

SasView Tutorials

Simultaneous 1D Data Fitting in SasView Version 5

www.sasview.org

Preamble

SasView was originally developed by the University of Tennessee as part of the Distributed Data Analysis of Neutron Scattering Experiments (DANSE) project funded by the US National Science Foundation (NSF), but is currently being developed as an Open Source project hosted on GitHub and managed by a consortium of scattering facilities. Participating facilities include (in alphabetical order): the Australian National Science & Technology Centre for Neutron Scattering, the Diamond Light Source, the European Spallation Source, the Federal Institute for Materials Research and Testing, the Institut Laue Langevin, the ISIS Pulsed Neutron & Muon Source, the National Institute of Standards & Technology Center for Neutron Research, the Oak Ridge National Laboratory Neutron Sciences Directorate, and the Technical University Delft Reactor Institute.

SasView is distributed under a 'Three-clause' BSD licence which you may read here: <https://github.com/SasView/sasview/blob/master/LICENSE.TXT>

SasView is free to download and use, including for commercial purposes.

© 2009-2019 UMD, UTK, NIST, ORNL, ISIS, ESS, ANSTO, ILL, TUD, DLS, BAM

If you make use of SasView

If you use SasView to do productive scientific research that leads to a publication, we ask that you acknowledge use of the program with the following text:

This work benefited from the use of the SasView application, originally developed under NSF Award DMR-0520547. SasView also contains code developed with funding from the EU Horizon 2020 programme under the SINE2020 project Grant No 654000.

Contributors to this Tutorial

Steve King (stephen.king@stfc.ac.uk)

Piotr Rozyczko (piotr.rozyczko@esss.se)

Wojciech Potrzebowski (wojciech.potrzebowski@esss.se)

Last revised: 19 May 2019

Learning Objective

This tutorial will demonstrate how to simultaneously fit 1D ('intensity' versus Q) datasets in SasView. Batch fitting of multiple 1D datasets, and the fitting of 2D datasets, are considered in separate tutorials.

It is assumed that the reader has some familiarity with the purpose and principles of data fitting. If not, these Wikipedia articles provide an overview:

- https://en.wikipedia.org/wiki/Curve_fitting
- https://en.wikipedia.org/wiki/Mathematical_optimization

It is also assumed that the reader is familiar with using SasView to fit individual datasets (Single Fitting Mode) as covered in the tutorial ***basic_1d_fitting_in_sasview***.

The program interface shown in this tutorial is SasView Version 5.0 running on a Windows platform but, apart from a few small differences in look and functionality, this tutorial is generally applicable to any earlier version of SasView on any platform.

Glossary

<i>a priori</i> information	Known facts about the system whose datasets are being modelled that can guide the selection of model or model parameters.
Chi-square (X^2 , 'Chi2')	<p>A statistical test of how well a chosen model fits the data with a given set of model parameters. In SasView this means</p> $X^2 = \sum ((I(Q)_{meas} - I(Q)_{calc})^2 / E(Q)^2)$ <p>where I is the scattering intensity and E is the error on the intensity value. Clearly, as $X^2 \rightarrow 0$, the better the model fit is.</p> <p>NB: The SasView interface actually reports a variation of chi-square called the <u>reduced chi-square</u>, sometimes referred to as the '<u>goodness-of-fit</u>'</p> $X^2_{Reduced} = \sum ((I(Q)_{meas} - I(Q)_{calc})^2 / E(Q)^2) / (N_{pts} - N_{params})$ <p>where N_{pts} is the number of data points in the dataset and N_{params} is the number of model parameters being optimised (which may be less than the total number of parameters in the model!). As $X_R^2 \rightarrow 1$, the better the model fit is.</p>
Compute [button]	Perform a direct calculation of the model with the current parameters but <u>without any optimisation</u>
Correlated parameters	To obtain the best solutions the optimiser needs to explore the

	<p>widest possible parameter space. That is best achieved if every parameter can be considered independent of every other parameter. However, there will be instances where this is not the case and one or more parameters may be correlated. In such cases the best approach is to fix the values of some of correlated parameters using <i>a priori</i> information and then optimise the remaining values.</p> <p>A particularly common instance of correlated parameters encountered when model-fitting SAS data is when the components of the forward scattering intensity</p> $I(0) = \phi V (\Delta\rho)^2$ <p>where ϕ is the volume fraction of scatterers, V is the volume of one scatterer (and so dependent on size parameters), and $\Delta\rho$ is the contrast (difference in SLDs), are separated out as individual parameters in a model.</p> <p>Correlated parameters often manifest themselves in the SasView interface as having very large uncertainties.</p>
Fit [button]	Perform a calculation of the model <u>optimising the selected parameters</u>
Model-fitting	The process of finding a good yet <u>physically-realistic</u> mathematical description ('solution') for a dataset or collection of datasets. The procedure employed to achieve this is called optimisation.
OpenCL	A low-level software framework that allows calculations to be distributed between any compatible processors (eg, GPUs as well as CPUs) on the host computer. OpenCL can speed up demanding model-fits if suitable hardware is available.
OpenMP	A software framework that permits shared-memory multi-processing (ie, parallelisation) of calculations. OpenMP can speed up demanding model-fits if suitable hardware is available.
Optimiser	The mathematical algorithm used to perform the model-fitting.
Polydispersity	<p>Where one or more model parameters have a distribution of values.</p> <p>SasView allows for 2 types of polydispersity:</p> <p><i>Size</i> where, for example, the radii and/or lengths of the scatterers have a distribution of values</p> <p style="text-align: center;"><i>This will apply in most instances</i></p> <p><i>Angular</i> where the scatterers are anisometric in</p>

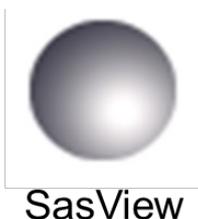
	<p>shape (eg, cylinders) and exhibit preferred orientations with respect to Q, for example, in under shear or in a magnetic field</p>
Reduced Chi2	See Chi-square
Residual	<p>The difference between the measured and calculated values of a function at a given point. In SasView this means:</p> $R(Q) = I(Q)_{meas} - I(Q)_{calc}$ <p>where I is the scattering intensity.</p> <p>NB: The SasView interface actually reports a variation of the residual called the <u>normalised residual</u></p> $R(Q)_{Normalised} = (I(Q)_{meas} - I(Q)_{calc}) / E(Q)$ <p>where E is the error on the intensity value.</p> <p>A normalised residual can be thought of as the number of standard deviations between the measured value and the calculated value. Thus, for a good fit, 68% of the values will lie within $-1 < R(Q)_{Normalised} < +1$, and 95% within ± 2.</p> <p>Residuals larger than ± 3 indicate significant problems with either the input data or the choice of model or model parameters.</p>
SLD	<p>Abbreviation for Scattering Length Density, a measure of the ability of a molecule to scatter. Strictly speaking, SLD is a SANS quantity, so if fitting SAXS data use electron density values in their place.</p> <p>SLD values (neutron and X-ray) can be calculated with the SLD Calculator Tool in SasView.</p>
Smearing	<p>Sometimes the instrumental geometry used to acquire the experimental data has an impact on the clarity of features in the scattering curve. For example, peaks or fringes might be slightly broadened. This is known as Q-resolution smearing.</p> <p>To compensate for this effect SasView can add a resolution contribution into a model calculation (which by definition will be exact) to make it more representative of what has been measured experimentally.</p> <p>SasView provides 4 smearing options:</p> <p><i>None</i> no smearing correction is applied</p>

	<p><i>Use dQ Data</i> the measured Q-resolution for each data point is used to apply a smearing correction <i>This is the default if dQ data is present</i></p> <p><i>Custom Pinhole</i> SasView will apply a smearing correction suitable for data measured on 'pinhole geometry' instruments (eg, most SAXS/SANS instruments)</p> <p><i>Custom Slit</i> SasView will apply a smearing correction suitable for data measured on 'slit geometry' instruments (eg, most USAXS/USANS instruments)</p>
Theory	The name SasView gives to a model calculation.
Uncertainties	<p>Every experimental measurement, including the measurement of $I(Q)$, is subject to some degree of error (which will, ideally, be included in the dataset). Similarly, the parameters returned by optimisation will have some associated range of uncertainty.</p> <p>Parameters with uncertainties that are more than 95% of the parameter value should be viewed with deep suspicion.</p> <p>NB: Unhelpfully, but for reasons of space, the SasView interface actually labels parameter uncertainties as errors.</p>
Weighting	<p>An optimiser can be instructed to pay less or more attention to data points in a dataset by changing the weighting of those data points.</p> <p>SasView provides 4 weighting options:</p> <p><i>No Weighting</i> all data points will be weighted equally</p> <p><i>Use dI Data</i> the data points will be inversely weighted according to their measured intensity errors (ie, less prominence will be given to data points with large errors) <i>This is the default if dI data is present</i></p> <p><i>Use sqrt (I Data) </i> the data points will be inversely weighted according to the square root of their intensity values</p> <p><i>Use (I Data) </i> the data points will be inversely weighted according to their intensity values</p>

Running SasView

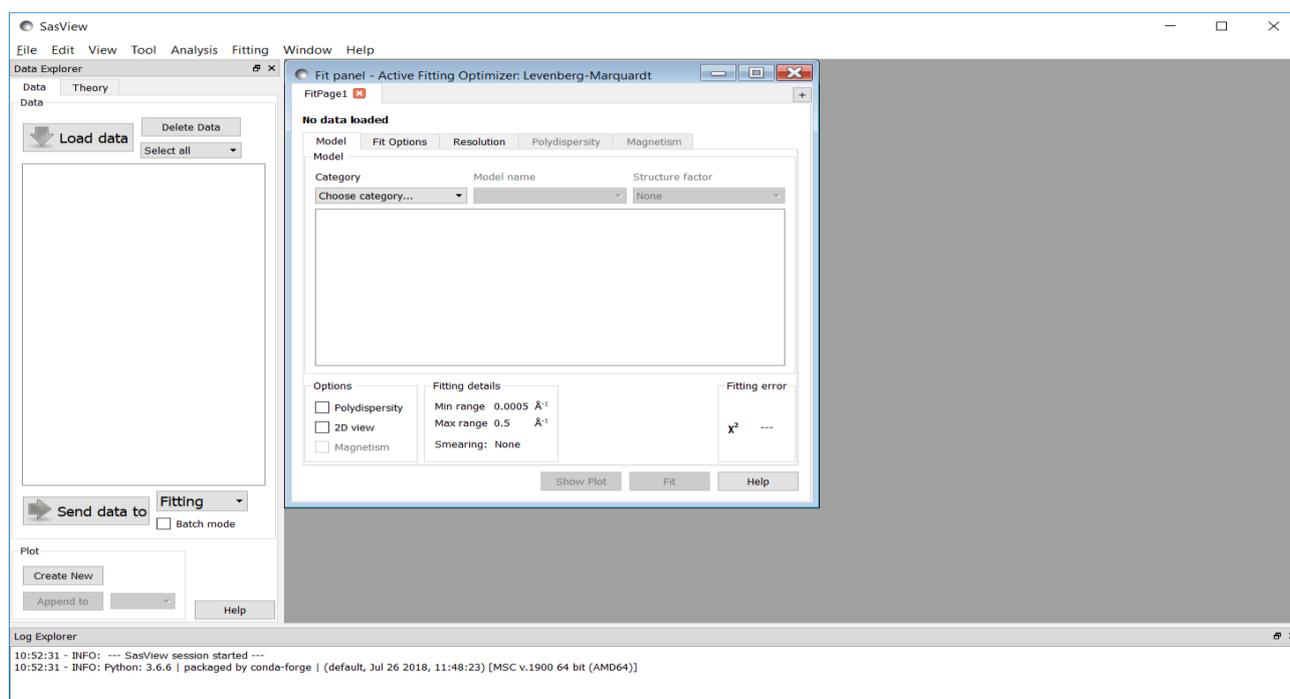
Windows

Either select SasView from 'Start' > 'All Programs' or, if you asked the installer to create one, double-click on the SasView desktop icon.



Mac OS

Go in to your 'Applications' folder and select SasView.



Example

This demonstrates a simultaneous model fit to three datasets with constraints and the use of the DREAM optimiser.

Simultaneous Fitting is an extension of the Single Fitting that allows for some relatively extensive constraints between several FitPage's (eg, to constrain all parameters in a series of datasets at different contrasts to be the same except for the solvent SLD parameter, or to fix the radius of the sphere in one FitPage to be the same as the radius of

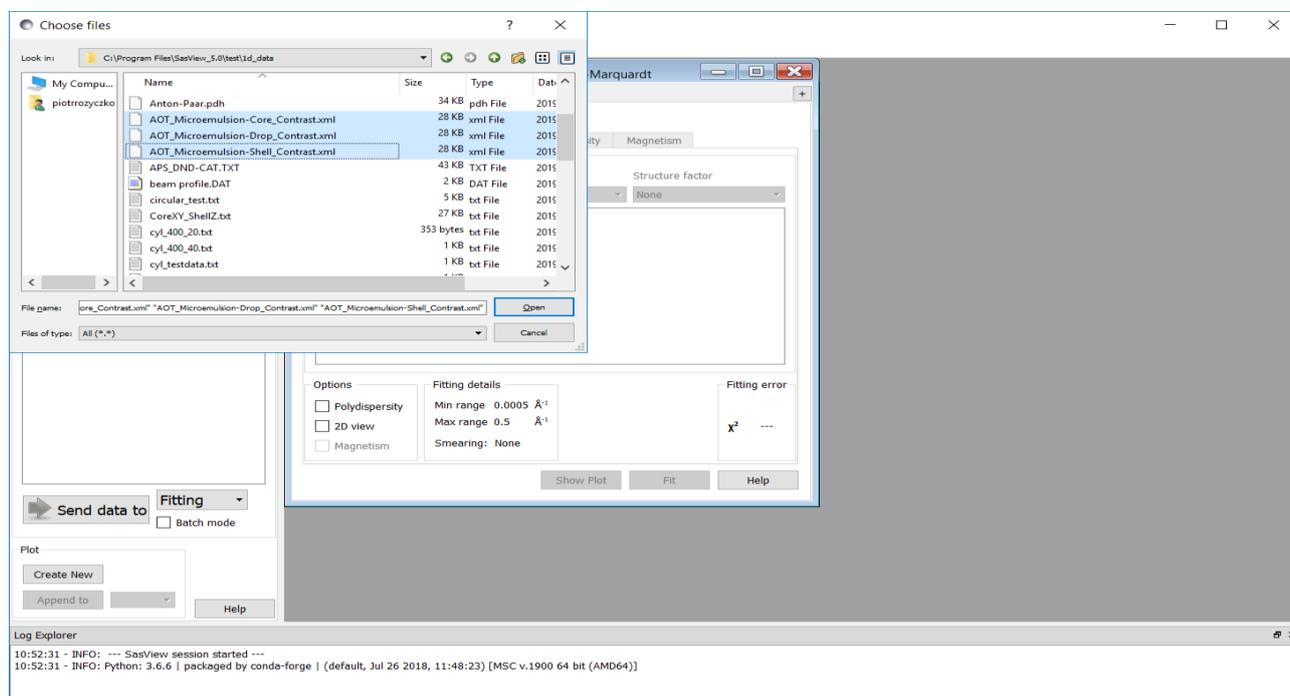
the cylinder in a second FitPage, etc), or between parameters in a single FitPage (eg, to constrain the length to be twice that of the radius, etc).

Tip: If you need to use a custom Plugin Model, you must ensure that model is available to SasView before proceeding.

In the Data Explorer panel, click the **Load Data** button, and navigate to the `\test1d_data` folder in the SasView installation directory.

Select the datasets **AOT_Microemulsion-Core_Contrast.xml**, **AOT_Microemulsion-Shell_Contrast.xml**, and **AOT_Microemulsion-Drop_Contrast.xml** and click the **Open** button.

These data are from a water-in-cyclohexane microemulsion system stabilised by a layer of the surfactant Aerosol OT® (AOT) located at the water/oil interface. The three datasets correspond to three different contrast conditions highlighting three different structural components: the 'core' (just the water droplets), the 'shell' (just the surfactant layer), and 'drop' (the water droplets plus the surfactant layer). The different contrasts (differences in SLD values) are achieved by changing which components are deuterated.

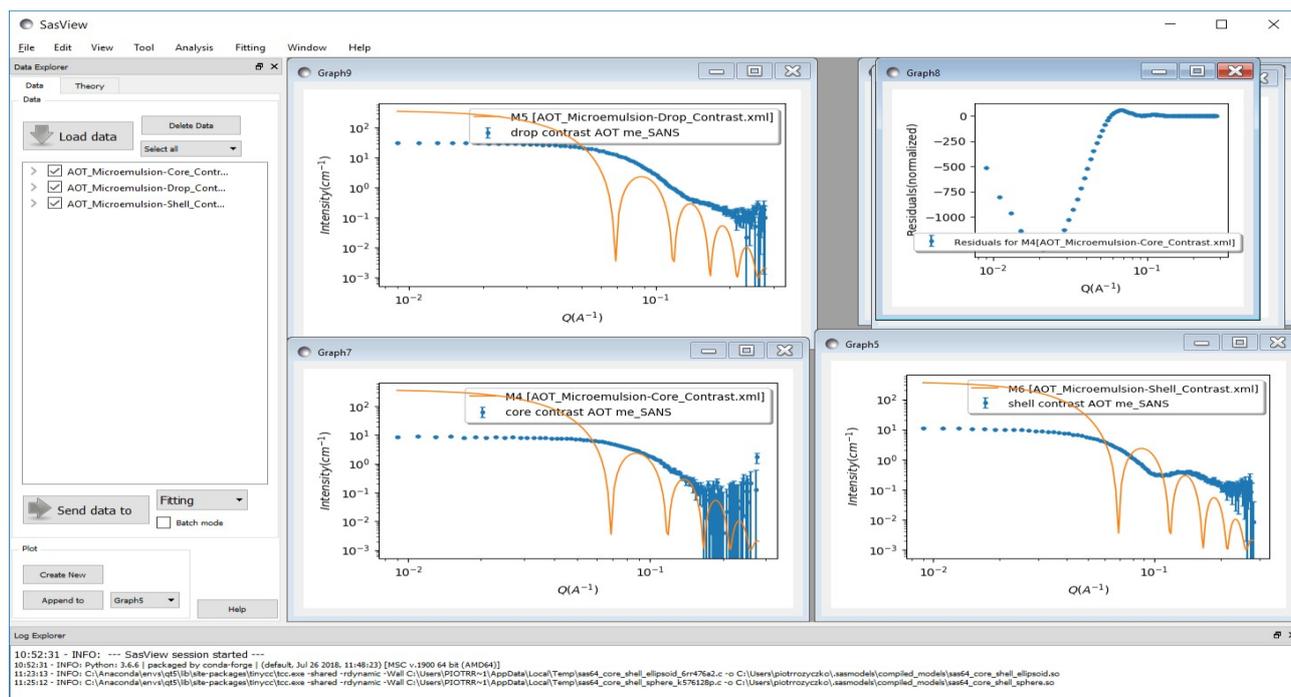


At the bottom of the Data Explorer panel, click the **Send To** button. As the load will have left all three datasets checked in the Data Explorer panel, this should result in the creation of three FitPage's.

Tip: If the multiple datasets are in one file, load that file, Unselect all Data (at the top of the Data Explorer panel), reselect just those datasets to be fitted, and then **Send To Fitting**.

Go to each FitPage in turn and select the **core_shell_sphere** model.

Then click on **Show Plot** on each page to see the graphs of the three theory curves added to dataset points.



Enter the following SLD values in the appropriate FitPage's:

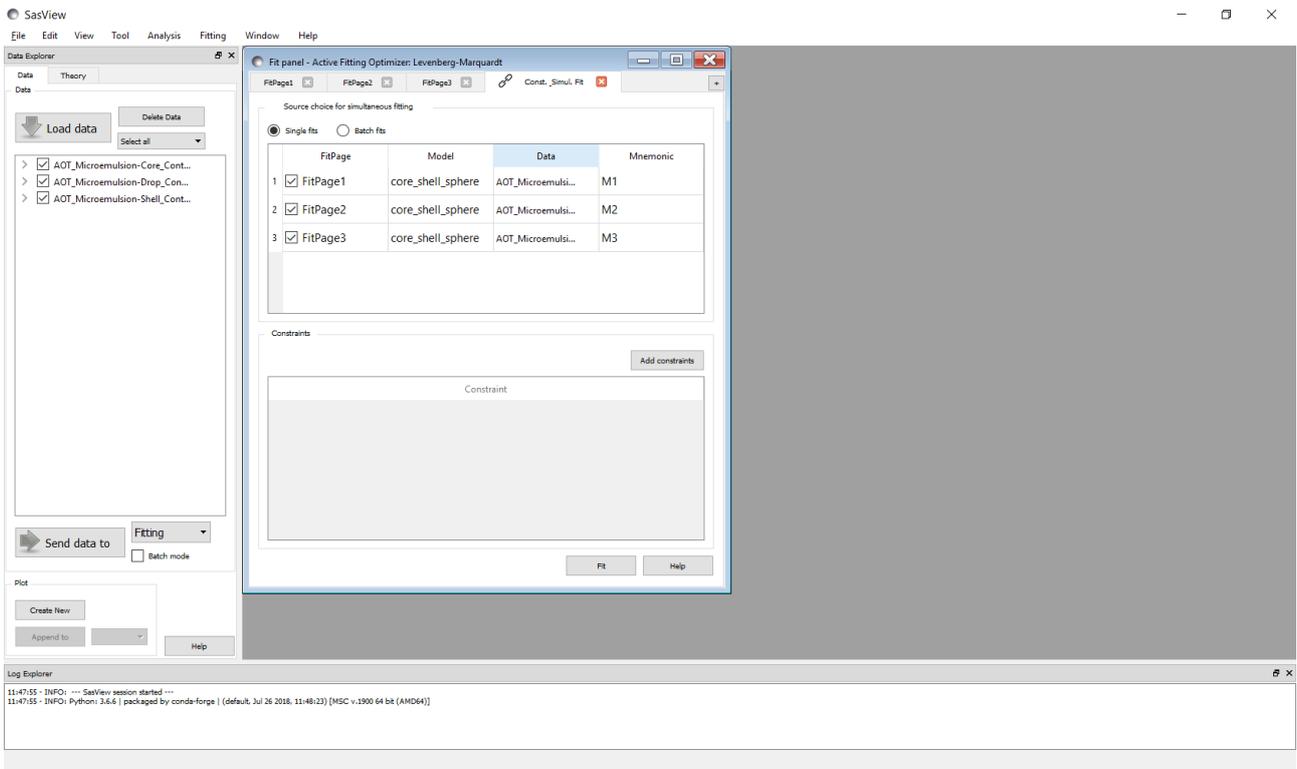
Contrast / SLD	<i>sld_core</i>	<i>sld_shell</i>	<i>sld_solvent</i>
'core'	$+6.39 \times 10^{-6} \text{ \AA}^{-2}$	$+0.62 \times 10^{-6} \text{ \AA}^{-2}$	$-0.28 \times 10^{-6} \text{ \AA}^{-2}$
'shell'	$+6.39 \times 10^{-6} \text{ \AA}^{-2}$	$+0.62 \times 10^{-6} \text{ \AA}^{-2}$	$+6.68 \times 10^{-6} \text{ \AA}^{-2}$
'drop'	$-0.55 \times 10^{-6} \text{ \AA}^{-2}$	$+0.62 \times 10^{-6} \text{ \AA}^{-2}$	$+6.68 \times 10^{-6} \text{ \AA}^{-2}$

The objective of this simultaneous fit will be to find common values for the *radius* (of the droplet cores) and *thickness* (of the surfactant layer) parameters that provide a good solution to all three datasets.

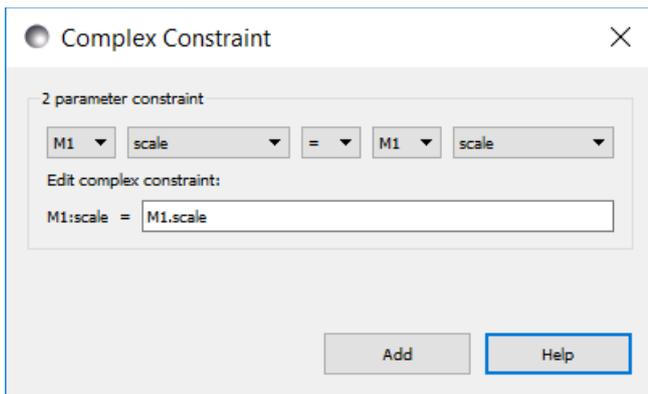
(Depending on how meticulous the person making the samples was, it might also be possible to find a common scale parameter, but for this example we shall assume that the volume fraction of droplets in each sample was different.)

We now need to tell SasView that the *radius* for the 'core' contrast dataset is the same as the *radius* for the 'shell' contrast dataset and is the same as the *radius* for the 'drop' contrast dataset, and so on. This is called introducing constraints. In each FitPage, check the *radius* and *thickness* parameters. However, we also need to optimise the *scale* and *background* parameters so check those too.

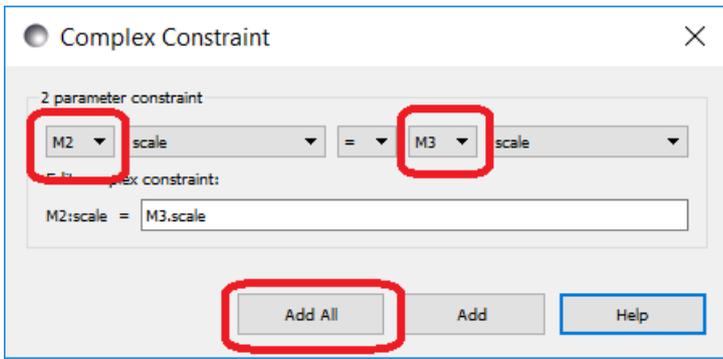
Now go to the Menu Bar and click **Fitting** followed by **Constrained or Simultaneous Fit**. In the Const & Simul Fit page, make sure that the boxes next to FitPage name are selected for those theories that you want to construct constraints for. For this example, all three theories should be selected.



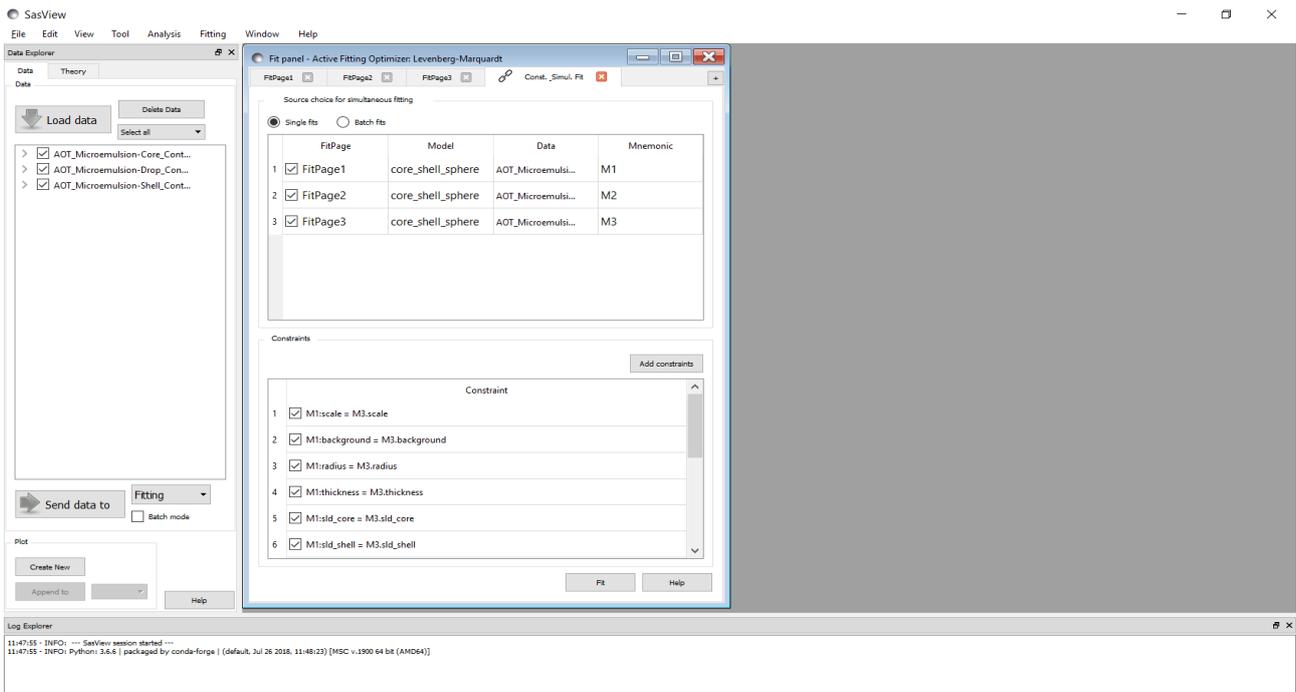
Then, in the section of the page called Constraints, click the **Add constraints** button. A new **Complex Constraint** dialog shows up.



To constrain all identically named parameters to fit simultaneously to the same value across all the selected theories we can use the **Add All** button. To enable it, two different models need to be selected. There are, however, several ways that we can set up the constraint equalities. Here we shall use the 'core' contrast (M3) as the reference. So set **M2 = M3** and click **Set All**. Then set **M1 = M3** and click **Set All**.



This will set up all possible constraints, which are listed in the Constraint table in the lower part of the tab.



Tip: If you need to scale parameter values between FitPage's then use the free-form constraint box below Easy Setup. The right-hand side of the equality can be of the form: $scalar \cdot [Mj].parameter_name$.

As we are assuming the volume fraction of droplets in each sample was different, remove the two *scale* constraints by right-clicking on the corresponding row and selecting **Remove constraint**. And because the different datasets represent samples containing different amounts of H & D (and therefore have different incoherent scattering contributions), also remove the two *background* constraints.

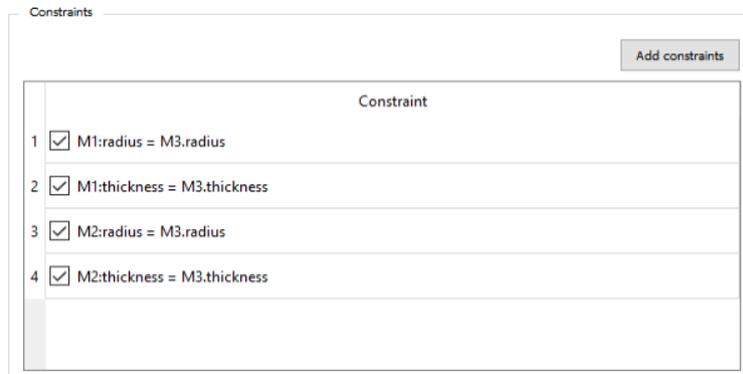
We are not fitting the sld parameters, so remove all rows with *sld*, that is: *sld_core*, *sld_shell* and *sld_solvent*.

But as it is the 'shell' contrast (M1) which should provide the most sensitivity to the *thickness* parameter, we need to change the constraint equalities dealing with *thickness* to read:

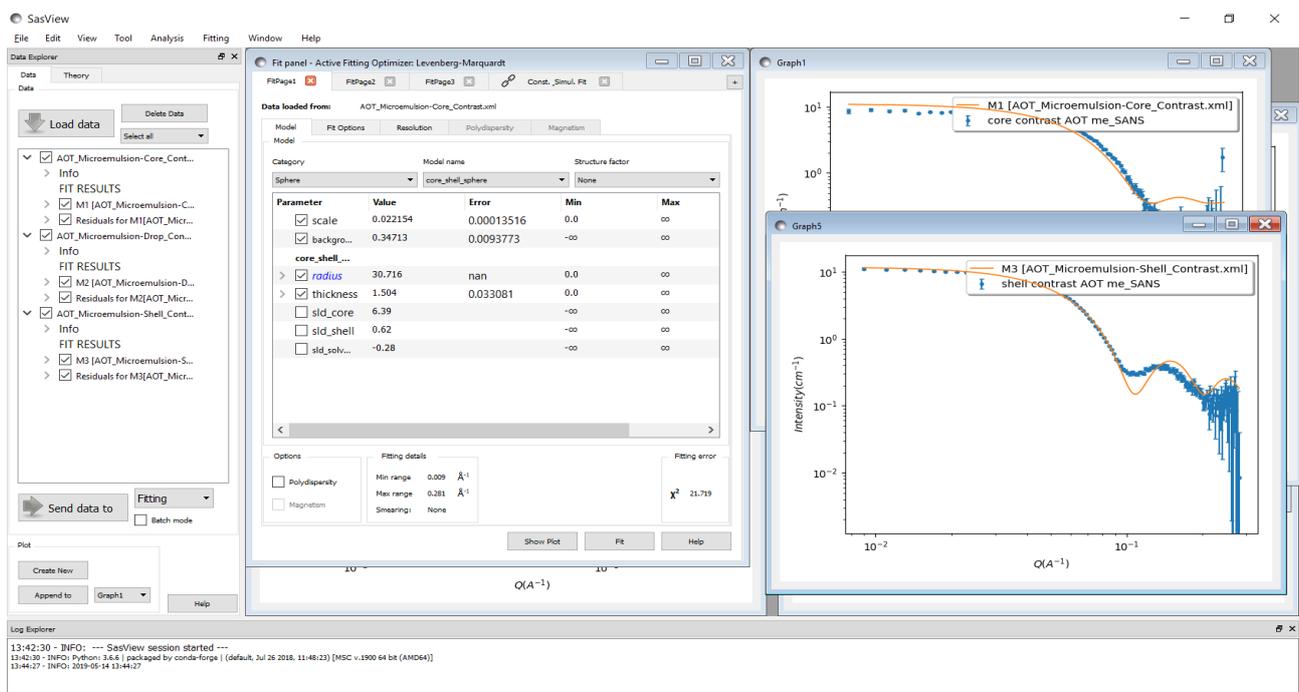
M2:thickness = M1.thickness
M3:thickness = M1.thickness

Double click on the corresponding row and edit the constraint directly to update it.

So we now have constraints that look like this:



When ready, click the **Fit** button on the Const. Simult. Fit tab. DO NOT click the Fit buttons on the individual FitPages.



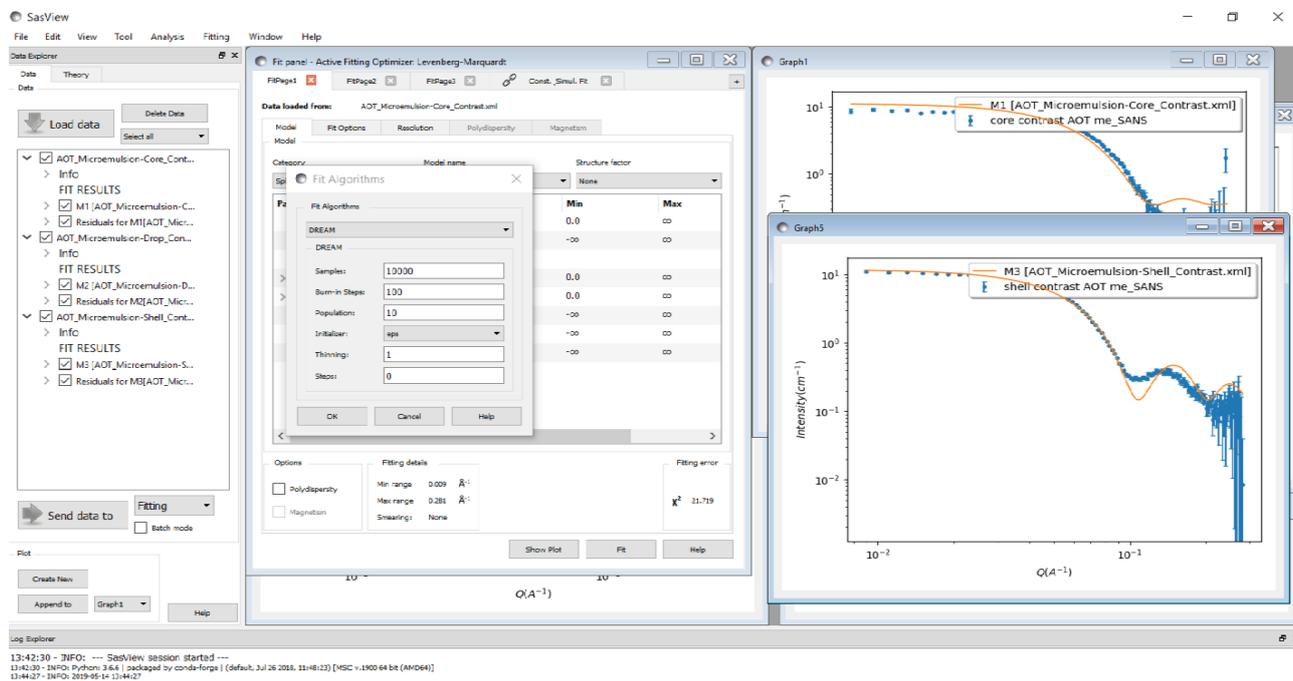
The results of the model-fitting are returned to each of the individual FitPage's. Note that the *Reduced Chi2* value returned is the sum of the *Reduced Chi2* of each of the three fits.

Tip: To see the *Reduced Chi2* value for a specific FitPage, click the Compute button at the bottom of that FitPage to recalculate it. Note that in doing so the degrees of freedom

will be set to N_{pts} . Moreover, in the case of constrained fits the number of degrees of freedom are less than one might think due to those constraints.

The fits returned leave room for improvement, particularly in the case of the 'drop' contrast. This might be because the default optimiser is struggling to handle three datasets. To test this, we can change the optimiser.

Go to the Menu Bar and click **Fitting** followed by **Fit Algorithms**.



Tip: Detailed information about the different optimisers available in SasView can be found in the Help Documentation.

The default optimiser, called the Levenberg-Marquardt, is a fast optimiser, but is prone to finding what are called 'local solutions' in parameter-space, rather than the true 'global solution'. More robust optimisers will be more likely to find the 'global solution' but will be slower. The most robust optimiser in SasView is the one called DREAM.

Select DREAM from the drop-down list of algorithms and click **Ok**.

Return to the Const & Simul Fit page and click **Fit**.

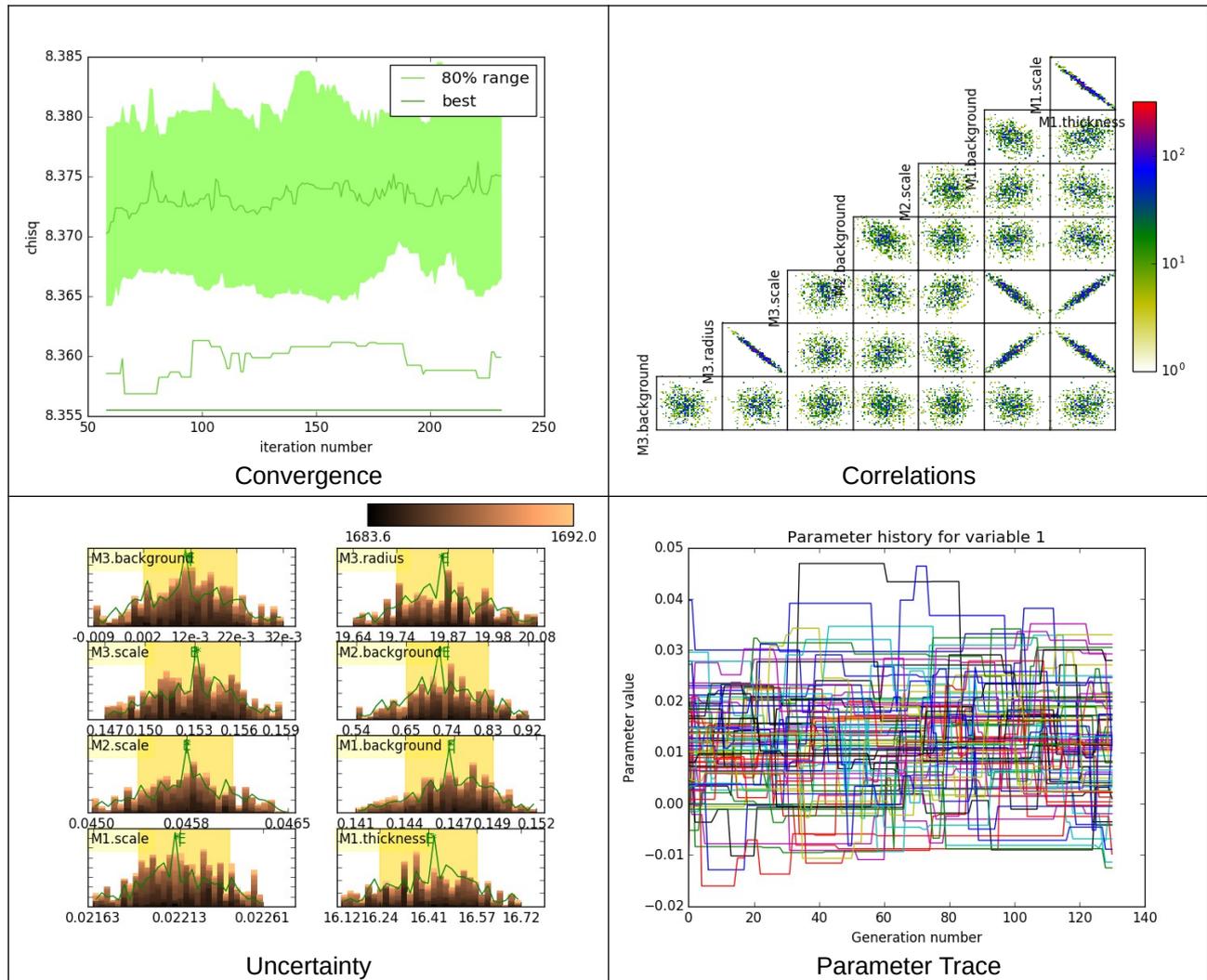
When the fit has completed, go to the Menu Bar and click **Fitting** and then select **Fit Results**.

The Result Panel is unique to DREAM and provides detailed insight into the quality of the fit through four displays:

- **Convergence:** how χ^2 has evolved during the fit
- **Correlations:** 'maps' depicting the relationships between parameters

- *Uncertainty*: the distribution of best fit values for each parameter
- *Parameter Trace*: the relative variation of each parameter

As can be seen, DREAM has highlighted strong correlations between the *scale* and *radius* parameters, in part because they are constrained between theories, but also because changing the radius of the droplets affects their volume fraction.



It is also evident that DREAM has been unable to improve the solution: *chisq* (how DREAM labels the *Reduced Chi2*) is essentially unchanged.

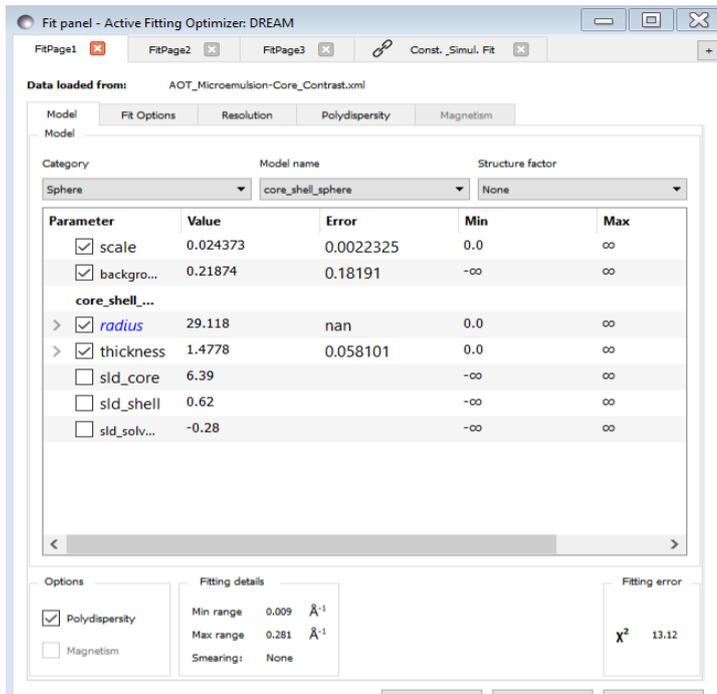
Go to each FitPage and turn Polydispersity **On**. In each case, set a *Distribution of radius* parameter of 0.15 with a **lognormal** distribution and check it so it will optimise.

Return to the Const & Simul Fit page click **Fit**.

Similarly, go to each FitPage and set a *Distribution of thickness* parameter of 0.2 with a **gaussian** distribution and check it so it will optimise. Then run the fit again.

WARNING: This fit may take ~20 mins to complete!!!

The resulting solution is the best that can be achieved with these data, suggesting the microemulsion droplets have a water core of *radius* $\sim 29 \pm 0.3 \text{ \AA}$ (with 12% Log-Normal polydispersity) and a surfactant shell layer $\sim 1.5 \pm 0.8 \text{ \AA}$ thick (with 36% Gaussian polydispersity).



Further Information

For further information, please consult the

SasView Tutorial Series

or

<http://www.sasview.org>

or email

help@sasview.org