Journal of
Applied Crystallography

ISSN 0021-8898

Received 31 May 2006
Accepted 30 August 2006

Reduction and analysis of SANS and USANS data using IGOR Pro

Steven R. Kline

NIST Center for Neutron Research, Stop 8562, Gaithersburg, MD 20899, USA. Correspondence e-mail: steven.kline@nist.gov

A software package is presented for performing reduction and analysis of small-angle neutron scattering (SANS) and ultra-small-angle neutron scattering (USANS) data. A graphical interface has been developed to visualize and quickly reduce raw SANS and USANS data into one- or two-dimensional formats for interpretation. The resulting reduced data can then be analyzed using model-independent methods or non-linear fitting to one of a large and growing catalog of included structural models. The different instrumental smearing effects for slit-smeared USANS and pinhole-smeared SANS data are handled automatically during analysis. In addition, any number of SANS and USANS data sets can be analyzed simultaneously. The reduction operations and analysis models are written in a modular format for extensibility, allowing users to contribute code and models for distribution to all users. The software package is based on Igor Pro, providing freely distributable and modifiable code that runs on Macintosh and Windows operating systems.

1. Overview

Small-angle scattering measurements have long been a powerful technique for the structural investigation of materials. Increasingly powerful sources have made it possible to collect large amounts of data in an experiment lasting only a few days. Having well designed software to process the large amount of raw data collected allows the experimenter to judge the quality of their data quickly, enabling them to make appropriate changes to their experimental plan, if necessary. An increasing number of experiments combine SANS and USANS measurements to extend the overall q range, and hence the length scales probed (1 nm to 10000 nm). Simultaneous analysis of USANS and SANS data makes the most of the information in the data.

For the reduction of data collected at a scattering facility, the procedures and code can often be somewhat site-specific for an instrument's raw data format. Allowing enough flexibility to work with data in other formats to accommodate the growing number of users that perform experiments at multiple facilities is a desirable feature. Providing software that can accommodate users from other facilities (ANSTO, ORNL, HANARO, for example) and easily incorporate suggestions for improvements [see canSAS (http://www.ill.fr/lss/canSAS/) and NOBUGS (http://nobugs.dl.ac.uk) user meetings] provides a great service to the scattering community. The package presented here has been separated into reduction and analysis modules since reduction modules are often facility-specific. The analysis module is inherently more generic, capable of reading in more general reduced ASCII data sets. The set of analysis models can also easily be extended to SAXS by changing the contrast term, or by following the supplied template to write SAXS-specific models.

The reduction and analysis capabilities of this package are similar in function to other programs, such as *FISH* (Heenan, 1999), *Grasp* (Dewhurst, 2002) and *ATSAS* (Konarev *et al.*, 2006), in providing a convenient interface for both reduction and analysis of experimental data. This package does not duplicate any indirect Fourier transform techniques or *ab initio* methods, such as those developed by Moore (1980), Brunner-Popela & Glatter (1997), Svergun (1999) and Chacon *et al.* (1998). The emphasis here is to provide reduction and analysis modules designed to be easy to use for novice users and also have features for more expert users. Standard experimental data can be processed with very minimal user intervention. There is also enough flexibility to handle non-standard experiments which are becoming much more common. This software allows users to reduce, analyze and produce publication graphics, all from a single interface. Directing the experimenter's focus on understanding the analysis, not the details of programming, maximizes the information that can be obtained from scattering data.

2. Key features

Over nearly seven years of use (since 1999) at the NIST Center for Neutron Research (NCNR), many new features and user suggestions have been incorporated into the package to tailor it to the needs of SANS instrument users. This ready extensibility of the package, even by the end user, remains one of the strengths of the software. All of the source code is visible, and help is provided for every button, panel and model function, so that users can gain a real understanding of what they are doing with their data.

2.1. Experiment planning

Before samples are even prepared, the analysis models can be used to simulate experimental data in a realistic way. The scattering from the expected structure can be calculated using physically relevant parameters to determine if the experimental conditions (sizes, concentration, contrast, *etc.*) will produce a measurable signal with identifiable features. Planning such as this provides a realistic view of what can be expected from an experiment, providing the best chance

for success. The best analysis methods cannot overcome a poorly designed experiment.

2.2. Reduction of SANS data

The mathematics of reducing raw SANS data has been refined at NIST over many years since the NCNR's SANS instruments were constructed (Glinka *et al.*, 1998). The graphical user interface (GUI) presented here is layered on this Fortran code, which was easily translated into the Igor Pro programming language. In addition, the GUI allows for many new features unavailable in the original VAX-based reduction software due to limited graphic capabilities. Making use of a modern data analysis package allows the visualization of the data to be a central part of the experiment. In fact, a major benefit of the GUI is to allow the user to see their data earlier, to identify problems, subsequently modify experimental plans, and ultimately collect better data.

From a main control panel, raw data files can be displayed and manipulated. Two-dimensional data are displayed in a window with controls to adjust the color table, scroll quickly through data files, and regroup the data to be displayed in one dimension. An example of this is shown in Fig. 1. A typical reduction session begins with a listing of all of the data files, presented in a familiar spreadsheet format. The file listing contains the pertinent information from the raw data header and is used interactively to select the proper files to construct a reduction protocol. The reduction protocol is a definition of the sequence of operations desired to produce corrected data. Typically, raw data are corrected for empty cell and background scattering, detector pixel non-uniformity, and converted to absolute intensity scaling. In the final step, reduced data can be written to disk after averaging to one dimension [$I(q)$ or $I(\varphi)$], or as two-dimensional ASCII files [I(pixel) or $I(q_x, q_y)$], or saved as a graphics image (PNG). A listing of the steps, files and parameters used to reduce the data are written to the reduced-file header, providing a record of how each file was processed. Once a reduction protocol is defined, raw data can be reduced one at a time, or in a batch mode. In either case, a two-dimensional image of each step of the reduction is briefly displayed on screen, allowing the experimenter to identify quickly any data file that may be abnormal. If any step in the reduction process seems to be incorrect, a one-page printable schematic can be produced that contains a thumbnail image of the two-dimensional data set at each step in the protocol, along with the parameters and equations that were used. This greatly simplifies troubleshooting of errors in data processing.

A large fraction of SANS samples scatter isotropically, and can be simply reduced to one-dimensional data, $I(q)$. SANS data that have been reduced to one dimension [$I(q)$] are often collected with more than one instrument configuration for each sample. A utility is available to take up to three $I(q)$ data sets, trim data points from the ends of each set, find the scaling of the overlap regions (or not), and combine them into one data file (Fig. 2). One-dimensional data can further be processed with a utility that allows subtraction of two one-dimensional data sets (subtracting buffer scattering from a protein solution, for example), subtraction of a constant, or rescaling.

In addition to the 'standard' data processing utilities, there are a variety of two-dimensional operations where two-dimensional data sets can be directly manipulated using simple arithmetic operations. Data sets that are anisotropic can also be averaged in angular sectors or rectangular stripes, or as a function of azimuthal angle at a constant q value. Anisotropic data can also be kept in a two-dimensional format for analysis. A simple two-dimensional Gaussian fit is provided for peak location. More two-dimensional analysis features will be added in the future as requested.

2.3. Reduction of USANS data

The USANS instrument at NIST is of Bonse–Hart type using triple-bounce silicon crystals for the monochromator and analyzer (Barker *et al.*, 2005). As a result, data are collected point-by-point in

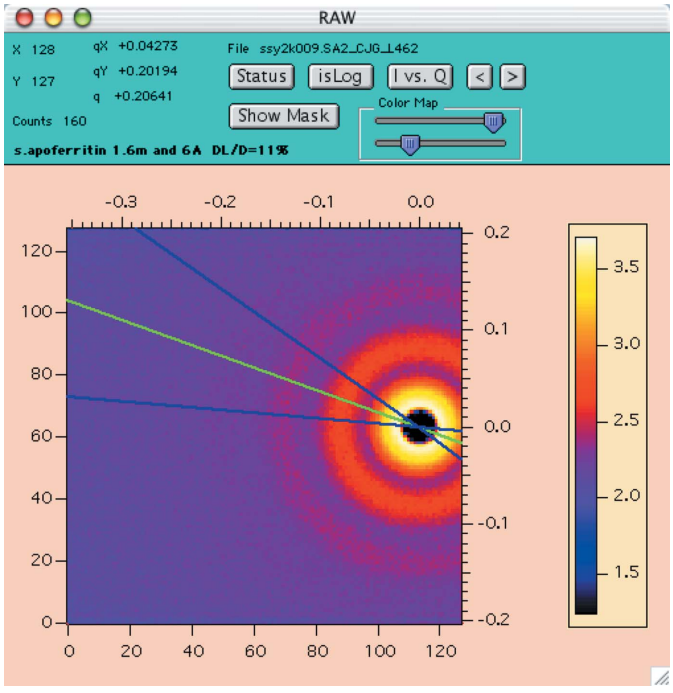

Figure 1
The SANS data display presents the two-dimensional detector data in a raw pixel format or can display any of the steps during data reduction. Sectors and annulus for averaging to one dimension are selected interactively. Color mapping can also be adjusted to better visualize features in the data.

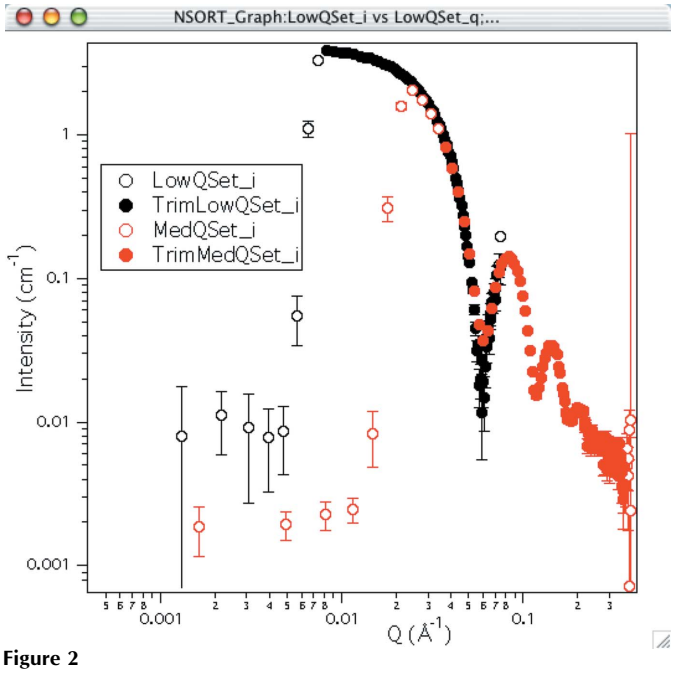

Figure 2
Reduced SANS data from up to three q ranges can be trimmed, overlapped, sorted and rescaled to provide a single data set for analysis. Open symbols are trimmed from the final data set, while the filled symbols are rescaled to overlap, if desired.

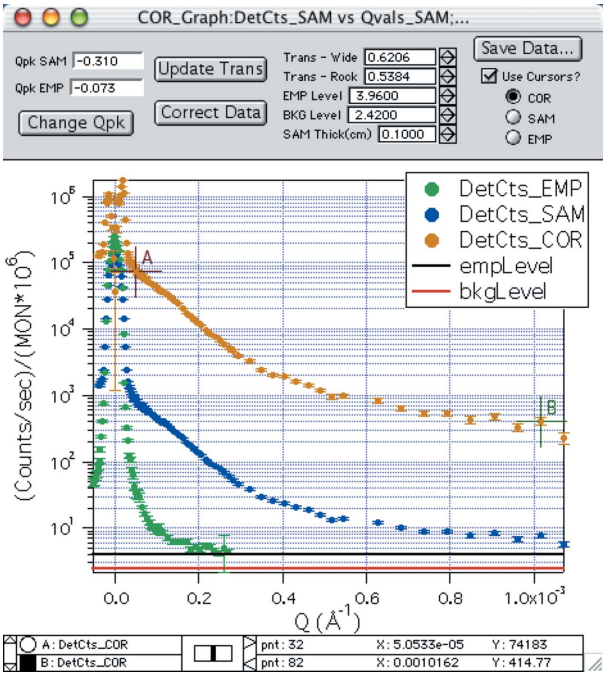

Figure 3
Correction of raw USANS data works with the raw one-dimensional data, producing corrected slit-smeared data and transmissions for estimating multiple scattering contributions. $Q = 0$ is found automatically from the peak intensity of the main beam. Extrapolated background and empty cell scattering can be used for the correction, and their contributions are graphed along with the sample scattering. The final corrected data can be exported as an ASCII file.

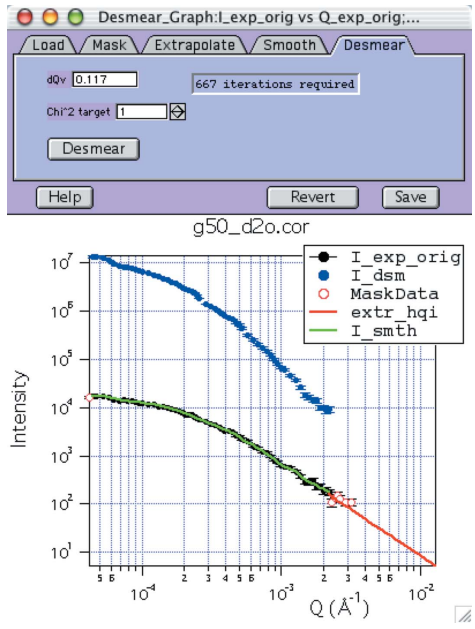

Figure 4
USANS data can be desmeared in an interactive panel. Slit-smeared data are loaded, masked, extrapolated to high Q, and smoothed if necessary before desmearing. Any of the pre-treatment steps can be fine-tuned or skipped if desired, and all steps are shown on the graph to ensure that the original data are not severely distorted. Desmearing is done using Lake's iterative technique.

discrete angular steps, making the raw data inherently one-dimensional. In a typical experiment, a series of files are collected at different scattering angle steps and must be spliced together. The lowest-angle set is rocked through the main beam to define zero scattering angle and then larger angle data are collected, usually on logarithmic angle spacing. The sets of files that comprise the sample or cell scattering are selected as a group and plotted. At this step, the raw data sets are joined, sorted, shifted to zero angle, and transmissions are automatically calculated from the data. The user can then set the angle-independent background level and the large-angle level of the cell scattering. Then one click produces corrected data. An example of this is shown in Fig. 3, where the sample (SAM) and empty cell (EMP) scattering are shown along with the corrected (COR) data set. Any subset of the raw or corrected data can be saved in ASCII format by interactively selecting the range of data to save. As a diagnostic, the multiple scattering probability can be easily calculated from the ratio of the wide-angle and zero-angle transmissions.

Desmearing of the data can be performed using Lake's iterative technique (Lake, 1967). The slit-smeared data can be interactively masked and smoothed before desmearing, if desired (Fig. 4). Often the best results are obtained without smoothing, which can introduce additional artefacts. Desmearing is typically not necessary, since the slit-smeared USANS data can be used 'as is' by fitting to appropriately smeared structural models as described below.

2.4. Analysis of reduced data

The analysis package provides tools for model-independent analysis and for using non-linear least-squares methods to fit models of scattered intensity to experimental data. Descriptions of other analysis techniques and model functions can be found in texts (Glatter & Kratky, 1982; Feigin & Svergun, 1987; Roe, 2000) and review articles (Pedersen, 1997).

2.4.1. Model-independent analysis. Modules for calculation of the invariant, linearized fits, global analysis and summed models can be dynamically loaded for use. The model-independent methods require no *a priori* knowledge about the sample, and most importantly, no experimenter bias of assumed structure. Model-independent methods include linearized fitting (Guinier, Kratky, Porod, power-law fits, and more) to extract structural information. Invariant calculations can also be made to calculate sample parameters, such as volume fraction or specific surface area, or area per head group in a surfactant aggregate. An example of a Guinier fit to a cylinder to determine the radius of gyration is shown in Fig. 5. Fitting ranges are defined interactively, and the results with their statistical uncertainties are displayed on the graph.

2.4.2. Non-linear model fitting. Models of a wide variety of structures and systems can be plotted with experimental data for quick comparison. Model functions can be included and plotted from a control panel that also provides direct access to help files for every model function. This is by no means an exhaustive collection, but contains a large number of functions that can quickly be compared with scattering data (see Table 1 for a partial listing of the available models). If a good fit cannot be found with these models, they will likely point the experimenter in the proper direction to search for a better description of their sample. Plotting a model function generates a graph of the scattering function, and a spreadsheet-like table of the free structural parameters of the model. Model calculations and plots are interactive, where changing any of the parameter values in the table produces an automatic recalculation of the model and an update of the graph. This provides immediate feedback about how a change in one of the physical parameters of the model (radius, contrast, concentration, *etc.*) affects the model scattered intensity. The visualization provides quick feedback as to whether or not the chosen model is an accurate description of the data. If so, the

Table 1
Representative list of available model functions.

This is only a partial listing of included models. All of the intensity models can be smeared with either pinhole or slit resolution.

Form factors	Polydisperse spheres	Core–shell spheres
	Multilamellar vesicles	Rigid cylinders
	Polydisperse cylinders	Flexible cylinders
	Core–shell ellipsoids	Triaxial ellipsoids
	Parallelepipeds	Lamellae
	Tactoids	
Structure factors†	Hard spheres	Screened coulomb
	Square well	Sticky hard spheres
Hard spheres (exact)	Polydisperse hard spheres	Binary hard spheres
Polymers	Debye	Polyelectrolytes
	RPA for copolymers (ten cases)	
Two-phase materials	Peaks (Lorentzian, Gaussian)	Power law
	Debye–Anderson–Brumberger	Fractal
	Unified power law + R_g	Teubner–Strey

† Structure factors can be combined (with approximations) to most of the polydisperse or anisotropic form factors.

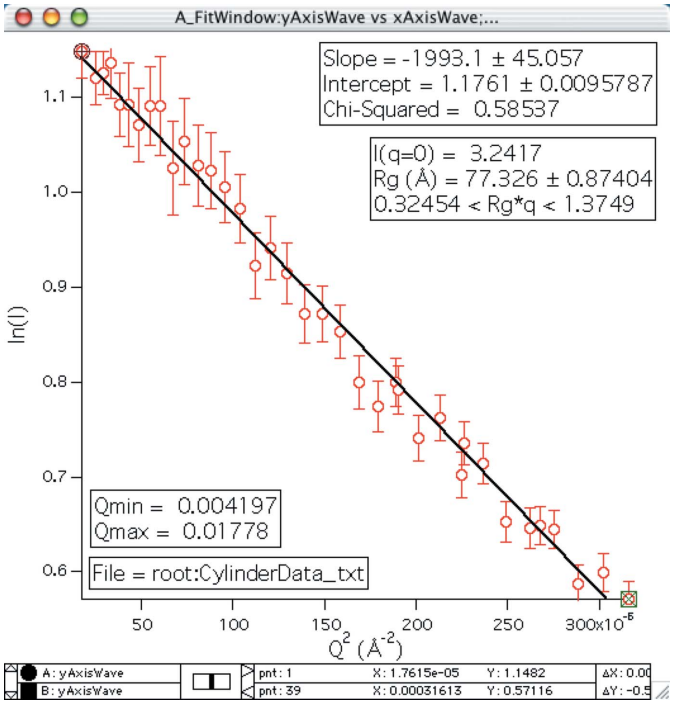

Figure 5
Linearized fit (Guinier) of scattering from a cylindrical particle. The fitting range is chosen interactively, and the results and statistics of the fit are displayed on the graph.

manually entered values provide a very good initial guess for the optimized fitting.

Models are included to calculate the scattered intensity for monodisperse, polydisperse, and interacting particles, as well as scattering from polymers and co-polymers, and general two-phase systems. Users can quickly compare the scattering from a large variety of different structures without investing large amounts of time programming. This does not absolve the user from understanding what they are doing. Some model functions have limited ranges of applicability due to assumptions in their derivation. Help files are included for every model function, showing exactly what is calculated and including the original reference for the model. The user must, of course, make sure that the fitted results are physically meaningful and do not violate the applicable range or any of the assumptions of the model.

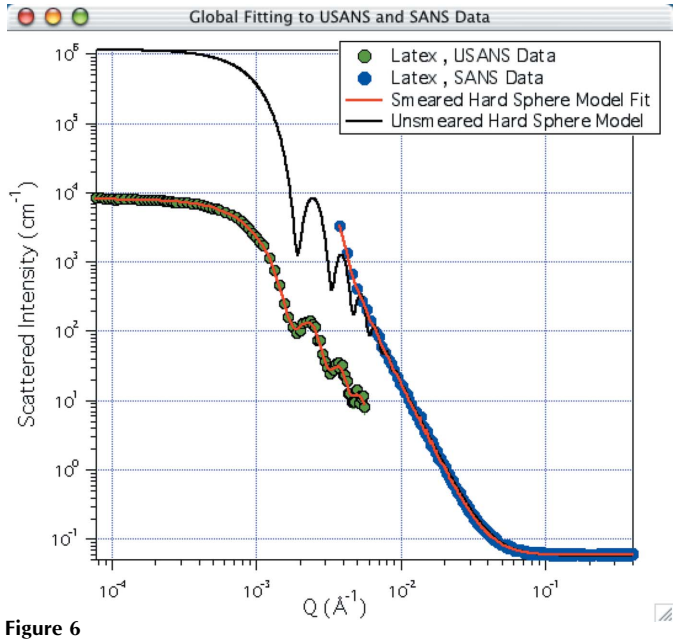

Figure 6
Global fitting can be performed for a single model function to SANS and USANS data simultaneously. The data points are experimental data from a dilute dispersion of nominal 500 nm diameter latex particles in heavy water. Note that the scale and slope of the slit-smeared USANS data do not overlap with SANS data, but the fitting correctly incorporates resolution. The unsmeared model function of polydisperse hard spheres is also plotted over the full q range to show the different effects of slit and pinhole smearing.

Once the raw data have been reduced to one-dimensional form, there is no difference in how USANS or SANS data are analyzed. All of the models that calculate form-factor scattering or total intensity can be smeared with either pinhole or slit resolution information, which is an integral part of the reduced data files. The ASCII output of both SANS and USANS reduction is in a six-column format, where the first three columns are q values, intensity, and the standard deviation of the intensity. Both types of data are smeared with instrument resolution, so the next three columns describe the resolution function. For SANS, these three columns are the standard deviation of the q resolution (σ_q), the mean q value ($\bar{q}$), and a beamstop shadowing factor (f_s), respectively. The instrument resolution function for (pinhole) SANS is approximated as a Gaussian function (Barker & Pedersen, 1995), so the convolution is performed using Gaussian quadrature quite efficiently. The smeared intensity is calculated at the measured q values, q_0, as

$$I_s(q_0) = \int_0^\infty R(q, q_0) I(q)\, dq. \quad (1)$$

The Gaussian approximation to the resolution function is

$$R(q, q_0) \equiv \frac{f_s}{(2\pi\sigma_q^2)^{1/2}} \exp\left[\frac{-(q-\bar{q})^2}{2\sigma_q^2}\right] \quad (2)$$

and is calculated from the values contained in the data file. If the resolution function extends to negative q values for a given q_0, the resolution function is truncated at $q = 0$. This typically only occurs for q values that are shadowed or are behind the beamstop and should be discarded from the data set. USANS data files have a single value, $-\Delta q_v$, written in all values of the three resolution columns, where Δq_v is the maximum extent of the vertical resolution and is a constant value for the instrument. The slit-smearing for USANS is then

performed using an adaptive trapezoidal integration (Press *et al.*, 1992) according to

$$I_s(q) = \frac{1}{\Delta q_v} \int_0^{\Delta q_v} I\left[(q_v^2 + q^2)^{1/2}\right] dq_v. \quad (3)$$

The effect of slit smearing on the appearance of the measured data is much more significant than pinhole smearing. Scattered intensity over a broader q range contributes to each slit-smeared q value more than for pinhole smearing where the resolution is confined to a much more narrow Gaussian at each q value.

Pinhole smearing of SANS data is not severe in many cases, and may appear to be fitted well with an unsmeared model function, but should always be fitted with a smeared model as the last step. The parameters found from the unsmeared model fit provide an excellent starting guess for the smeared fit, saving computation time. Accounting for the instrument resolution is especially important when a sample has been measured on both the SANS and USANS instruments. Visually, plotting USANS (slit smeared) and SANS (pinhole) data together looks wrong (data points on Fig. 6) since the data will not overlap due to the vastly different levels of smearing. In fact, even desmeared USANS and (pinhole) smeared SANS data should typically not overlap. Resolution smearing effects at the lowest q values of the SANS data are not insignificant and typically increase the observed slope from the true (unsmeared) value, which will be observed at the highest q values of the desmeared USANS data. Fig. 6 shows the results of globally fitting a single-resolution smeared model function to SANS and USANS data from nominal 500 nm diameter latex particles in D_2O. Model functions can also be fitted globally to a set of scattering data such as a series of measurements taken at different scattering contrast, to better refine complex structures. Global model fits are done through an interactive panel, without the need for modification of the original data files.

Igor Pro's nonlinear fitting uses the Levenberg–Marquardt technique, and Igor Pro's curve fitting supports simple bounding constraints as well as more complex constraints on the values of any linear combination of the fit coefficients. The goodness of fit is reported as χ^2:

$$V_{\chi^2} = \chi^2 = \sum_{i=1}^{N} \left(\frac{y - y_i}{\sigma_i}\right)^2, \quad (4)$$

with the usual definitions for the model y, the observed y_i, and its associated uncertainty, σ_i. A reduced χ^2 is reported for the linear fits (see Fig. 5) by normalizing χ^2 by $(N - p)$, where N is the number of data points and p is the number of free parameters in the fit. χ^2 has statistical significance only if proper uncertainties are used during the fitting. Covariance and confidence intervals for the fit can also be calculated, and are very useful for judging whether fitted parameters are independent and well determined. The computation time required for model calculations increases with the complexity of the model, especially the nested levels of numerical integration used for anisotropic structures and particle size distributions. Most fits, including those with resolution smearing, are completed within a few seconds.

2.5. Extensibility

One of the most useful features of the package is that the code is completely open and extensible by the end user. Curious users can inspect the code to see exactly what calculations are being performed in any reduction or analysis step. Tutorials and sample data sets are provided for all reduction and analysis operations. Help files and documentation are also an integral part of the package, and can be added to compliment a new analysis model. Modifications and new models coded by end users are encouraged since additions can easily be incorporated into the distribution package for all users in the future. For this purpose, templates with line-by-line instructions are provided for users that want to write their own model function. Using the supplied templates with step-by-step instructions, a user can easily code in a functional form not yet available in the package. Very little programming knowledge is required.

3. Hardware and data requirements

These procedures were developed using Igor Pro version 4.09, and are fully compatible with Igor Pro 4 and 5. Igor Pro is available for Macintosh OS 9.1+, OS X 10.2+, and Windows 98, 2000, ME and XP (http://www.wavemetrics.com). It is important to note that all reduction and analysis operations can be used with the free demonstration version of Igor Pro. Reduced data and model calculations are exported as ASCII files for use with other graphics packages, if desired. Currently, the reduction module distributed from NIST reads only NCNR SANS and USANS raw data formats. However, the reduction package has been easily modified by J. Schulz to read AUSANS (Australia) data and by T.-H. Kim to read HANARO (S. Korea) SANS data. These modifications are available from the respective institutions. Work is underway to generalize the file loaders to accommodate other raw data formats including NeXus (NeXus Committee, 2006), XML and ASCII data formats. The analysis modules require standard ASCII data with a minimum three-column format [q, $I(q)$, σ_I]. Optional resolution information can be added to non-NCNR data following Barker & Pedersen (1995) and the documentation included with the analysis modules.

4. Distribution

The reduction and analysis routines, documentation, examples, tutorial data sets, and model function templates are freely available at http://www.ncnr.nist.gov/programs/sans/data/red_anal.html. Source code for all of the analysis model calculations is provided as Igor Pro procedure files (text format) and is easily readable (syntax similar to C or Fortran).

5. Summary

The small-angle-scattering reduction and analysis package presented here provides any easy to use suite of tools for processing and analyzing scattering data. It is designed for ease of use, streamlining the process of reducing data while allowing flexibility for custom data manipulation. Analysis tools include model-independent calculations, linearized fitting, non-linear model fitting, and global analysis of multiple data sets. Documentation is included for every processing step and analysis model, along with detailed examples and test data. As an experimental planning tool, it provides a valuable teaching aid to help those new to scattering to develop a feel for how different structures will scatter, and how changing experimental conditions will affect the scattering. Finally, the source code for the entire package is freely available, allowing anyone to develop additional tools and models that can be made available for the entire user community.

The NIST SANS group is thanked for contribution of several analysis models, and for extensive suggestions, testing and error

checking. Several of the models included in this package are based on code developed by students of Dr Eric Kaler at the University of Delaware. The mention of commercial products does not imply endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose. This material is based upon activities supported by the National Science Foundation under Agreement No. DMR-9986442.

References

Barker, J. G., Glinka, C. J., Moyer, J. J., Kim, M. H., Drews, A. R. & Agamalian, M. (2005). *J. Appl. Cryst.* **38**, 1004–1011.
Barker, J. G. & Pedersen, J. S. (1995). *J. Appl. Cryst.* **28**, 105–114.
Brunner-Popela, J. & Glatter, O. (1997). *J. Appl. Cryst.* **30**, 431–442.
Chacon, P., Moran, F., Diaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.
Dewhurst, C. (2002). *GRASP*. http://www.ill.fr/lss/grasp/grasp_main.html.
Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. New York: Plenum.
Glatter, O. & Kratky, O. (1982). Editors. *Small-Angle X-ray Scattering*. New York: Academic Press.
Glinka, C. J., Barker, J. G., Hammouda, B., Krueger, S., Moyer, J. J. & Orts, W. J. (1998). *J. Appl. Cryst.* **31**, 430–445.
Heenan, R. K. (1999). *FISH*. http://www.isis.rl.ac.uk/largescale/loq/FISH/FISH_intro.htm.
Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.
Lake, J. A. (1967). *Acta Cryst.* **23**, 191–194.
Moore, P. B. (1980). *J. Appl. Cryst.* **13**, 168–175.
NeXus Committee (2006). NeXus data format for neutron science: http://www.nexus.anl.gov/index.html.
Pedersen, J. S. (1997). *Adv. Coll. Interface Sci.* **70**, 171–210.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C*. New York: Cambridge University Press.
Roe, R.-J. (2000). *Methods of X-ray and Neutron Scattering in Polymer Science*. New York: Oxford Press.
Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.