra” - if checked the FIDs and spectra (either 1D or 2D) will be displayed simultaneously. Otherwise Gsim will overlap data with equivalent types only. For 2D case the data types in both dimensions should match. *Note: in any case the absolute levels for 2D display are the same for all 2D datasets. If this option is active that could lead to apparently 'empty' display for dataset with lower intensity.*
“2D projections”. Chooses the appearance of the 'usual' projections (see 'Smart sideviews' feature for the explanation what is the usual projections). Can be positive 'skylines' or, in other words, the maximum, taken from the corresponding row/column. Another option is a simple sum of rows/columns.
“Animated markers” switch on the animation of the label on the main and secondary marker position.
“Line thickness” - determines the sickness of the line on the printouts. Sometimes can be needed for high-resolution laser printers.
“Use original parameters if available” - tries to insert parameters as in vendor's software. The code has an experimental status.
“Use external image (SVG or bitmap) as a print logo”. You can use a custom logo, which appears at the top right corner of the PDF output and on the printouts. For instance, you can put you group/university/company logo or even your personal photo! In order to ensure the best picture quality and small size in outputs the logo can be saved in SVG vector graphic format (see “File->Export vector graphics” for list of supported software). But don’t be too sophisticated – GSim (or Qt-library to be precise) support the minimalistic SVG set. For instance, convert fonts to pathes to

ensure the same appearance on different computers. Other bitmap graphics format supported by Qt also can be used. If the file name field is empty a default GSim logo is used.

Tab ‘Axis’ controls the axis appearance

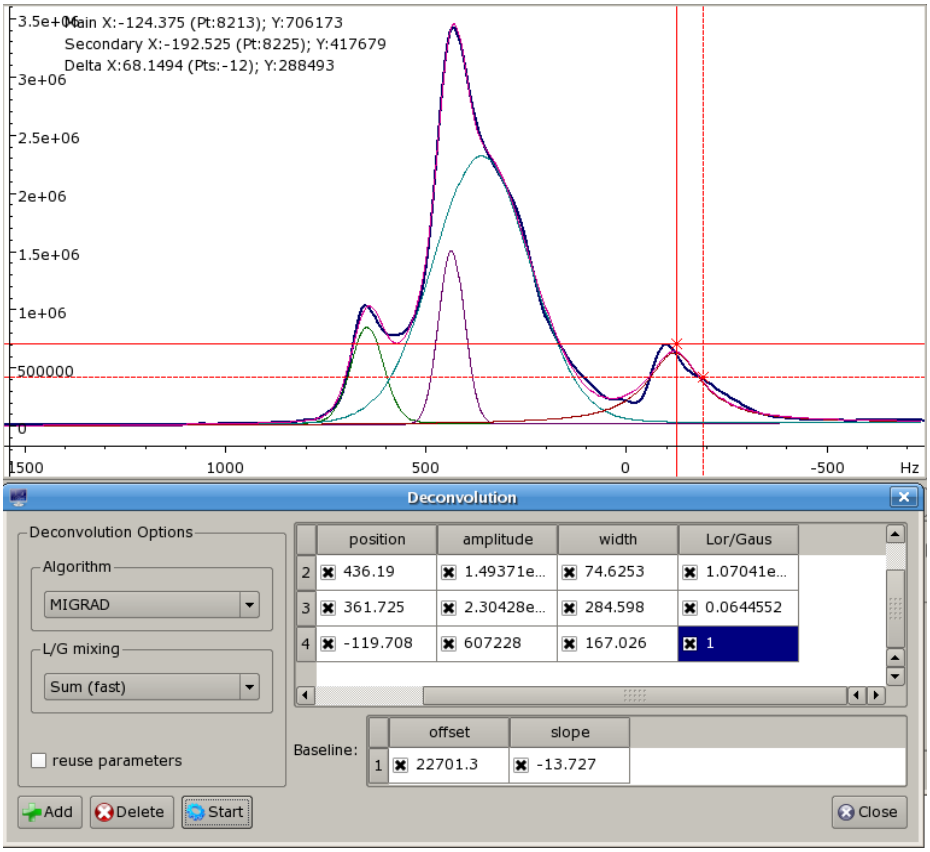


Average number of major marks controls how many major marks will be displayed on each axis. GSim has a smart algorithm to choose an actual axis step but keeping the overall number around specified here.
Show horizontal/vertical axis controls visibility of the axis on plots.
Vertical axis can be drawn with labels on the left of the axis (good for preparing publictions) or on the right (better mode for the screen). The check box “Draw labels on the left of the vertical axis” controls this behaviour.
Each major step can be subdivided into minor steps with total number specified in minor marks number box.
Screen font button allows to change the font, used in the plot area.
Each axis has two default units for frequency and time domain. All open datasets are displayed in this units. For frequency domain: if dataset cannot be displayed in ppm, scale is automatically changed to Hz.

Tab “2D processing”
Here you can choose whether the hypercomplex data should be saved or not during 2D data processing.
Saving hypercomplex data allows you to do some extra operations (for example, correct phases of the direct dimension of a 2D spectrum, Fourier transformed in both directions). But it requires as twice as more memory and make most of the 2D operations 2 times slower.
Hypercomplex data can be viewed via Edit->Data menu entry. Only few processing functions recognise hypercomplex part at the moment: FT, phasing, shift, inversion.

Analysis->Deconvolution

Deconvolution (peak fitting) procedure. It works only on 1D spectra (although it could be an extracted row from 2D spectrum and GSim can do automatic deconvolution for corresponding 2D dataset row by row). First, set the main marker on the top of the expected line. Secondary marker should be set on the half intensity of the line on either (left or right) side. GSim would try to guess where is this position and set the secondary marker accordingly. Correct it if the automatic procedure has failed. Thus, the markers specify the initial parameters: amplitude of the line (*y* position of the main marker) and its width (difference between *x* positions of the main and secondary markers). Pressing "Add line" creates a new line in the table. The created line is immediately appears on the screen. Add all desired lines by this way. You can correct initial parameters by editing corresponding cells in the table. Screen will be updated automatically reflecting you input. You can choose some parameters to be fixed during fitting. Simply uncheck them from the table. The baseline offset and slope could be fitted as well. If you don't want to do this, simply uncheck them from the baseline table. If data is a row extracted from 2D dataset and you would like to apply fitting to the whole 2D spectrum than two modes can be chosen. (1)Initial parameters as specified in the table could be reused for each consequent row, or (2) for each next row of the dataset the optimised parameters obtained in fitting of the previous row are used. Specify this by checking/unchecking of the "Reuse parameters" checkbox. Pressing "start" button the deconvolution is executed. Only the visible part of the spectrum will be used for the deconvolution. If the active spectrum is a row extracted from 2D dataset than user will be prompted either apply deconvolution to this row only or to entire 2D dataset, row by row. As result of the deconvolution the individual model lines and their sum are created as independent datasets. In the case of 2D these datasets are checked as invisible in order to not affect the overall program performance. The result of deconvolution is stored as arrays for each resulting line. Dataset 'sum' contains arrays with all results of the deconvolution as well as χ^2value. Normally user should start "Array manager" after deconvolution if (s)he likes to analyse the deconvolution results.

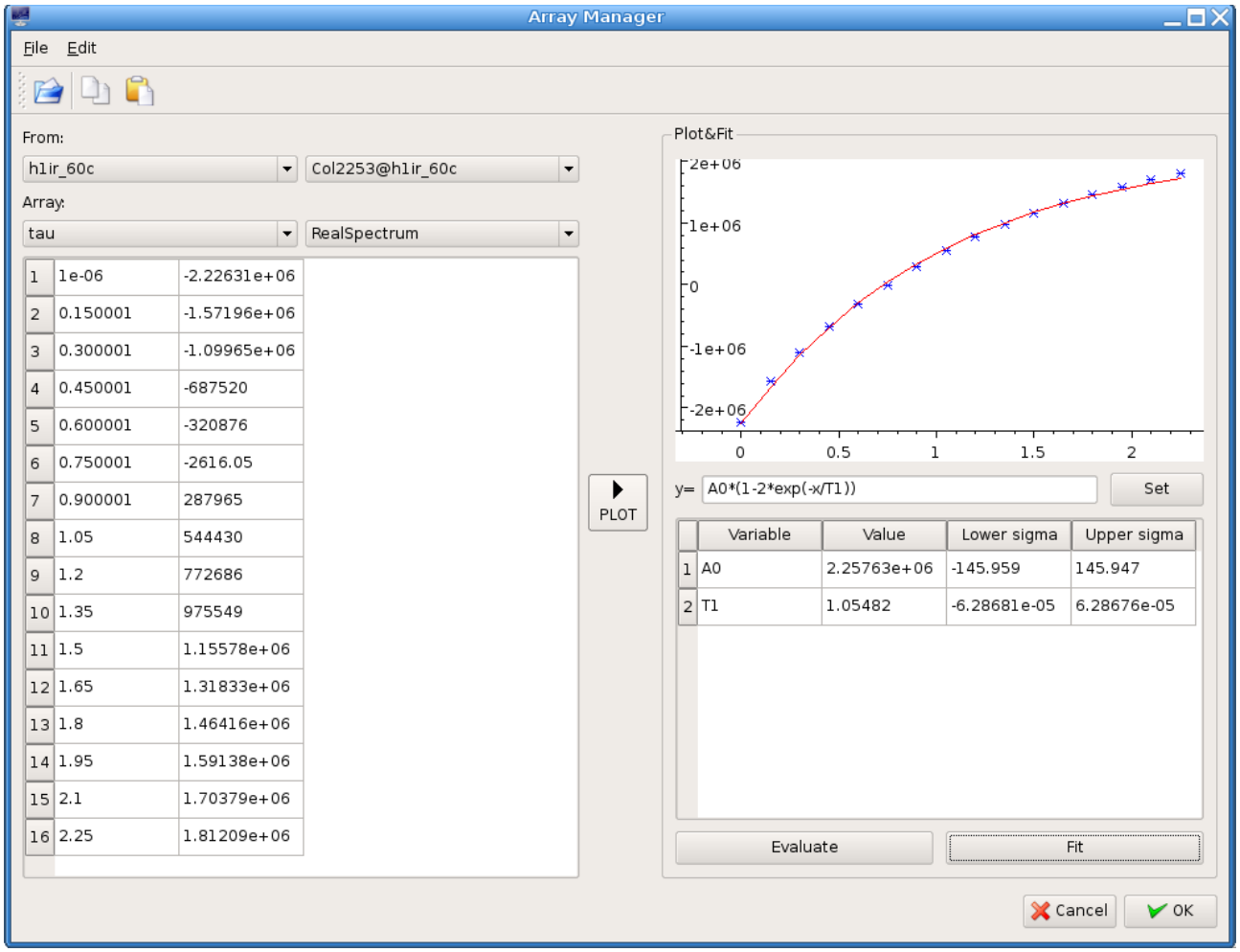


GSim uses 'MinUIT2' library for fitting (http://project-mathlibs.web.cern.ch/project-mathlibs/sw/html/Minuit2.html).
Three fitting algorithm can be used: MINGRAD, simplex or combined. From MinUIT documentation:
'MIGRAD' is the most efficient and complete single method, recommended for general functions.
Simplex is a function minimization method using the simplex method of Nelder and Mead. As 'SIMPLEX' is a stepping method it does not produce a covariance matrix.
'Combined' method causes minimization of the function by the method of MIGRAD but switches to the SIMPLEX method if MIGRAD fails to converge."
User can also choose a method of the Lorentzian/Gaussian mixing for the line.
Let assume that Lorentzian and Gaussian lineshape with amplitude equal to 1 can be generate in functions: $lor(\nu,\omega)$and $gau(\nu,\omega)$where ν- frequency of the line and ω- its width, then method "SUM" corresponds to:
$A \cdot (1-f) \cdot gau(\nu,\omega) + A \cdot f \cdot lor(\nu,\omega)$
where A- the intensity of the line and f – Lorentzian/Gaussian fraction. In this case fraction f actually represents the contribution to the *intensity*. This method is fast and used in the most NMR processing programs. However, it is not completely correct form the theoretical point of view.
More correct method involves the convolution of the lines:
$A \cdot lor(\nu, \omega * f)cplus \wedge gau(\nu, \omega * (1-f))$
In this case the fraction *f* represents the contribution to the *linewidth*. However, this method requires numerical convolution and therefore much slower and less robust the method "Sum". Its GSim implementation is at experimental stage.

Analysis->Array manager

Each NMR spectrum can contain associated arrays. For instance, if 2D dataset contains a relaxation measurements then an array can contain delays. For nutation experiment it could be pulse lengths

and so on... Literally speaking, each commercial program has its own implementation of arrays. For example, in Xwin-NMR they are stored, say, in vdlist (delays), vplist(pulse length), vtlist(temperatures) etc. GSim can create its own arrays, for example after deconvolution procedure or peak picking (see above).
Array manager allows viewing and copying of this arrays as well as more sophisticated plotting and data fitting. “Copy-Paste” should work with major spreadsheets applications as MS Excel or OpenOffice Calc. You can store all available arrays in one single file in Matlab 5 format using “File->Export to MatLab” menu. Each array in MatLab appears as a vector with name ‘<filename>_<arrayname>’. For example, array “vdlist” from dataset “adamantane_1” will appear under name “adamantane_1_vdlist”.
Arrays can be loaded from an external text file. The file should be a simple text file with one number per line. Suffixes ‘s’, ‘us’, ‘ms’ can be used after each number (without spaces) to determine seconds, microseconds and milliseconds units (Bruker-like syntax). However, all data will be converted to milliseconds when loaded.
Whereas you can “Copy-Paste” arrays to more specialised programs as commercial ‘*Origin*‘ (http://www.originlab.com) or free ‘*QtiPlot*’ (http://http://soft.proindependent.com/qtiplot.html) and do statistical analysis there, Gsim itself contains some plotting and fitting features.



1	1e-06	-2.22631e+06
2	0.150001	-1.57196e+06
3	0.300001	-1.09965e+06
4	0.450001	-687520
5	0.600001	-320876
6	0.750001	-2616.05
7	0.900001	287965
8	1.05	544430
9	1.2	772686
10	1.35	975549
11	1.5	1.15578e+06
12	1.65	1.31833e+06
13	1.8	1.46416e+06
14	1.95	1.59138e+06
15	2.1	1.70379e+06
16	2.25	1.81209e+06

Let discuss these features on an example of the fitting relaxation data. Suppose you have a pseudo-2D relaxation experiment where each row associated with delay value, stored in array called 'tau'. You can use a following protocol for obtaining relaxation parameters:

1. Process direct dimension of the original data
2. Take a column near the top of the line you want to analyse. New 'colNNN@..” file will

contain the intensities.

3. Call Array Manager
4. Select dataset with extracted column for the left column in the main table.
5. Use 'File->Create special->from spectrum' menu entry. This will create several arrays associated with this dataset.
6. Assign 'tau' array from the original dataset to the left column of the main table and 'RealSpectrum' array from the extracted 1D dataset for the right column.
7. Press 'PLOT' data. This will display peak intensities as a function of the tau-delays.
8. In the string 'y=' type an expression for the theoretical curve. Argument of the function should be 'x'. GSim will complain if the function would not contains 'x'. For example, for inversion-recovery experiment you should type something like: A0*(1-2*exp(-x/T1))
9. Press 'Set' button. Array Manager will analyse the given expression and generate a table of variables.
10. Set initial values for the variables in the table. You can visually estimate how good is these values by pressing 'Evaluate' button
11. Press 'Fit' button. If fitting is successful, the values in the table will be substituted by optimal values. Two extra column in the variable table shows standard deviation for each variable in negative and possible range. For non-linear fitting in general there are not equivalent. Whereas most of the fitting programs ignore this fact, Gsim is able to use MINOS algorithm from MinUIT2 library under Unices[*], which is doing accurate analysis of the errors. See MinUIT documentation for details.

You also can use an integral intensities of the peaks instead of their amplitude. When you are doing an integration of 2D data, Gsim automatically creates arrays 'IntegRR*XX*' where *XX* is the number of the integral starting from zero. It contains the integrals value, taken from each row in the integrated area. It could be used directly for the fitting.

Deconvolution procedure also creates arrays with integral intensities which can be used for fitting.

In addition to standard function list, presented in 'muParser' library (*e.g.* functions like **sin**, **cos**, **exp**) a couple of new functions has been added:

redor(*x*,*d*) – calculates REDOR response (S/S_0) as a function of the time *x*. *d* - dipolar coupling (Hz). The “redor” processing function can be used for the preparation of the experimental data for fitting in the Array Manager.

abs(*x*) – takes absolute value of *x*.

Analysis->pNMRsim Simulation

Basic concepts

pNMRsim is a general solid-state NMR simulation program similar to SIMPSON or SPINEVOLUTION. It can be freely downloaded from http://www.dur.ac.uk/paul.hodgkinson/pNMRsim. GSim provides a simple inteface to this program allowing to set some simlation parameters interactively and see result of the simulation immediately. One of the area where pNMRsim integration can be very useful is fitting of the experimental NMR spectra using complex NMR-related models. A number of computer programs are dedicated to fitting experimental static or MAS CSA, dipolar or qudrupolar pattern. But when the model is physically ‘build-in’ in the program itself the user can find at some stage that he needs more sophisticated model in order to fit his experimental data. Due to this reason GSim doesn’t contains any sophisticated fitting procedures (apart from the spectrum deconvolution using

[*]Due to a bug in 3rd party MinUIT2 library, MINOS doesn't work under Windows. The Windows version of GSim reports the errors taken from the conventional gradient method. By this reason the upper and lower errors are identical in this version.

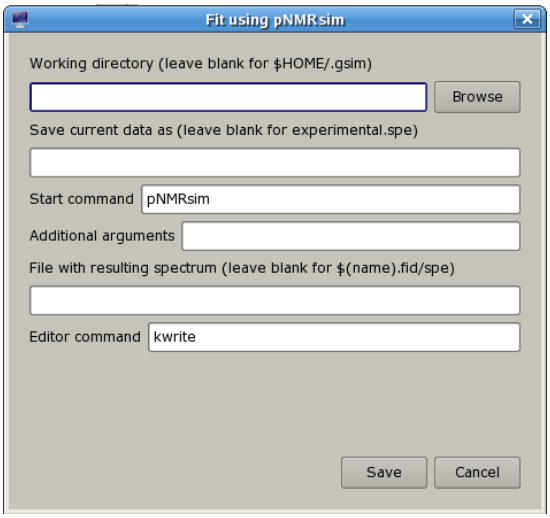
gaussian/lorentzian lineshapes) but provides a way to use external tools which can do this work with better quality.
GSim interface for pNMRsim is not restricted to the fitting problems only and can be used for any interactive simulation using pNMRsim.
In order to use pNMRsim together with GSim user should do following steps:

1. Prepare pNMRsim input file (with extension .in) with GSim instructions and put it in 'working directory'.
2. Invoke 'Analysis->pNMRsim simulation' and set up parameters.
3. Start simulation. Results normally should appears in GSim as soon as the simulation is completed.

Setup GSim inteface for pNMRsim

In order to call setup dialog press 'Setup' button after starting 'Analysis->GSim simulation'.

Working directory is a filesystem directory where GSim is looking for pNMRsim input files. Files



should have an extension '.in' and be prepared as discussed below. If the 'Working directory' field is left blank then a default directory will be used. For instance, for Linux it will be something like '/home/jsmith/.gsim' or for Windows 'C:\Documents and Settings\JSmith\.gsim. Note that the direcory is 'hidden' on Unix-systems.

GSim can save the current active dataset in working directory before starting pNMRsim simulation. It is needed, for example, for fitting procedure when pNMRsim loads experimental spectrum and trying to fit it. If the corresponding field is empty the file 'experimental.spe' will be saved. The saved file is in SIMPLOT format.

Start command is used to start pNMRsim. Change it if pNMRsim is not within your normal executable path.

'Additional arguments' field is used to provide command line arguments for pNMRsim. For instance, '-disable:phasemodulation' will force phase modulation algorithm to not be used. See pNMRsim documentation for details. There is no needs to provide input file name here as this is done automatically.

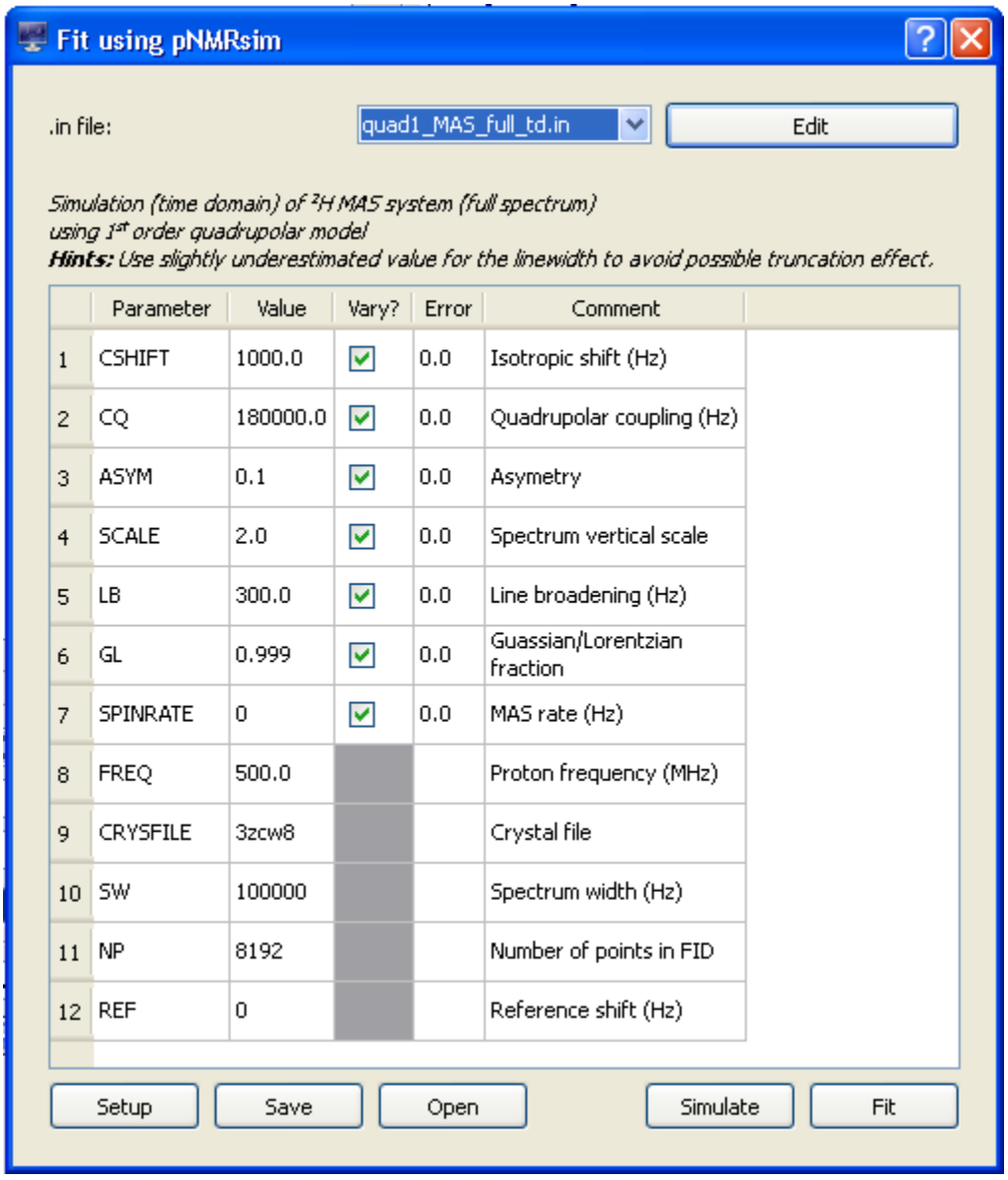
'File with resulting spectrum' field specifies which file should be opened when simulation has done. By default, the file with the same name as the input file and with extensions .fid or .spe will be loaded. If both .fid and .spe files are present, the .spe file is used.

Editor command contains command to start editor which should be used for editing the input file.

Clicking ‘Save’ the setup parameters will be saved.

Running simulation

In order to start simulation the user have to choose one of the input files placed in the working directory using the pull-down menu. Simulation parameters specified in the file should appear in the table. Change the values as required. Some scripts can work with 'arrayed values' (see next section). In this case several values, separated by comma can be assigned for the same parameter. Parameters values can be saved in the separate file and loaded later using ‘Save’ and ‘Open’ buttons. When loading GSim checks if the current input file is identical to that used when parameters were saved. If not, the warning message will appear, however, all parameters with the names identical to those in saved file will be loaded.



	Parameter	Value	Vary?	Error	Comment
1	CSHIFT	1000.0	☑	0.0	Isotropic shift (Hz)
2	CQ	180000.0	☑	0.0	Quadrupolar coupling (Hz)
3	ASYM	0.1	☑	0.0	Asymetry
4	SCALE	2.0	☑	0.0	Spectrum vertical scale
5	LB	300.0	☑	0.0	Line broadening (Hz)
6	GL	0.999	☑	0.0	Guassian/Lorentzian fraction
7	SPINRATE	0	☑	0.0	MAS rate (Hz)
8	FREQ	500.0			Proton frequency (MHz)
9	CRYSFILE	3zcw8			Crystal file
10	SW	100000			Spectrum width (Hz)
11	NP	8192			Number of points in FID
12	REF	0			Reference shift (Hz)

Field ‘Vary?’ in the table specifies should or not the parameter be varied during simulation. It is identical to appending ‘V’ to the value of the parameter in pNMRsim input file.

You can also specify the error margins for parameters that can be varied. The error values are used

by pNMRsim in order to determine the initial steps which can be applied for the parameters. If an error is set to zero then default value is applied which is calculated as a 10% of the parameter initial value. However, if the initial value is equal to 0 itself, then user must specify error manually.

Clicking ‘Simulate’ the actual simulation can be started. pNMRsim output is displayed in the GSim window. As soon as simulation is complete, the resulting spectrum should appear in Gsim. Fitting of the exisiting data can be done by clicking 'Fit' button

Preparing input file for use with GSim

In order to transfer parameters to/from GSim the pNMRsim input file should be prepared in a special way. pNMRsim is able to use environment variables as parameters and GSim interface is based on this feature. See pNMRsim documentation for details. Briefly, a variable in the input file will be considered to be ‘system’ if its name contains only capital letters. Examples are: $SPINRATE, $GAMAANGELS, etc. In order to be visible in GSim the variable should be declared in special string like this somewhere in the input file:

##par:SPINRATE;10e3;variable;MAS rate

Prefix ‘##par:’ specifies that the string contains a parameter for GSim interface. SPINRATE is the name of the variable used in the input file. ‘10e3’ specifies the default value. Next string can contains either ‘variable’ or ‘fixed’ field. Field ‘variable’ means the value can be optimised in fitting procedure. For this variable the ‘Vary?’ check box will appear in GSim interface. ‘fixed’ means the row in the GSim interface will not contain checkbox in the ‘Vary?’ field. Last entry is a comment string. Note that all fields are listed without spaces between them. However, spaces can be used in the last comment string. Each ‘##par:’ string ***must*** contains 4 mentioned entries separated by ‘;’.

Other type of string which GSim can use is a string like this:

##title:Simple test calculation on MAS

The contents of the string after prefix ‘##title:’ will appear as a description in the GSim window. Input file can contains several ‘##title:’ strings which will appear as separate lines in GSim interface window.

Now we can explain how GSim inteface is working. When .in file is selected from the pull-down menu, GSim is analysing its contents looking for ‘##’ prefixes and placing parameters to the table and titles to the description string. When ‘Start’ is clicked GSim sets parameters as system variables and starts pNMRsim simulation with current input file. As far as all Gsim-related strings starts from ‘#’ they simply treated as comments by pNMRsim itself.

In some cases you can use arrayed variable for some parameters. To set up array a user should input the values, separated by comma in the simulation table. Pressing the 'Start' button these values will be either directly transferred to the script if they are not variable (fixed or the corresponding box 'Vary' is unchecked) or symbol 'V' will be added at the end of the each value. In order to use these mechanism, the system variable should be treated as an array in your script. For example

shift 1 ($CSHIFT)

with $CSHIFT=100,200 will setup '()'-type array in pNMRsim.

Process

Process menu is a method to perform processing operation which is complementary to using “Processing” side panel. It is convenient to use “Process” menu if you want to apply a single processing operation rather then a series of them where “Processing” panel is much more useful.

“Process” menu contains a link to all available processing function regardless are they included in the table of the “Processing” panel or not.

Activating any of the entries in “Process” menu is actually equivalent to the consequent clicks to the corresponding interactive input button (see about “  ” button in “Processing” panel section) and “apply” button in “Processing” panel. For more information about available processing functions see section with “Processing” panel description.

User can place some of the processing functions as a button on processing tool bar. Customisation of the toolbar can be done using ‘Edit’ command on ‘Processing’ panel. Pressing any of these buttons are equivalent to activation of the corresponding function in ‘Process’ menu.

Windows

Gives a list of the present plot windows and performs a switch between them.

Help->Documentation

Opens PDF documentation. A PDF viewer (for example, Adobe Reader) should be installed. In case of custom installation check that file README_GSIM.pdf is located in the same directory as GSim executables.

Help->Quick Start

Opens ‘Quick Start’ tutorial. File ‘quickstart.pdf’ should be present in the program directory.

Help->Go to webpage

Opens GSim web-page (http://gsim.sf.net) using the default system web-browser.

Help->About

Logo, program info...

Help->About QT

Information about QT-library used by the program.

4. Plot Tool Bar:

“Grid on/of” - add grids on plots

Three plot manipulation buttons: see above in “Mouse action” section.

GSim is able to do simple drawing operations. Each graphic object is associated with the particular dataset. Clicking on buttons, described below creates an object associated with the active dataset.

Set the main marker at the position where you want to add a text string. Then press this button. You

will be asked to enter text string then. Text position and contents can be changed in ‘Select/Edit graphic mode’

Draws a line between positions of the main and secondary markers. You have to set both markers before pressing this button. The line position and the positions of the beginning and the end can be changed in ‘Select/Edit graphic mode’.

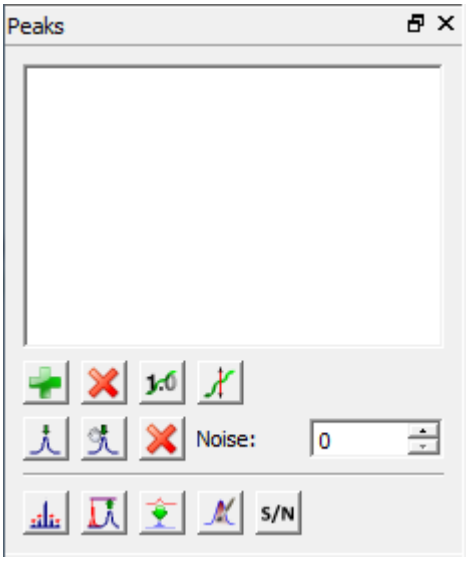
Draw bitmap images setting the top left corner at the main marker positions. User will be asked to open an image from the file in one of the supported formats. GSim relies on Qt-library when retrieving the list of supported formats. For example, Qt can be compiled with extra support of GIF files (not by default). ‘Save to vector formats’ only partially recognises the presence of bitmaps in GSim. EPS output simply ignores any bitmaps. SVG output has a more complete implementation. The image position and scale can be changed in ‘Select/Edit graphic mode’. See also ‘Edit->Paste image’ description for other ways to attach image to the graph.

“Reset all” - rescale window to fit all spectra, resetting all scaling/shifting to 1 and 0 respectively, remove all markers.

5. Operation, performed by side panels:

Panel "Peaks"

This tab is supposed to be used for some operations which analyse spectral peaks, like integration or peak peaking.



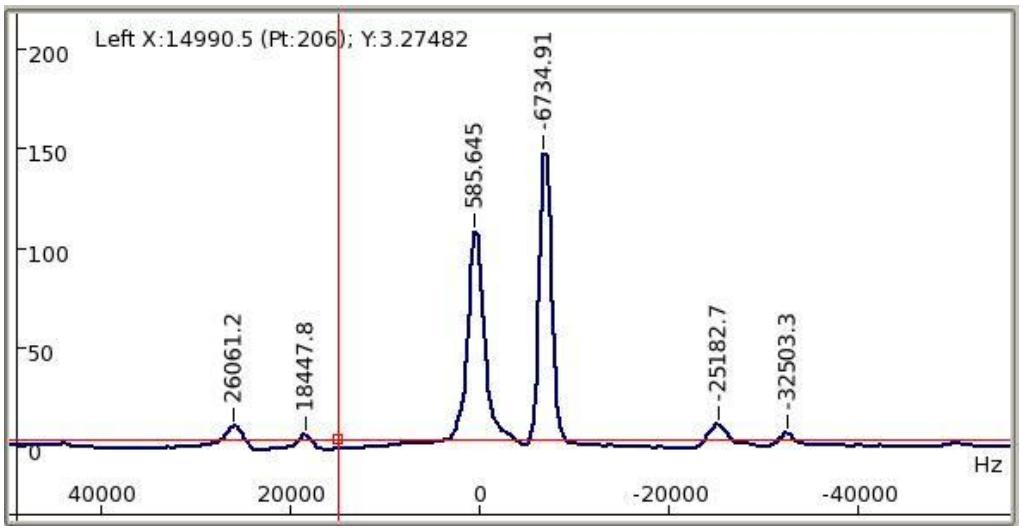
In case of 1D spectra, "Set integrals" button creates an integral curve for the line of the active 1D spectrum, located between the main and secondary markers. Info window shows extra information about selected line, as moments of line, linewidth calculated from general linewidth theory and measured directly at the half height of the maximal intensity.
For 2D spectra it calculates integral within a rectangle, specified by the markers. In this case info window contains only total peak intensity.
In both cases total peak intensities can be found as an array "IntegList", assocoiated with the corresponding data plot. For 2D data an array "IntegRR*XX*" also contains integrals taken from each row within the integration area for the integral *XX*.

"Delete integrals". If the main marker is set, then the prompt dialog should appear asking what integral(s) you want to delete. 'All integrals' delete all integrals whereas 'Selected only' would delete integral form the active dataset which is closest to the horizontal position of the main marker. If marker is not set, 'All integrals' option is assumed.

"Calibrate" button allows changing the integral values displayed in the plot. First set the primary marker within the range of the integral to be changed. Then press "Calibrate" and type new value for the the selected integral. All other integral values in the same spectrum will be scaled accordingly.

Opens a new dialog where the vertical position and the vertical size of the integral curves can be adjusted. Also tilt and slope of the integral can be set here. Relevant for 1D integration procedure only.

“Automatic peak peaking.” Set main marker above threshold level (noise level) and press “Peak picking” button. “Noise factor” box controls the sensitivity to the noise at the top of the peak. If noise factor is more then 0 then ‘+-noise factor’ points are averaged (smoothed) before maximum search. List of the peak positions (in current units, Hz or ppm) is stored as an array and available *via* “Array manager”.



“Manual peak picking”. Set the main marker at the position you want to label and click this button.

“Delete peak labels” button deletes all peak labels.

“Sticks” - change a representation of the 1D spectra in a cycle: “Line”/”Sticks”/”Line&dots”.

“Whitewash” - if ‘on’ then the area under 1D spectra will be ‘white washed’. Especially useful for pseudo-3D spectra, created by ‘Take->stack’ function (see panel ‘2D’ description).

“Sidebands picking.” Set the main marker on the top of the centreband. Optionally you can set the secondary marker so its horizontal bar specifies the threshold for the sideband intensities. Press the

button. You will be prompted to input the sideband label, spinning rate and suggested precision for the spinning rate. Sidebands labels will be set at the top of the discovered sidebands. The labels are editable via 'Select/Edit graphic objects'. If spinning rate error is too low, then some sidebands will marked not at their top point. Too high error causes the picker to pick up wrong peak.
In the same dialog you can specify whether the sidebands envelope should be created as a separated dataset. The envelope can be used, for example, in fitting procedure using pNMRsim.

"Normalise". Normalises the intensity of all opened datasets. Actually, it normalise all open dataset, regardless those are 1D or 2D.

S/N

Calculate signal-to-noise ratio. Works for 1D datasets only. Clicking the button calls a dialog where user can input a region on which noise calculation will be performed and a region where peak with maximum intensity will be searched for the signal-to-noise calculation. The default values, appearing in the dialog after pressing S/N button, correspond to the regions with 10% of the spectral width and contain the minimal noise level and maximum peak intensity respectively. These regions can be used if user wants to estimate the maximum value of the signal-to-noise ratio available on the spectrum.
The noise level is computed according to equations:

where

$$a = \sum_{i=-n}^{n} Re_i$$

$$b = \sum_{i=-n}^{n} Re_i^2$$

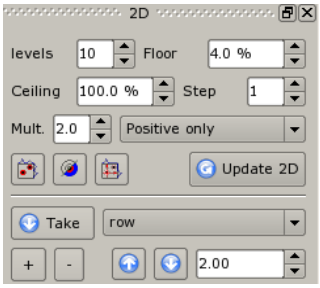
$$c = \sum_{i=1}^{n} i(Re_i - Re_{-i})$$

N- is the number (forced to be odd in the routine), i – index of the point which runs between –n for the leftmost to n to the rightmost point of the noise region. Re_i – the real part of the intensity of the point i.
Signal-to-noise ratio then estimated as:

Where R_{max}- is the maximum of the peak intensity, found in the specified peak region.

Panel "2D"

Aimed to perform different visual operation on 2D spectra.



“Levels”, “Floor”, “Positive/Negative”, “Ceiling”, “Multiplier”
Set the parameters for contours calculation. Press to “Update2D” to perform recalculation. ‘Step’ box allows to set a step used in contour and raster calculations. Bigger step – faster calculations but worse the overall quality. The value of ‘1’ provides the best quality.

“Contours on/off” switches between contour and raster modes. Calculating contours can take some time whereas it is normally faster in plot manipulations. The countours can be drawn with different colours be pressing 'C' button on the keyboard or by adding legend to the 2D plot (see 'Add legend' button description).

“Projections on/off” adds projections (sum or 'skylines' of all rows/colimns depending on settings in View->Options dialog) as sideviews of the plot. The intensity of the sideviews can be scaled by using mouse wheel or ‘PageUp’/’PageDown’ keys.

“Smart sideviews”. If this button is toggled than the sideviews appearance depends whether the markers are present or not. There are three possible options:

- If no markers are present then the usual projections (sums or skylines) are being displayed.
- If the only main marker is present then instead of the projections the sideviews display slices, taken from the position of the main marker.
- If both, the main and the secondary markers are present then the sideviews display partial projections, *i.e.* projections taken between markers positions. For example, horizontal partial projection will be taken for the part of the spectrum, located between red horizontal lines of the main and secondary markers.

A level at the left upper corner of the plot shows what exactly is displayed in this mode: projections, slices or partial projections.
It should be noted, that this mode can be slow for large 2D datasets.

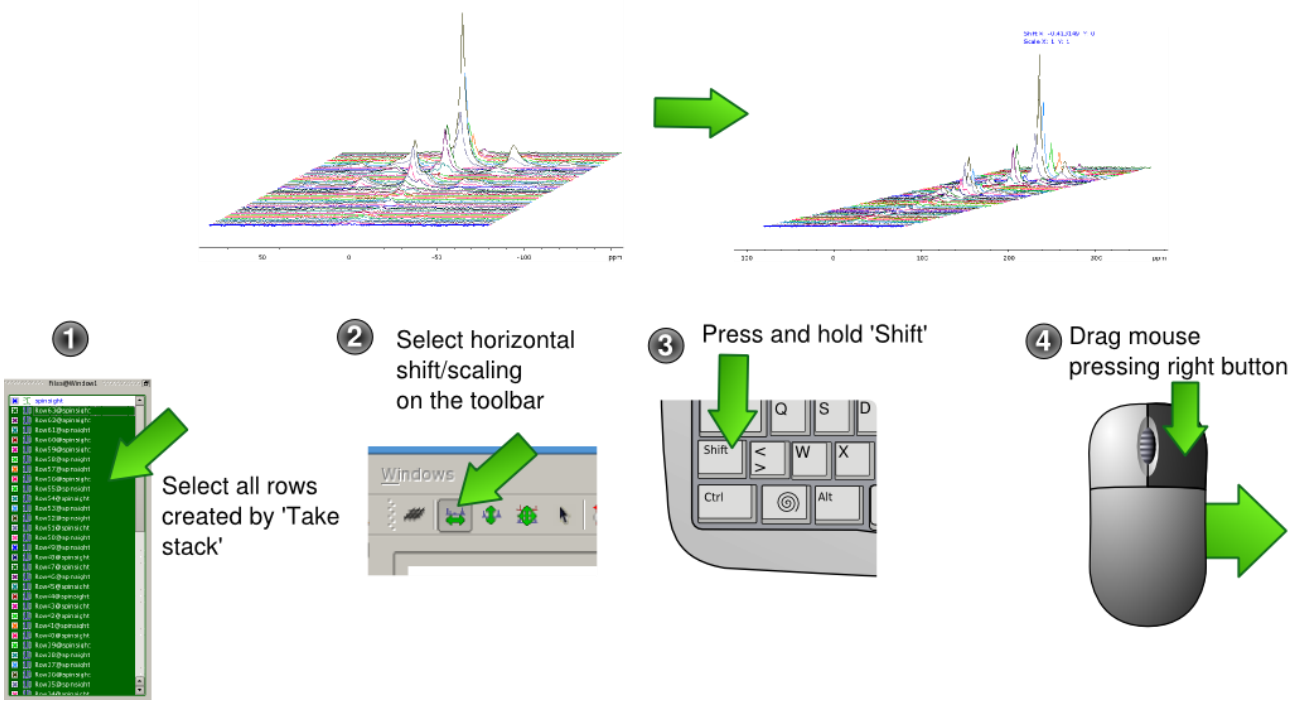
“Add/remove legend”. Adds or removes legend with the absolute values of the level. For the contour plots switches to the colour representation of the contours.
“Legend”: add remove legend or 2D raster/contour plots.

“Arrow up” and “arrow down”: GSIM uses the same absolute values of the levels for all displayed

2D spectra. As result, it is difficult to compare 2D spectra with very different intensities. These buttons allow doing multiplication/division of active spectrum intensities ("z-axis" intensities) by factor, displayed in the field next to the buttons (by default, 2.0).[3] If by the reason of different intensities you are not able to see second 2D spectrum, it is advisable first normalise the intensities (see "Peaks" panel) and then adjust them using these buttons.

"Take"
Extracting specified 1D spectrum from active 2D dataset. Extracted spectrum can contain:
-row/column – row or column, pointed by main marker (so it is required the main marker to be set).

- horizontal/vertical projection – actually it is sum of rows/columns
- horizontal/vertical (+/-) skyline – takes max/min points form rows/columns.
- horizontal train – convert 2D spectrum into single 1D where rows follow each other. It also creates a separating lines between 'former' rows. GSim also looks for data arrays, with the length equal to the number of rows in 2D dataset and puts labels with the values over the corresponding part of the 'train'. The separation lines and labels are graphical objects and can be edited/deleted in 'graphic' mode.
- stack – converts 2D spectra into a series of 1D spectra. Each 1D spectrum appears with a small horizontal and vertical shift, so it looks like 3D view. It is useful to use 'Whitewash' button ('1D' panel) to modify the appearance of the 'pseudo-3D' spectrum, created in this way. You also can change a 'viewing angle' (in reality, change synchronously the vertical and horizontal scaling) by selecting all relevant 'rows' in the file list, selecting either 'vertical' or 'horizontal shift/scaling' on the toolbar, and drag mouse with right mouse button pressed holding



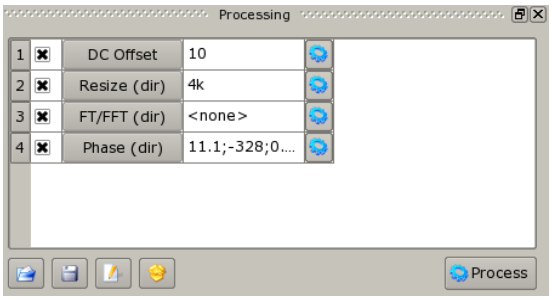
simultaneously the keyboards 'Shift' button pressed (see figure below).

"+" and "-" buttons. If you take row or column from 2D spectrum, this button allows browsing over rows or column.

[3]NB: it actually changes the spectral data. If you save spectrum after these operations the saved data will contains new scaled intensities.

How does this work: if row or column has taken, the new 1D dataset remember the index of its "parent" in file list. By pressing "+" or "-" buttons previous (active) 1D data is deleted and the new data with higher/lower row/column is creating. That means if active row is not the last in the file list, after incrementing/decrementing row/column number it will be moved to the end of the list and it colour will be changed – don't panic if it happens! It also means, if "parent" being moved in files list or deleted then program likely to crash!

Panel "Processing"



All processing functions are concentrated here. Each processing function is represented by the row in processing table. Each row contains 4 cells:
-Checkbox which specify would or not the function applied by clicking "Process" button.
-Function button with function name on it. Clicking this button the processing function would be applied immediately on the selected datasets using options specified in the next cell. Mark "(dir)" means the function is applied to direct spectral dimension, "(indir)" - indirect or vertical dimension of 2D spectrum.
-Option cell. This is a simple text cell. All required options are specified in the single line and separated by ";". For example, in case of the exponential linebroadening of 100 Hz the option string for "LB (dir)" function should be: "lorentz;100.0". If you are not sure what is the allowed range for options then use the "..." button from the next cell.

-Button "  " allows an interactive input of the processing options. Normally it is asking step by step about all needed parameters.

All functions, checked in their first column, could be applied in "single shot" by pressing "Process" button. Order is from top to bottom of the processing table. Already processed data can be reload from disk and reprocessed clicking "  " box. Simple editor for the processing list could be called pressing "  (edit)" button. The same editor can be used to place some of the processing buttons on the processing toolbar (see description of the menu command 'Process'). Current processing list can be stored on disk using "  (save)" and loaded later by "  (load)" button. GSim can try to extract some processing parameters from commercial data formats as Spinsight or Xwin-NMR. However, don't expect too much from this -function – spectrometer manufacturers are not going to describe in a clear way how they stores processing information and this function is based mainly on my personal "hacking". (If you know a bit more about this subject – let me know).
Generally processing list are independent of the datasets. This approach has several advantages, for example, you can easy apply the same processing on several dataset by one click. Drawback of this approach is that you have separated 'files' for processing and data respectively. However, you can 'attach' processing file to original dataset. To do this, select all datasets you want to be attached to the current processing in processing table, call save processing list dialog and check the box "Attach to selected dataset. When next time you will try to open one of these datasets, GSim will recognise

that the dataset has an attached processing and will prompt you to load, ignore or remove attached processing. For performance reasons Gsim stores no more then 100 'links between datasets and processing files. The size is keept in good fit by deleting the oldest links. Links are stored as a text information in file $HOME/.gsim/AttachedProclist.

Below is a short description of the individual processing functions.

“Sort rows”

Under some specific circumstances the rows of 2D experiments can be obtained in non-standard order. In order to process such a data one has to resort rows accordingly to some specified experimental vector (containing, say, randomised t_1 values for the evolution time). Pressing this button user will be asked which vector he wants to use for resorting and then actual resorting takes place.

“Rearrange”

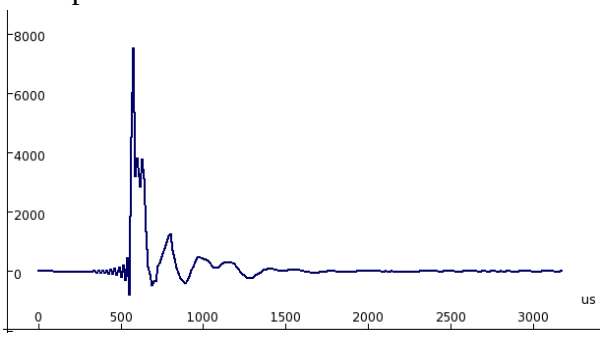
This function changes the number of points in rows and column. Only the number of points in direct dimension is needed. The number of points in indirect dimension is equal to number of total points divided by the number of points in direct dimension. The later value should be an integer otherwise the procedure produces an error message. It's useful, say, for some SPINSIGHT experiments recorded in STATES mode, where sin and cos FIDs are recorded into the same row.

“DC offset”

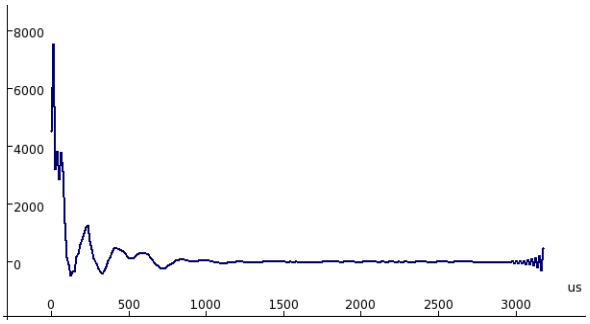
Performs DC offset correction. The option field specified how many points from the tail of FID will be used to determine total offset (10% of total length by default). This performs independently for real and imaginary data, as well as for each rows in 2D dataset.

“auto Bruker”

Bruker Avance data can be acquired in “digital” mode. The FID taken in this mode has a form near beginning. Simple Fourier transformation leads to the spectrum whit huge 1st order phase distortins. Cyclic shift of FID by certain number of points fixes this problem. The “auto Bruker” function tries to determine the number of points for shift using parameters in ‘acqus’ file and perform this shift. Note: after the shift linebroadening, DC offset correction or resizing can gives some of the artefact on the spectrum.



Before “auto Bruker”



After

“Resize (dir) or (indir)”

Performs resizing of the data in direct or indirect dimension. If new data size is larger than previous this function is doing zero filling. In opposite case data will be truncated. This is the normal way to do a zero-filling in GSIM.

“Shift (dir) and (indir)”

Perform shift of data in direct/indirect dimension by specified number of points. Shift is cyclic. That means, the points removed from one side appears at the other side. Direct shift could be useful, say, for processing of the full spin-echo FID. Indirect shift allows processing folded 2D spectra.

“Shearing”

performs left incrementing shearing of the data. Usually performs on data in the frequency domain. See “TShearing” for the shearing in time domain. First parameters specifies the number of points which would be removed in the first row; second parameter is an increment; third parameter is a step using in incrementing. Step equal to 1 means that incrementing will be applied to each row, step 2 means that incrementing affects odd rows only. For example if the parameters have values of 10, 20, 2 respectively and 2D dataset contains 10 rows, then 10, 10, 30, 30, 50, 50, 70, 70, 90, 90 points will be removed (or moved to the end, see further explanation) from the beginning of each row respectively. Forth parameter specifies whether the ending points of each row in the sheared spectrum filled by zeros (‘cut’ option) or by cut part (‘shift’ option).

"t1-dep PHC"

Performs phase correction of 2D dataset with correction coefficients that varies linearly from row to row. The actual applied phase is calculated as

$$phase = (PHC0 + PHC0_{inc} \cdot i) + (PHC1 + PHC1_{inc} \cdot i)\, pivot - \frac{point}{N}$$

where PHC0, PHC1- values of the zero and first-order phase correction coefficients as applied for the first row, $PHC0_{inc}$, $PHC1_{inc}$ - increments for PHC0 and PHC1; *i*-is a row counter. Row counter starts from zero. Special parameter ‘step’ controls when this counter has to be changed. If ‘step’=1 then *i* is incremented for each row. If ‘step=2’ then *i* is incremented only for even rows, *etc*. Normally the ‘step’ should follow the applied increment method for t_1-dimension in 2D NMR. For STATES and STATES-TPPI step should be equal to 2 and for TPPI it should be equal to 1.

‘Point’ is an index of point in the row. N- the total number of points in each row. ‘Pivot’ is a pivot point (in fraction) for the first-order phase correction. For example, ½ shows that the first order phase correction is applied relatively to the centre of the spectrum. Interactive inputs asks parameters PHCO, $PHC0_{inc}$, PHC1, $PHC1_{inc}$, step and pivot respectively.

Note that the t1-dep PHC do not take into account the type of 2D data and therefore do not produce correct shearing results, e.g. for the 2D datasets acquired using STATES method. Use “TShearing” instead.

"TShearing"

Performs shearing in time domain. This is the most common shearing operation for experiments like MQMAS, etc. At the moment only one type of shearing is supported – shearing for MQMAS spectra. For MQMAS experiments the shearing operation should be placed between Fourier transformations in direct and indirect dimensions. The option string requires two parameters: a “shearing factor” and 2D spectrum type. Shearing factor is measured in degrees and used to calculate the phase of each point in the spectrum using the equation:

$$phase=daslp*((i-np/2.0)/np)*j/ni$$

where *i* runs over all points within the row and *j* runs over time increments in indirect dimension.

Parameters *np* and *ni* store the number of points in direct and indirect dimension.

2D type should specify the type of 2D dataset. At the moment only STATES is supported for this parameter.

To simplify the calculation of the shearing factor the interactive input provides the fields where user can set the nuclei spin, coherence order and 2D type. For example, for 3Q-MAS experiment for ^{27}Al the user should input "5/2" as the spin number of ^{27}Al and "3" as the observed coherence order. The value of the computed shearing factor is presented in the second dialog window and can be corrected by the user if needed.

Note that the shearing operation is very similar to t1-dep PHC described above but the t1-dep PHC do not take into account the type of 2D data and therefore do not produce correct shearing results, e.g. for the 2D datasets acquired using STATES method.

"LB (dir/indir)"

performs apodisation in direct or indirect dimension. Lorentzian ('lorentz'), gaussian ('gauss'), sinebell ('sinebell') and squared sinebell ('sq_sinebell') functions are available for simple apodisation. The only parameter is a linebroadening in Hz or angle for sinebell-type apodisation (0 – for pure sine and 90 – for pure cosine). For full-echo measurements functions 'shift_gauss' and 'shift_lorentz' can be used. In this case the third parameter in the option list should specify a position of the (echo) maximum on the FID, measured in microseconds. Bruker data with digital filtering should be processed in a special way (see 'autoBruker' function). In order to get a right apodisation for such Bruker FIDs one can use functions 'bruker_lorentz' and 'bruker_gauss'. Intrinsically they are identical to 'shift_gauss' and 'shift_lorentz' but the required shift is determined automatically in a similar way to 'autoBruker'.

"t1-dep GLB"

Performs t_1-dependant Gaussian line broadening. The area of applicability for this function is similar to that of "t1-dep PHC" (see above). The only one of possible parameters is varying – the centre of the gaussian apodisation function whereas linebroadening coefficient is constant. In a particular case of MQ-MAS the centre of the gaussian function should follow the echo maximum. See specialised literature how to calculate it. It is likely that in further releases of GSim more automatic functions will be implemented in addition to "t1-dep GLB" and "t1-dep PHC".

The interactive input for this function requires: line broadeng in Hz, initial value of the centre in microseconds, increment for the centre position in microseconds and a step which controls when incrementing will be applied (see discussion about 'step' parameter for "t1-dep PHC").

"FT(direct)"

doing Fourier transformation in the direct dimension if data is FID. If data is "SPECTRUM' than invert FT is applying. For curious persons: first point multiplication by 0.5 and exchanging of two halves of spectrum is already included in this function. If data size is not equal to 2^n than data will be resized to the closest value of 2^n.

"Phase (dir/indir)"

doing a spectrum phasing. Parameters are: zero-order phase and first order phase. The pivot point of the first order phase is internally fixed at the spectrum centre. However, it still can be set at the main marker position in the interactive mode (see below).
It is more convenient to do phasing using interactive input button. Pressing this button the phasing

dialog appears. Simply move sliders to get a desired phase. For the first order phase correction put the main marker at the reference point. In order to apply phase correction press corresponding button in the table with function name on it ("Phase (dir/indir)"). Of course, you can use an entry in the "Process" main menu as well.
For 2D dataset you will be prompted to take slices (rows for direct phase corrections and columns for indirect one). Put the main marker at the position of the slice we want to use and press 'Take' button. After interactive input all slices will be automatically deleted. User can also do phasing on several selected datasets simultaneously.

"auto Phase (dir)"

Automatic phase correction. The dataset should be a 1D spectrum or 2D dataset where FT has been applied in direct dimension but not applied in indirect dimension. In the later case the phase will be determined from the first row and applied for all rows in the spectrum. Such behaviour is designed to be useful for arrayed experiments (like relaxation measurements) or in the intemediate step of 2D spectrum processing. Can be a bit slow.

"auto BL correction"

performs a baseline correction. Polynomial, Fourier and Bernstein corrections are supported ('poly', 'fourier' or 'bern' should be the first option in the option list). The second option is an order of the series. You can either use automatic peak determination (experimental code) or specify the regions that should be excluded from the correction (peaks regions) manually. Each regions should be specified in the option list by a pair of numbers (in current units) "from point" - "to point". The exact order (from-left-to-right *vs*. from-right-to-left) is not important. Interactive input allows to do this interactively by setting by markers around the desired region. In order to use automatic mode the user should provide 'auto' (or '/auto') as the third option.

"Magnitude"

calculates the magnitude of the spectrum.

"FT (indirect)"

performs Fourier transformation in the indirect dimension. Five modes are accessible: complex, STATES, TPPI, STATES-TPPI and echo-antiecho. If STATES, TPPI, STATES-TPPI or echo-antiecho modes are chosen the data size in the indirect dimension after FT is a half of the initial data size. Hypercomplex data is not stored, so, please, perform phasing in direct dimension before pressing this button. Invert FT potentially works (but untested) for complex transformation, but cannot reverse the FID after STATES or TPPI.
For curious persons: if STATES is chosen the new dataset is created with half-size of indirect dimension. Real part of each points of the new dataset is equal to real part of odd points in old one. Imaginary part of the new dataset is equal to real part of odd rows in old data. Then complex FT is calling.
If TPPI mode is chosen then imaginary part is filled by zeros and complex FT is calling. This mimics a real FT transformation. Then the half of rows throws away.
In STATES-TPPI mode first the same processing applied as in STATES mode. Then sign is alterated for each new even complex datapoint.
If "echo-antiecho" is chosen then the procedure taken from http://rmn.iqfr.csic.es/guide/nmr/smanual/8_4_proc2d.html#Echo-Antiecho is applied:

Each two consecutive FIDs are replaced by

re0=-im1-im0
im0=re1+re0
re1=re1-re0
im1=im1-im0

processing then continues as in the STATES case.

“addNoise”

adds random noise to data. Option parameters specifies the level of noise (in%) to the maximal peak intensity.

“Smoothing”

Smoothing spectrum using given number of “+/-” points *n*. Each points *i* in the spectrum becomes equal to an average value taken between *i-n* and *i+n* points.

“Transpose”

Swap rows and columns in 2D dataset.

“Crop”

Crop data according to selected boundaries: 2 parameters (left/right) for 1D spectra and 2/4 parameters for 2D. In the case of 2D spectrum 4 parameters corresponds to left, top, right and bottom position. In case of 2 parameters no croping in indirect dimension has applied and points corresponds to left and bottom position as in 1D case. The parameters can have a units suffix. For example, parameter “100pts” means that the truncation will be done at 100 points. The allowed suffixes are (case-sensitive) “pts”, “us”, “ms”, “s”, “Hz”, “kHz”, “ppm”. If parameter doesn't have the units suffix it's assumed that parameters have the same units as the current units of the plot. If parameter has units, incompatible with the data, for example, “1.0ms” applied on the spectrum with the current “ppm” scale, the units suffix will be ignored and the parameter will be assumed to have the same units as the data (“ppm” in our example).
Interactive input simply takes the positions of the markers.

“Append data”

Allows to combine several data sets. All involved dataset names are taken from the option list that can be set interactively. Result stores in the dataset “combination” with parameters (sw, etc.) taken from the first dataset listed in option string. In order to have a right result all datasets in files list should have different names. If two or more datasets have identical names then option parser will pick up the first one.
If specify the single option “/all” or set interactively “100” for the datasets number then all opened datasets will be appended in order of their appearance in the ‘Files’ panel.

“Add spectra”

Performs linear combination of several dataset. The option list should contain names of the datasets which should be added/subtracted and the weighting coefficients for the datasets. For example, coefficient “1.0” causes a simple addition, and “-1.0” - subtraction. The result is stored as a “sum” dataset that inherits most of the options from the active dataset.
Example: if the option string for “Add spectra” contains:

first;1.0;second; 1.0; third; -1.0; forth; 0.5
(of course datasets"first", "second", "third", "forth" should be open).

then for each complex point of the corresponding dataset:

sum=1*first+1*second-1*third+0.5*forth

If combined spectra have different spectral ranges, the function tries to be "cleaver" and make a sum over the overlapping region only. The function is also able to sum datasets with different numerical resolutions by using cubic interpolation. The resolution of the resulting sum will be equal to the highest numerical resolution of the involved datasets.
If specify the single option "/all" or interactively set "100" for the datasets number then all opened datasets will be added with coefficients 1.0 in order of their appearance in the 'Files' panel.

"Load file"

Load file from the disk. Note a significant difference between this function and the functionality of 'File->Open' action. If 'Load File' processing function is applied, the loaded file substituted the file which is under processing. However, file attributes, stored within GSim (as filename, for instance) are preserved from initial file. Thus, if you, say, press 'Reload' button then the initial file will be loaded, not the file, specified as an option of the 'Load File' processing function. This sort of behaviour is chosen in order to fit 'External Command' functionality (see below).

The option string should contains a file path. File format is determined automatically as for "Any known file format" in File->Open. White spaces should be avoided in the file path.

The 'External Command' section describes how to extend GSim processing functionality via external tools.

"Save file"

Saves processing file on disk. Similar to 'Load file' (see above), the file attributes are preserved from the initial file, not from the saved one. Option string should contains two parameters: filename and format which is one of the following: simplot, matlab or spinsight. If format is ommited, the 'matlab' format is assumed.

The 'External Command' section describes how to extend GSim processing functionality via external tools.

"External command"

This processing function simply executes external (operation system) command written as an option string. Because GSim doesn't like white spaces in the option string, replace them by symbol ';'. It will be done automatically if you are using the interactive button.

This function, in combination with 'Load file' and 'Save file' functions (see above), can be used for extending the functionality of GSim.

Earlier or later one can find that built-in functionality is not enough for his/her specific dataset. The sophisticated NMR processing sofware should provide a way to extend it using custom functions. In most cases it is achievable via macroses, automative functions, user scripts, etc.

The functionalities of GSim can be extended by several ways. **First of all**, GSim is an open-source project, so everyone can write an additional processing function in C++ and recompile GSim. In this case the new function will absolutely 'native' for GSim, however, a substantial programming

knowledge is necessary.

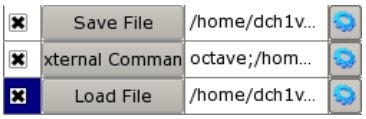
Other way would be to extend the functionality via scripting. This way is not realised at the moment but probably will be realised in the future via QScript language, provided by Qt library. This option is fast and don't require full recompilation. However, you still have to learn a specific 'language' in order to write a script.

"External command" provides **a third way**, in which you can use your favourite software (or programming language) for creating an extension.

Below is an example of the function which is written using free 'Octave' program (www.octave.org). The 'Octave' provides a very similar environment to that of the commercial 'MatLAB' program.

The example, presented below, simply inverts the intensities of the spectra.

First of all, you should create a block of three functions in the processing table:



First function, 'Save file' saves a temporary file (with the name '/home/dch1vz/temp.mat' in this example) in MATLAB format.

Second function performs actual intensity inversion. For this example its option string is 'octave;/home/dch1vz/inv.m' which do the same thing as a command

```
octave /home/dch1vz/inv.m
```

typed in the command line of your OS.

Last function, 'Load file' loads the transformed file back in GSim. Option string is identical to that in 'Save file' function: '/home/dch1vz/temp.mat'.

Before starting actual processing we have to create our "script" and save it in the file '/home/dch1vz/inv.m'.

In this example the script contains only three strings:

```
load /home/dch1vz/temp.mat;
spectrum1.data=(-1)*spectrum1.data;
save -mat-binary /home/dch1vz/temp.mat spectrum1;
```

First line loaded the file, saved previously in GSim. Second line inverts the intensities and, finally, third line saves the transformed dataset (which is stored in variable 'spectrum1') into the same file on the disk in binary MATLAB format.

MATLAB format can be easily read by several mathematical software packages. See description of the GSim datastructure in section where MATLAB format is introduced.

Of course, other formats can be used. For example, SIMPLOT format stores data as a ASCII text. It should be quite simple to write I/O procedures for this format in languages as Python, Perl or Tcl.

This approach is the slowest one (it requires several disk operation and loading of an external, sometimes bulky tools as Octave or MatLAB). However, it is the most flexible one. A possibility to

include ‘External command’ in processing table in GSim opens a way to do processing anywhere and in any way you like and use GSim to visualise the results just by clicking one ‘Process’ button. It worth mentioning that the processing table can contain a series of ‘External command’ strings between ‘Save file’ and ‘Load file’ functions.

This approach also have some ‘legal’ implication. Writing C++ procedure for GSim or even script for QScript you are actually using Qt library distributed by the company called ‘Trolltech’. GSim uses free version of Qt which is published under ‘general public license’ (GPL). According to GPL, any code which is uses GPL code should be published under GPL license too. That means you have to make your script/procedure available for free download and publish its sources. In contrast, using ‘external command’ you are writing an independent ‘program’ which is not linked with GSim or Qt whatsoever. So, you can choose whether to make the latter code available for everyone or not (of course, if the license of your programming environment allows this).

“XY balance (indir)”

Experimental processing function designed for specific needs of static 2H EXSY experiment, using original phase cycle from C.Shmidt, B.Blümich and H.W.Spiess, *JMR*, **79** (1988) 269. The experiment creates somewhat unusual 2D data which needs extra processing step. The described function applies 90° phase shift in indirect dimension and multiplies each imaginary point by the coefficient, specified in the option string. For the mentioned experiment the function should be applied after FT in direct dimension. After that, an FT in indirect dimension in States mode should produce 2D spectra without artefacts. This function is likely to be removed from future Gsim releases.

“redor”

Creates REDOR curve from the arrayed data set. It is assumed that the dataset has 'reference' Hahn-echo experiments in even rows and REDOR experiments in odd. The columns from the resulting 'redor' dataset can be converted to the arrays and fitted in Array Manager. Array Manager has a dedicated 'redor' function for the fitting of the REDOR curves in 'spin pair' approximation.

“Sum Sidebands (dir)”

Performs sidebands summation. Can be useful to improve signal-to-noise ratio for solid-state NMR spectra with sidebands manifold.

This processing function should be applied on the time-domain FID data, before Fourier transformation (FT). After FT on the resulting FID, the final spectrum can be expressed as:

$$\sum_{i=-N}^{N} S(f + i \cdot srate)$$

Where *S(f)* - initial spectrum as a function of the frequency *f*, *srate* – repetition rate provided as the first option for the ‘Sum Sidebands’ function, N – the highest sideband order which should be provided as the second option. For example, *S(f+1*srate)* will corresponds to the initial spectrum shifted by the *srate* value (expressed in Hz) in the cyclic way.

The final sidebands sum can be obtained by cropping the spectrum around the centreband region.

6. Operation system command line options

GSim can perform some operations from the command line interface, provided by the operation system installed on your computer:

Open files: just include all filenames after 'gsim' command. Example:

```
gsim *.fid
```

will open all file with type 'fid' (Simplot format). Automatic format determination is used.

Process according to the given processing list: using the option '-p [processing list filename]' the all loaded file will be processed according to the given processing list. You can also use the last processing list used in GSim under graphical mode. GSim saves it under <you home directory>/.gsim/proclist.prc.

Save file: there are three options to save the active dataset: '-smatlab', '-ssimplot' and '-sspinsight' which save in MATLAB, Simplot and Spinsight formats respectively. The filename should follow the '-s...' option. If one of the save options is provided, GSim is not launching in the graphic mode. However, X11-server should run on Unix system (at least for possible error reports which appears as graphical windows). There are no problems on Windows and MacOS as far as graphics is always on for these systems.

Only single file can be saved. If you would like to process several datasets you can do it via command shell scripting (say, batch or c-shell).

Examples of command line operations:

```
gsim test1.fid test2.fid
```

open GSim in graphic mode with data test1.fid and test2.fid.

```
gsim *.fid -p /home/myacc/.gsim/proclist.prc
```

Open all files in the current directory with extension '.fid', process them according to the previously used processing list and open GSim in graphical mode with the processed data.

```
gsim varian/fid -ssimplot test.fid
```

Converts Varian file in Simplot test.fid in command line mode.

```
gsim test.fid -p myproc.prc -ssimplot test.spe
```

Command-line processing: open file test.fid, process it according to the myproc.prc processing list and save in test.spe.

If you have found these programs useful, please, make a references to this web-page in your publications:
Vadim E. Zorin, "*GSim - visualisation and processing tool for NMR experiments and simulations*", URL: http://gsim.sourceforge.net