

# **Data Reduction for the SANS Instruments at Oak Ridge National Laboratory**

William T. Heller\*, Changwoo Do, Christopher B. Stanley, Wei-Ren Chen, Sai Venkatesh Pingali, Shuo Qian, Volker S. Urban, Lisa DeBeer-Schmitt, Kenneth C. Littrell, Lilin He, Ricardo M. Ferraz Leal, Jose M. Borreguero Calvo, Mathieu Doucet, Garrett E. Granroth, Michael R. Fitzsimmons, Tianhao Wang, Bin Wu, Matthew Cuneo

Neutron Scattering Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831.

\*contact email: [hellerwt@ornl.gov](mailto:hellerwt@ornl.gov)

## **Abstract**

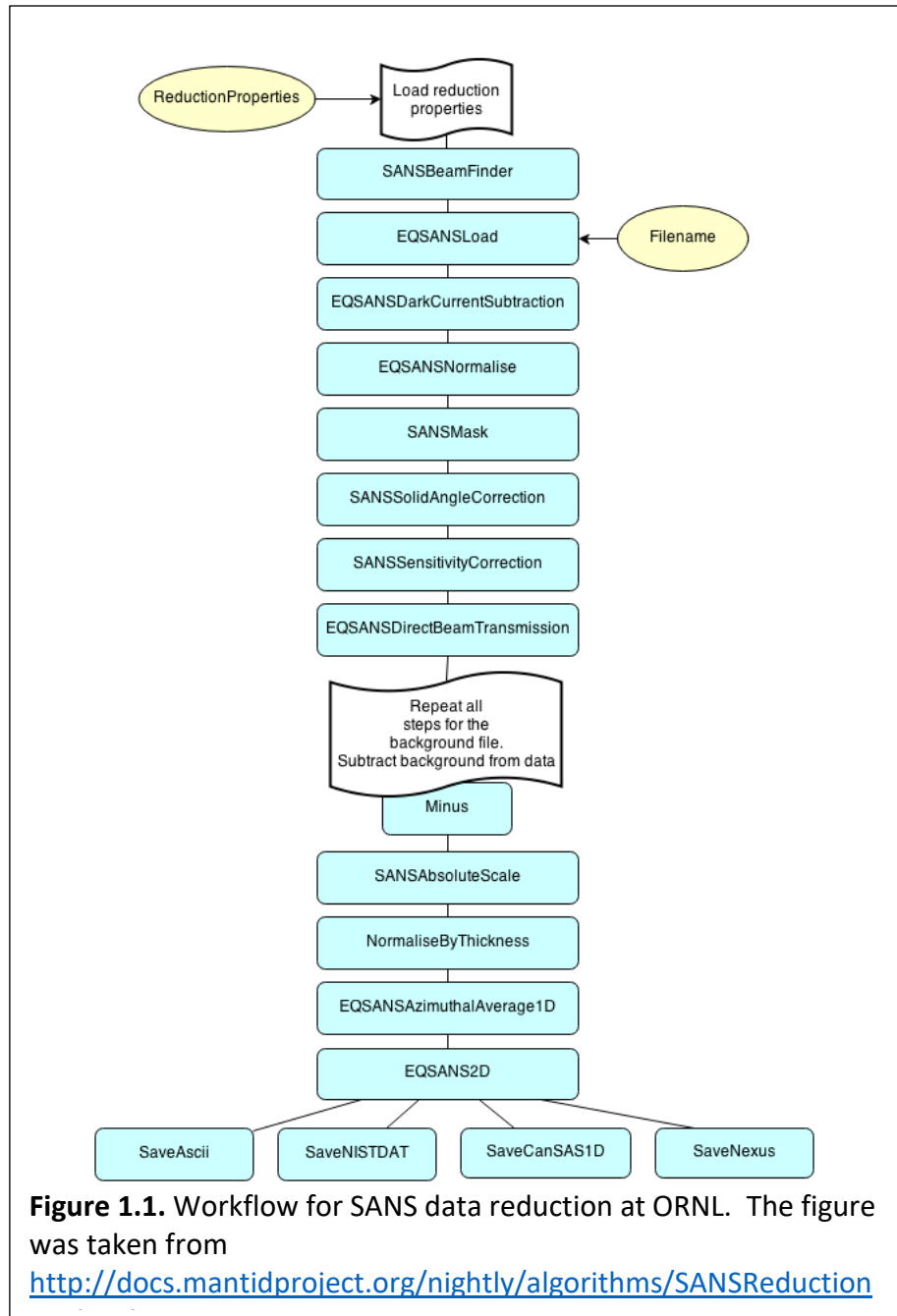
The neutron scattering facilities at Oak Ridge National Laboratory are home to three small-angle neutron scattering (SANS) instruments. The General Purpose SANS (GP-SANS) and Biological SANS (Bio-SANS) are located at the High Flux Isotope Reactor, while the Extended Q-range SANS (EQ-SANS) is located at the Spallation Neutron Source. Recently, an effort to unify the data reduction methods across the three instruments was undertaken. Herein, the data reduction methods that have been implemented in Mantid is presented. Details of the various steps in the process are provided, as is how the methods described herein work together to provide reduced SANS data.

## 1. Introduction

The suite of small-angle neutron scattering (SANS) instruments at Oak Ridge National Laboratory (ORNL) [8] includes three pin-hole SANS instruments. The High Flux Isotope Reactor (HFIR) is home to the General Purpose SANS (GP-SANS) and Biological SANS (Bio-SANS) instruments, while the Extended Q-range SANS (EQ-SANS) instrument, which is a time-of-flight (TOF) instrument, is located at the Spallation Neutron Source (SNS). Historically, the data reduction employed at the two facilities was performed using approaches implemented by different people in different software packages. Data reduction of EQ-SANS utilized Mantid [1], while the data reduction for the HFIR SANS was originally implemented in IgorPro (Wavemetrics, Inc., Portland, OR). GrASP, which was developed at the ILL [2], is also used on GP-SANS. Mantid-based data reduction software was also available for the HFIR instruments, but it was less commonly used.

Here, we present the methodology to post process data (i.e., the process of data reduction) with the aim to provide the user with the differential scattering cross-section as a function of wavevector transfer,  $\frac{d\sigma}{d\Omega} = S(\vec{Q})$ .  $S(\vec{Q})$  may also be a function of time for an experiment when the state of the instrument and/or sample is changing with time. The

methodology was provided to a software programming team—the job of this team was to produce a unified software package implementing the data reduction methodology. The text following describes the operations of the operations shown in the workflow plan (Figure 1.1). The methods presented here were implemented in the



Mantid software package [1], and a layer above the underlying methods was developed to facilitate its use by the users of the ORNL neutron scattering facilities.

## 2. Read the Required Data Files

SANS data reduction requires considerable information beyond the scattering data. In addition to the scattering data, the following additional files are required sample's transmission, that of a reference sample and one or more measurements of background (from the instrument, neighboring instruments, sample environment, etc.). Measurement of detector noise—a so-called dark current measurement—or with the beam blocked at the sample position may also be used for the data reduction as another form of background measurement. The position of the center of the beam, possibly as a function of wavelength, needs to be collected. In all cases, these files are measured as part of an experiment.

Sample scattering data files are collected when the sample is placed in the sample position and the beam-stop is positioned to shield the footprint of the direct beam on the detector array. The transmission data files are collected with the beam-stop being moved away from the footprint of the direct beam on the detector array. The direct beam is normally attenuated in order to protect the detectors during the transmission measurements. In certain conditions, such as when using long wavelength neutrons with the detector positioned far away from the sample, it is possible to simultaneously measure scattering and transmission. The Bio-SANS also has a semi-transparent beam-stop, which enables a simultaneous measurement of scattering and transmission beams. In these cases, a single data file with both scattering and transmission data is saved. If one uses either an empty cell, a pure solvent or other tunable parameters such as temperature as the background, the background data files are collected no differently than the sample scattering and transmission data files. A blocked beam measurement, where a strong absorber is placed at the sample position or in place of the

sample aperture may be collected for use as a background. A dark current measurement, representing electronic noise and cosmic radiation, may be used instead of a blocked beam measurement. This data would be collected with either the shutter closed, or during a maintenance period when no beam is being produced. An empty direct beam measurement is required for every instrument configuration used during an experiment. This measurement is often performed with nothing in the sample position, which also makes the measurement well-suited for use as the transmission reference measurement, as well as for determining the beam position, but large, complex sample environments may make a true empty beam measurement impossible.

Two different formats for the data files are used at ORNL. Measurements made with the HFIR instruments are saved as XML files that contain both the raw data and metadata (i.e., measurements related to conditions of the experiment, e.g., beam polarization, motor position, sample environment parameters) that is related to the measurement. The raw data is saved as a matrix of counts in each pixel of the detectors in the instrument. The XML file makes it possible to view the metadata and the final detector image, also saved by the data acquisition system as a conveniently-formatted web page. In contrast, the time-of-flight EQ-SANS (and the GPSANS machine using a modified HFIR-lite software) provides data as an event stream in which the pixel and arrival time of each neutron captured by the detector are saved. Experiments using a polarized beam on GP-SANS are saved as an event stream. Similarly, studies of kinetics having a clear trigger may also benefit from having data saved as a stream of events. Within a year, data will be recorded using event mode as a default for all ORNL SANS machines. Regardless of the format, the data files contain extensive metadata that are essential

for performing the data reduction. An overview of the metadata saved in the data files is provided in Appendix 1, but some information in the metadata can vary, such as when the sample environment changes.

Several files related to the instrument that can be common across several experiments must also be read. Unlike the various kinds of files mentioned above, these files are either generated by beamline staff or a user or have been processed prior to their use. The sensitivity file, which is a measure of the uniformity of the response of the detector, is one example of such a file. Another example of this kind of file is a mask that identifies regions of the detector that should not be included during the data reduction. A general configuration file for the data reduction may be required, and a flux spectrum file is typically used when working with data from the EQ-SANS.

## **2.2 Measurements Employing Half and Full Polarization**

Measurements that employ either half or full polarization require special attention, which begins when the data file is read. As is noted above, such data are saved as event streams. The data file must be divided into either two or four data structures at the time the file is read that are processed through the desired data reduction steps as any other file until after the Subtract Background (Section 8, below). The data structures contain data collected when the first flipper and the polarization analyzer are in specific states, which are related to the various spin state SANS data sets of interest in the experiment. A description of the measurements performed and the process of determining the spin state SANS data are described in Section 9. The possibility that the beam position changes with the states of the



polarizer and analyzer is addressed by specifying different direct beam for each of the workspaces extracted from the event streams.

### **3. Establish the Instrument Geometry and Prepare the Data**

To reduce SANS data, the physical layout of the important pieces of hardware must be known with respect to the direction of the beam. Much of the information contained in the metadata record locations and states of the various pieces of hardware that define the physical layout of the instrument. Two additional pieces of information are required before a coordinate system can be established for the instrument, namely the location of the pixels within the detector and the location of the beam on the detector. The sizes and locations of the pixels are either specified in the instrument definition file from the known geometry of the linear position sensitive detectors and the number of pixels that the length of each tube is electronically divided into, or they are determined through a series of measurements, called the bar scan calibration, that involves additional measurements and further processing, which are described in Appendix 2. The bar scan calibration is not universally used across the suite of SANS instruments. Additional steps are taken to prepare the data before proceeding with the reduction. A correction for the solid angle subtended by each pixel is applied. The time-of-flight in EQ-SANS data is corrected for the scattering angle before data are binned into wavelength. Finally, the uncertainties in the measured counts are estimated. Details are provided below.

#### **3.1 Determine the Beam Center**

Finding or setting the beam center is an important step in establishing the location of the instrument components in real space because it links the coordinate system of the detector to the vector defined by the incident neutron beam. The most common means for determining

the beam center in pixel space is to measure the direct beam, which is one of the data files required during the reduction. When polarization and analysis are employed, it may also be necessary to specify a set of beam center files to reduce the set of measured polarization conditions. It is also possible to specify the coordinates of the direct beam, which is often done for diagnostic purposes during an experiment. Errors in the beam center—if minor—will primarily result in distortions to the data at extreme low  $Q$ , but data reduced with a slightly mis-specified beam center can be used while preparing the instrument at the start of an experiment.

While it is possible to fit a direct beam measurement with a function to determine the position of the beam center, a simple center of mass calculation is universally used across the suite of SANS at ORNL. Note that even though EQ-SANS uses a band of neutrons for performing TOF SANS, the band is narrow, being less than 4 Å wide at the shortest sample-to-detector distances. Therefore, a wavelength-dependent beam center is not used because the difference in the effect of gravity for the shortest and longest wavelength neutrons in the band is less than the size of a pixel. The process employed follows.

1. The raw data is read and the location and sizes of the pixels are established according to the instrument definition file provided during data acquisition and stored in the data file. The pixel locations and sizes can be overwritten using an IDF obtained from a bar scan calibration, if one is provided.
2. The TOF information, if it is present in the data, is flattened such that the data are only counts in pixels.

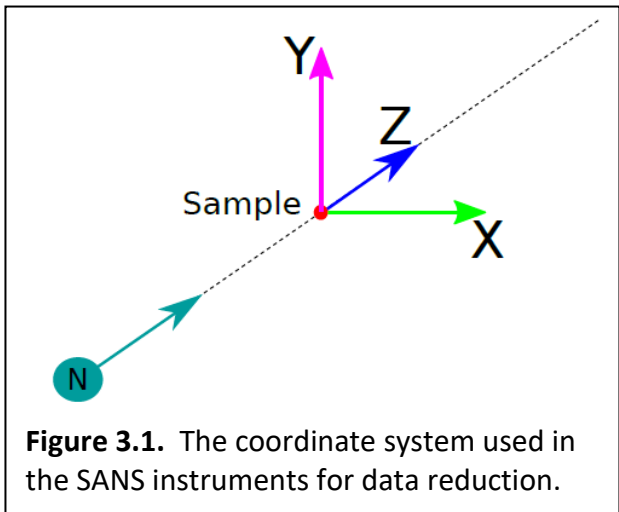
3. The required masks, if supplied, are applied to the data according to the description provided in Section 4 to exclude defective detector pixels from the calculation.
4. The data are normalized for the area of each pixel. If the height and width of the pixel having counts  $I(x, y, z)$  are given by  $(\Delta x, \Delta y)$ , then the data are normalized according to Equation 3.1.

$$I_{corr}(x, y, z) = \frac{I(x, y, z)}{\Delta x \Delta y} \quad (3.1)$$

5. The detector sensitivity can be applied to the data, if specified by the user, according to the method described in Section 5.
6. Finally, the center of mass  $\langle X, Y \rangle$  in pixel space is calculated using Equation 3.2. Note that the z-coordinate of the beam center is not calculated.

$$\langle X, Y \rangle = \left\langle \frac{\sum_n x_n I_{corr,n}(x, y, z)}{\sum_n I_{corr,n}(x, y, z)}, \frac{\sum_n y_n I_{corr,n}(x, y, z)}{\sum_n I_{corr,n}(x, y, z)} \right\rangle \quad (3.2)$$

The center of mass calculation is done in an iterative fashion. First, the center is roughly defined using a center-of-mass calculation from the entire detector. Then, the calculation is repeated within an area of the detector that is restricted around this nominal center. The restricted calculation is repeated until the beam center moves less than 0.25 pixels in either direction from one calculation to the next. The default radius for inclusion was set to 11 pixels. When a large number of pixels in the area of the beam center are defective, a peak fitting method may become necessary



**Figure 3.1.** The coordinate system used in the SANS instruments for data reduction.

for determining beam center as the excluded pixels may artificially shift the center of mass if they are asymmetrically distributed. Alternatively, the center of mass of a scattering pattern can be calculated according to Equation 3.2.

The method above is used for finding the beam center of the primary detectors of the EQ-SANS, Bio-SANS and GP-SANS, but cannot be used for the beam center for the data measured on the wing detector of the Bio-SANS. Two corrections must be applied to the beam center measured on the primary detector for use on the wing detector. First, the difference in the exact heights of the primary and wing detectors must be determined, which is accomplished by a direct beam measurement on the wing detector when it is rotated into the beam (set the motor WWDR to 0°) and with a second measurement of the beam on the primary detector when it is close to the sample position (sample-to-detector distance of 2.25 m is commonly used). This value has been measured to be 13.5 mm (roughly 3 pixels). Then, the trajectory of the neutron must be corrected for. The distance  $\Delta H$ , in mm, that the beam drops between the wing detector position and the primary detector position, which we denote  $\Delta L_2$  and is specified in mm, is given by Equation 3.3.

$$\Delta H = \frac{gm_n^2\lambda^2\Delta L_2^2}{2h^2} \quad (3.3)$$

where  $g$  is the gravitational constant (9.8 m/s<sup>2</sup>),  $m_n$  is the mass of the neutron (1.675E-27 kg),  $\lambda$  is the neutron wavelength in meters and  $h$  is Planck's constant (6.626E-34 m<sup>2</sup>kg/s).

### 3.2 Place the Instrument Components into Real Space

Determining the time-of-flight (TOF) of a neutron event and some data reduction operations, such as the solid angle correction, require that the components of the instrument

be positioned in 3-dimensional space. Doing so requires the information contained in the metadata from the scattering file, the beam center in pixel space, and the detector geometry from either the detector map obtained from the bar scan calibration or the known geometry of the detector. The coordinate system of the instrument is shown in Figure 3.1. The position of the sample is defined as the origin of the coordinate system. The positive Z-axis is in the direction of the neutron beam. The Y-axis is in the vertical direction, while the X-axis points horizontally to the right as viewed along the direction of the beam. While the locations of the detectors with respect to the sample are the primary concern, the locations of the sample and source apertures are also important because they are used to calculate the instrument resolution information, as described in Section 10. If the instrument is operating in TOF mode, which is always true for the EQ-SANS instrument and may be employed in the future if GP-SANS and Bio-SANS work in TOF mode with a chopper installed, the location of the source must be known because it is used to determine the TOF, and thus wavelength, of a detected neutron.

### 3.3 Assign the initial estimate of the uncertainty

The last step in preparing the data for reduction is to assign the initial estimate of the uncertainty in the measured counts. For a data set  $I(x, y, z, \lambda)$ , then the uncertainty associated with this measurement,  $\delta I(x, y, z, \lambda)$ , is calculated using Equation 3.4.

$$\delta I(x, y, z, \lambda) = \left( \left| I(x, y, z, \lambda) - \frac{1}{2} \right| + \frac{1}{2} \right)^{1/2} \quad (3.4)$$

Importantly, Equation 3.4 assigns the uncertainty of “0” counts in a pixel to be “1”, which is an appropriate account for the uncertainty of non-events.

### 3.4 Correct for dark current background signal

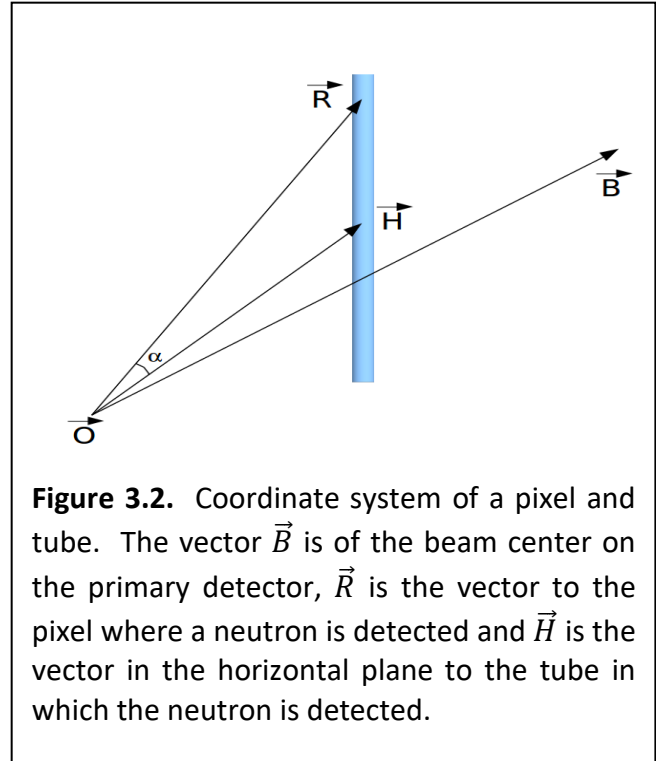
The dark current represents background signal that results from cosmic radiation, electronic noise, and sources of radiation not related to the neutron beam of the instrument. A dark current measurement that lasts for a few hours is performed during instrument calibration. Subtracting the dark current or blocked beam is optional. The EQ-SANS instrument almost always employs a dark current measurement during data reduction. In contrast, GP-SANS and Bio-SANS employ either a dark current measurement or correct for the blocked beam, which is described in Section 8. Data, which are denoted here as  $I(x, y, z, \lambda)$  with uncertainty  $\delta I(x, y, z, \lambda)$ , are corrected for the dark current background signal,  $I_{dc}(x, y, z, \lambda)$ , which has uncertainty  $\delta I_{dc}(x, y, z, \lambda)$  after being normalized for time, as described in Section 6 according to Equations 3.5 and 3.6.

$$I'(x, y, z, \lambda) = I(x, y, z, \lambda) - I_{dc}(x, y, z, \lambda) \quad (3.5)$$

$$\delta I'(x, y, z, \lambda) = \left[ (\delta I(x, y, z, \lambda))^2 + (\delta I_{dc}(x, y, z, \lambda))^2 \right]^{1/2} \quad (3.6)$$

### 3.5 Correct for the Solid Angle of a Pixel

The next step in preparing the data is correcting for the solid angle of a pixel, which arises from the fact that the detector pixels are not all the same distance from the sample position and may be different



sizes. Consider a scattered neutron that is collected in a pixel having dimensions  $p_x$  and  $p_y$  located at a position  $\vec{R} = \langle R_x, R_y, R_z \rangle$  in the coordinate system of the instrument. The sample is located at the origin of the coordinate system of the instrument,  $\vec{O} = \langle 0,0,0 \rangle$ . The suite of SANS instrument at ORNL are all equipped with area detectors that are arrays of linear position sensitive detectors (LPSDs), also referred to as tubes. A correction specifically for these tube detectors was derived by He and coworkers [4]. The detector pixels are actually volume elements within a tube. For the horizontal direction, the view of the tube from the sample position is independent of the scattering angle, which is in contrast to traditional, planar crossed-wire detectors [3]. In the vertical direction, it does vary with scattering angle. However, only the length of the pixel along the angle along the tube from the horizontal plane, which is denoted  $\alpha$ , must be corrected for its projection back to the sphere centered at the sample [4]. The most general formulation of the solid angle subtended by the pixel, which is suitable for both planar arrays of tubes and curved arrays of tubes, is given by Equation 3.7.

$$d\Omega_{tube}(\vec{R}) = \frac{p_x p_y \cos \alpha}{\|\vec{R}\|^2} = \frac{p_x p_y \cos \alpha}{(R_x^2 + R_y^2 + R_z^2)} \quad (3.7)$$

The angle  $\alpha$  is shown in Figure 3.2. The beam travels along the vector  $\vec{B}$  to the primary detector (i.e.  $\vec{B}$  is the vector to the beam center), and the position of the tube in the horizontal plane is given by the vector  $\vec{H} = \langle R_x, 0, R_z \rangle$ . The angle between  $\vec{R}$  and  $\vec{H}$  is  $\alpha$ , and  $\cos \alpha$  is calculated using Equation 3.8.

$$\cos \alpha = \frac{\vec{R} \cdot \vec{H}}{\|\vec{R}\| \|\vec{H}\|} \quad (3.8)$$

To correct the data for the subtended solid angle, the expression shown in Equation 3.9 is applied to the data.



$$I(x, y, z, t)_{corrected} = \frac{I(x, y, z, t)_{detected}}{d\Omega_{tube}(x, y)} \quad (3.9)$$

Note that there are no uncertainties related to the correction because the solid angle effect on the data is strictly a geometric effect. It is always applied to data read in by the data reduction process, unless it is a dark current measurement because the source of events is not related to the sample position.

### 3.4 Correct the time-of-flight for the scattering angle and histogram by wavelength

Prior to reducing TOF SANS data, they are converted into wavelength. An important part of this process is correcting the time-of-flight for the differences in path-length from the source to the pixels on the detector. Let  $L_1$  be the distance between the moderator-to-sample distance, and the vectors to the beam center and pixels be as defined in Figure 3.2. Then, if the measured time-of-flight of a neutron detected in the pixel at  $\vec{R}$  is  $t$ , the path-length-corrected time-of-flight of the neutron is given by Equation 3.10.

$$t' = \left( \frac{L_1 + \|\vec{B}\|}{L_1 + \|\vec{R}\|} \right) t \quad (3.10)$$

Next, the data are binned into wavelength histograms. Two options for the wavelength binning exist. Only neutrons not eliminated by the user-specified time-of-flight cut-offs are considered further. The original approach employed on EQ-SANS utilized constant wavelength bin widths of  $\Delta\lambda = 0.1 \text{ \AA}$ , and the option still exists. An alternative approach provided is to utilize constant  $\Delta\lambda/\lambda$ , which ensures the contribution of  $\Delta\lambda/\lambda$  to resolution in measurement of  $Q$  is constant; see Section 10. The minimum wavelength that is accessible on EQ-SANS is  $0.5 \text{ \AA}$ .

Select a width of the first bin  $\Delta\lambda_1$ , and the center of the first bin is  $\lambda_1 = 0.5 \text{ \AA} + 0.5\Delta\lambda_1$ . The centers and widths of the subsequent bins are given by Equations 3.11 and 3.12.

$$\lambda_{i+1} = \frac{(\lambda_i + 0.5\Delta\lambda_i)}{\left(1 - \frac{\Delta\lambda}{2\lambda}\right)} \quad (3.11)$$

$$\Delta\lambda_{i+1} = \lambda_{i+1} \frac{\Delta\lambda}{\lambda} \quad (3.12)$$

Assuming that  $L_1$  and  $L_2$  are provided in millimeters, then the wavelength of the neutron, specified in  $\text{\AA}$ , having the corrected time-of-flight  $t'$ , provided in microseconds, is given by Equation 3.13.

$$\lambda = \frac{3.9560346}{L_1 + L_2} t' \quad (3.13)$$

The neutron is then added to the appropriate wavelength histogram bin.

## 4. Apply Detector Mask

If a detector pixel, an entire LPSD, or an 8-pack of LPSDs (the electronic grouping used for the LPSDs) are performing poorly, including counts collected by a detector element that may adversely affect a measurement of  $S(\vec{Q})$ , it should be excluded from the data reduction process. Similarly, spurious signal that results from some aspect of the instrument configuration should not be included in the data reduction. Exclusion of detector elements from the data reduction is accomplished through “masking” of pixels. Rather than setting the values of the pixels to some value that the data reduction ignores during processing, a mask  $M(x, y)$  is a separate data structure having the same size and shape as the detector. Pixels to be used in the reduction are set to “1” in the mask, and those to be ignored are set to “0”. The data being reduced remain unaltered throughout the process. When the data are output after the reduction is finished, any 1D or 2D bins that contain only masked information are noted as “-inf” or “NaN” in the reduced data files.

There are several sources of masks, which are described below. It is possible to append masks together, which takes the form of a union of the set of masks. The data reduction log stores the final, assembled set of masks employed during the processing. The use of masks is optional.

- 1. Known Bad Detector Elements.** The LPSD arrays used on the ORNL SANS instruments are generally robust. However, individual LPSDs do fail, and there are times when an 8-pack of LPSDs does not function. In light of the redundancy of  $Q$ -space coverage, operations generally proceed provided that the malfunctioning tubes are not in the location of the direct beam. The EQ-SANS instrument employs a

configuration file that is linked to the run number employed by the data acquisition software that provides a means to identify regions of the detector that should be masked. A new configuration file is created at the run number where the operating condition of an LPSD or 8-pack is noted. The configuration file for the run being reduced is read during data reduction and the specified detector region is masked. This mask is also applied during the beam center determination, the transmission calculation and the creation of the sensitivity file. Generally speaking, this mask is always applied, but it can be omitted for diagnostic purposes.

2. **The Beam stop.** The quality of data taken at low- $q$  can be improved by masking out a region of pixels around the direct beam and beam stop. Doing so ensures that the contributions of parasitic scattering from the instrument and sample environment to the data reduction are minimized. On the EQ-SANS instrument, it is necessary to use such a mask because the edge of the beam stop in real space does not correspond to the same place in  $Q$ -space for the different wavelengths into which the data is binned. By failing to mask the region around the beam center, the low- $q$  becomes distorted when pixels behind the beam stop are averaged in  $q$ -space with those outside the beam stop. The set of beam stops on the SANS instruments are mounted immediately in front of the detector, and their motor positions can be used to determine which one of the set is being used for a given measurement. The radius of the beam stop, along with the measured beam center can be used to mask any pixels that are shadowed by the beam stop. The SANS instruments at HFIR

employ configurations where no beam stop, or a semitransparent beam stop, is used. This mask is optional.

3. **The Beam.** The radius of the beam,  $R_{beam}$ , can be directly determined from the size of the sample and source apertures ( $R_{sam\_ap}$  and  $R_{src\_ap}$ , respectively), their separation,  $SSD$ , and the distance from the sample aperture to the detector, which is approximately the same as the distance from the sample to the detector,  $SDD$ . The radius of the beam is given by Equation 4.1.

$$R_{beam} = R_{sam\_ap} + \frac{SDD(R_{sam\_ap} + R_{src\_ap})}{SSD} \quad (4.1)$$

In some instrument configurations, such as for long sample-to-detector distances with a long wavelength, the shape of the direct beam is not circular. Instead, it expands downward along the vertical direction due to gravity. In such cases, a manual mask, described below, can improve the data quality at low  $q$ . This mask is optional.

4. **Detector Edges.** The LPSD detectors do not respond uniformly along their length, particularly at the ends of the LPSD. The sensitivity of the LPSDs drops off dramatically near either end of the detector, making the data unreliable. Therefore, these regions are masked out. A mask for these detector elements is created by hand-drawing it, as described below. However, the ability to specify a number of pixels from the top and bottom edges to automatically mask exists. This mask is optional.
5. **Manually drawn mask.** Occasionally, a mask for a specific experiment or a specific sample may be required. For example, samples or a sample environment can create

spurious peaks or flares that should be removed prior to reducing the data. Therefore, the ability to draw arbitrarily shaped masks is required. These manually drawn masks are saved so that it can be used in a data reduction script. This is an optional mask.

- 6. Detector Bank Mask.** Presntly, the LPSD arrays in the detectors used for the ORNL SANS are arranged in a staggered pattern of front and back banks of LPSD that are spaced in such a way to eliminate gaps in the coverage. The arrangement creates a shadowing effect of the back bank by the front bank of LPSDs that requires non-trivial corrections during data reduction. These distortions depend very strongly on the *SDD*, and can become strong enough at short *SDD* to necessitate discarding the entire back bank from the data reduction. In particular, for scattering angles greater than  $25^\circ$ , a mask applied to the back bank of detectors ensures the best results. At present, it is possible to specify that either the front bank or back bank of the detector be masked during the data reduction. The ability to automatically mask the back bank of detectors when  $SDD < 1.5$  m and the detector offset is greater than 20 cm also exists, which will ensure that unreliable data is not included in the data reduction. This mask is optional.
- 7. Mask from sensitivity threshold.** The sensitivity file, which is prepared as described in Appendix 3, contains mask information that automatically excludes detector pixels that fall outside of an acceptable range of sensitivity about the mean value (see Appendix 3.2.3 for details). These masked pixels are extracted from the sensitivity file when it is read and appended to the mask. The use of this mask is not

optional if a sensitivity file is employed during the data reduction and the sensitivity file was prepared with thresholds as described in Appendix 3.

## 5. Correction for Sensitivity

Inherent differences exist in the neutron detection performance of the individual pixels, LPSDs and 8-packs of LPSDs in the detectors. The differences arise from the manufacture of the components, their age and can even be the result of localized damage due to over-exposure of the detector to a high flux of neutrons in a small area. In the case of the LPSD tube arrays employed at ORNL, the arrangement of the tubes and the electronic grouping into 8-packs are the largest source of deviations in detector response. The differences in pixel response can be removed from SANS data by performing what is known as the sensitivity correction, and it is applied to all files that are utilized in the data reduction. The preparation of the sensitivity file used for this correction is described in Appendix 3.

Application of the sensitivity correction is a straightforward division of the data that has already been treated according to Section 3. The data are divided by the sensitivity  $S(x, y)$  and associated uncertainty  $\delta S(x, y)$  on a per-pixel basis. Given a data set,  $F(x, y, \lambda)$  having an associated uncertainty  $\delta F(x, y, \lambda)$ , the sensitivity corrected data and associated uncertainties, which are assumed to be uncorrelated between the data sets, are given by Equations 5.1 and 5.2.

$$F'(x, y, \lambda) = \frac{F(x, y, \lambda)}{S(x, y)} \quad (5.1)$$

$$\delta F'(x, y, \lambda) = \frac{F(x, y, \lambda)}{S(x, y)} \left[ \left( \frac{\delta F(x, y, \lambda)}{F(x, y, \lambda)} \right)^2 + \left( \frac{\delta S(x, y)}{S(x, y)} \right)^2 \right]^{1/2} \quad (5.2)$$

It is possible to disable the sensitivity correction during a reduction, which is generally done for diagnostic purposes.



## 6. Normalize the Data

In order to use the various data sets properly during data reduction, the data must be normalized to established equivalent beam exposure. At HFIR and SNS, three normalization methods can be used.

1. By measurement duration
2. By beam monitor value or spectrum
3. By proton charge on target and measured flux spectrum

### 6.1 Normalization by Time

When normalizing by time, the impact of the duration of the measurement on the number of counts per pixel is removed. Normalization by time is not the most common data normalization method. The reason lies in the fact that the beams at SNS and HFIR do not possess constant strength, which makes the second and third methods listed above preferable when the signal strength scales with the beam, i.e. any measurement other than a true dark current measurement, which is independent of the neutron source. Normalization by time is available for normalizing any data set for both diagnostic purposes and in the event of a hardware failure that prevents using the other normalization methods. A data set  $I_{sam}(x, y)$  that was collected for a time  $t_{sam}$  is normalized according to Equation 6.1.

$$I_{sam,norm}(x, y) = \frac{I_{sam}(x, y)}{t_{sam}} \quad (6.1)$$

Normalization of a dark current measurement performed with the shutter closed, which is not to be confused with a blocked beam measurement, is a special case of normalizing data by time. For the HFIR SANS instruments, if the dark current data set  $I_{dc}(x, y)$  was measured for

a time  $t_{dc}$ , it must be normalized relative to the sample measurement time,  $t_{sam}$ , as is shown in Equation 6.2.

$$I_{dc,norm}(x, y) = \frac{t_{sam}}{t_{dc}} I_{dc}(x, y) \quad (6.2)$$

In the case of the EQ-SANS instrument, scaling the dark current by time is more involved than is shown Equation 6.2. The dark current is still independent of the source, and therefore the time-of-flight structure, so the total dark current counts collected for a time is evenly distributed over the entire TOF frame,  $t_{frame}$ . The data reduction does not use the whole frame. Instead, lower and upper TOF cutoffs are applied to the data, which are denoted  $t_{cut}^{low}$  and  $t_{cut}^{high}$ , respectively. Further, the data are binned into wavelength bins centered at  $\lambda_j$  with width  $\Delta\lambda_j$ . Let the minimum and maximum wavelengths be  $\lambda_{min}$  and  $\lambda_{max}$ , respectively. Consider a dark current data set, which is always integrated over the wavelength bins to remove the TOF dimension,  $I_{dc}(x, y)$ , measured for a time  $t_{dc}$ . Then, it is normalized relative to a data set  $I_{data}(x, y, \lambda)$  measured for a time  $t_{sam}$  according to Equation 6.3 for dark current subtraction.

$$I_{dc,norm}(x, y, \lambda_j) = \frac{t_{sam}}{t_{dc}} \left[ \frac{t_{frame} - t_{cut}^{low} - t_{cut}^{high}}{t_{frame}} \right] \left[ \frac{\Delta\lambda_j}{\lambda_{max} - \lambda_{min}} \right] I_{dc}(x, y) \quad (6.3)$$

For Eq 6.1 – Eq. 6.3, error propagates as

$$\delta I_{norm}(x, y, \lambda_j) = constant \times \delta I_{sam/dc}(x, y) \quad (6.4)$$

Here, the *constant* is the prefactor used for performing the normalization by time.

## 6.2 Normalizing by Beam Monitor

The beam monitor is a low efficiency detector in the beam just after the neutron source for the instrument, which is the velocity selector for the GP-SANS or Bio-SANS and the chopper furthest from the moderator in the case of the EQ-SANS. The monitors on the GP-SANS and Bio-SANS provide a single value,  $\Phi$ . If there is a data set from a sample,  $I_{sam}(x, y)$  with a monitor value  $\Phi_{sam}$ , then the normalization takes the form shown in Equation 6.4. The factor of  $10^8$  is a reference number of monitor counts selected by the instrument staff.

$$I_{sam,norm}(x, y) = \frac{10^8}{\Phi_{sam}} I_{sam}(x, y) \quad (6.4)$$

The monitor provides a direct measurement of the strength of the source of neutrons and is, in principle, the best way to normalize data collected with the beam (i.e. anything other than a dark current measurement). However, when EQ-SANS is operating in the 30 Hz “frame-skipping” mode where neutrons from two consecutive pulses are collected at the primary detector (see reference [7] for details), the slower neutrons from the earliest pulse cannot be differentiated from the faster neutrons produced in the subsequent pulse at the monitor position because they do not fully separate out in time until they reach the primary detector. Thus, monitor normalization cannot be used for EQ-SANS data sets collected in this manner. However, when a data set  $I_{sam}(x, y, \lambda)$  can be normalized by its monitor,  $\Phi_{sam}(\lambda)$ , the normalization takes the form shown in Equation 6.5. A pre-measured flux-to-monitor ratio spectrum,  $FM(\lambda)$ , must be provided.

$$I_{sam,norm}(x, y, \lambda) = \frac{I_{sam}(x, y, \lambda)}{\Phi_{sam}(\lambda) FM(\lambda)} \quad (6.5)$$

### 6.3 Normalizing by Proton Charge on Target and Measured Flux Spectrum

The third normalization option, to proton charge on target and a measured flux spectrum, is only suitable for the EQ-SANS instrument, which was adopted because of the 30 Hz “frame-skipping” mode of data acquisition that precludes the use of the beam monitor. Measurements at the SNS are typically performed in units of proton charge, which enables consistent exposures regardless of beam power or interruptions of beam production, rather than in units of time. The proton charge on target during the measurement of a sample,  $P_{sam}$ , expressed in units of  $\mu A \cdot h$ , can be used to normalize a measured data set  $I_{sam}(x, y, \lambda)$  if it is used in conjunction with a pre-measured instrument flux spectrum  $\phi(\lambda)$ , which is measured with the primary detector. The pre-measured flux spectrum is collected using the primary instrument detector, which has the benefit of automatically including the effect of the wavelength sensitivity of the main detector,  $E(\lambda)$ , on the data to be normalized. Equation 6.6 is used to perform data normalization in this manner.

$$I_{sam,norm}(x, y, \lambda) = \frac{I_{sam}(x, y, \lambda)}{P_{sam}\phi(\lambda)} \quad (6.6)$$

## 7. Correct for Transmission

As a neutron passes through a material, it has a chance to be absorbed or scattered from its original path, much like a photon, and SANS data must be corrected for absorption. The neutron absorption can be expressed using the Beer-Lambert Law if absorption edges are neglected. Throughout this chapter, absorption edges are not considered. If a beam of neutrons is incident on a sample with a wavelength-dependent intensity  $I_{inc}(\lambda)$ , the transmitted intensity,  $I_{trans}(\lambda)$ , is given by Equation 7.1.

$$I_{trans}(\lambda) = I_{inc}(\lambda)e^{-\mu(\lambda)d} \quad (7.1)$$

The parameter  $d$  is the distance that the neutron beam travels through the material, and  $\mu(\lambda)$  is the wavelength-dependent linear attenuation coefficient. For neutrons,  $\mu(\lambda)$  is the sum of the cross-sections of the atoms in the material, as is shown in Equation 7.2.

$$\mu(\lambda) = \sum_i \sigma_i(\lambda) = \sum_i (\sigma_{coh,i} + \sigma_{incoh,i} + \sigma_{abs,i}(\lambda)) \quad (7.2)$$

In Equation 7.2, the coherent neutron cross-section,  $\sigma_{coh}$ , is responsible for the signal of interest in a SANS measurement, while the incoherent neutron cross-section,  $\sigma_{incoh}$ , is generally the source of a flat background for a given wavelength during a SANS measurement. Hydrogen is the most common material encountered in SANS experiments that has a large  $\sigma_{incoh}$ . The absorption cross-section  $\sigma_{abs}(\lambda)$ , is due to the nuclear capture process, and is the most wavelength-dependent of the three cross-sections. Values of the cross-sections for all nuclei and their isotopes can be found on the NIST Center for Neutron Research website [<https://www.nist.gov/ncnr/planning-your-experiment>]. The transmission of the material,  $T(\lambda)$ , is given by Equation 7.3.

$$T(\lambda) = \frac{I_{trans}(\lambda)}{I_{inc}(\lambda)} = e^{-\mu(\lambda)d} \quad (7.3)$$

The path length of the neutron through the sample,  $d$  in Equation 7.3, depends on the sample geometry and the route that the neutron takes through it. The impact of the scattering angle  $2\theta$  on the transmission  $T(\lambda, 2\theta)$  through a planar sample in which the incident beam and the normal to the plane of the sample are colinear was derived by J B. Hayter and W. A. Hamilton (W. A. Hamilton, personal communication) and is shown in Equation 7.4.

$$T(\lambda, 2\theta) = T(\lambda, 2\theta = 0)^{(1+\sec(2\theta))/2} \quad (7.4)$$

Here,  $T(\lambda, 0)$  is the zero-angle transmission, which is measured during an experiment according to the method described below. Equation 7.4 is only applicable to a planar sample geometry. It is possible that analytical forms can be derived when the sample geometry is simple. However, in the case of more complex sample geometries, such as are encountered in shear cells and rheometers, or when using irregular, solid samples, Monte Carlo computational approaches are suitable for determining the path-length through a sample as a function of the scattering vector  $\vec{Q}$ , which makes it possible to calculate  $T(\lambda, x, y)$ . Mantid methods exist for performing these calculations, but a description of use of the methods is beyond the scope of this document.

## 7.1 Measuring the Zero-Angle Transmission

The zero-angle transmission is measured using the “direct beam” method on all three SANS at ORNL, although there are variants of the exact method depending on the instrument and instrument configuration being used. Most of the measurements involve attenuating the beam, moving the beam stop away from the position of the incident beam, and collecting a

data set with the attenuated beam incident on the detector for a few minutes. In cases where the flux is sufficiently low, the transmission may be collected simultaneously with the scattering data by simply not using a beam stop. The Bio-SANS also has a semi-transparent beam stop that makes it possible to collect the transmission data at the same time as the scattering data when using the commonly employed wavelength of 6 Å, but not for the commonly used long wavelength settings used (12 Å and higher) due to the strong wavelength dependence of the absorption of the beamstop.

Regardless of the details of the measurement, there are guidelines for ensuring that the measured transmissions are reliable and possess sufficient counting statistics. The primary concern when measuring the direct beam transmission is to ensure that the localized count rate is lower than the level that causes the LPSDs to saturate (on the order of a few hundred counts per pixel per second). On the HFIR SANS, the best practice is to measure the transmission for a particular wavelength using the configuration with the longest sample-to-detector distance used because the flux in that configuration will be the lowest. On the GP-SANS, the beam is attenuated by a factor of ~10k or ~2k when using the wavelength of 4.75Å. When  $\lambda \geq 12$  Å, the beam is attenuated by a factor of ~30 for the 40mm diameter source aperture, and no attenuation is used if the source aperture is 20mm. The transmission is collected simultaneously with the scattering when  $\lambda = 18$  Å. Similar guidelines are applicable to Bio-SANS, but no additional attenuation is needed when the semitransparent beam stop, which has an attenuation factor of ~2500, is used for a simultaneous scattering and transmission measurement that is saved as a single data file. On the EQ-SANS, the direct beam is attenuated by utilizing either a source aperture with attenuation capability or a source aperture with

smaller size, depending on the wavelength band being used and the sample-to-detector distance.

Independent of the instrument being used, the measurements must run sufficiently long to have at least 10k counts in the direct beam spot on GP-SANS and Bio-SANS, which equates to counting statistics of  $\sim 1\%$ . For EQ-SANS, the wavelength-dependent transmission necessitates that each wavelength bin has the same counting statistics, at least to the extent that the wavelength-dependent flux allows. For wavelength or wavelength band settings near the peak of the flux spectrum, attenuated beam transmission measurements require  $\sim 2$  minutes. At longer wavelengths, sufficient counting statistics may require up to 5 minutes. Such guidelines are not necessary when using the semi-transparent beam stop on Bio-SANS, or a long wavelength setting without any beam stop for making simultaneous scattering and transmission measurements on the HFIR SANS. The counting time is governed by the desired statistics in the scattering pattern.

The transmission calculation requires two measurements to determine the absorption of the sample – a transmission measurement for the sample and a reference measurement made without the sample. Most often, the reference measurement used is of the empty beam in which the sample environment is in place, but there is nothing in the sample position. Alternatively, an empty cell mounted into the sample environment can be used as the reference for the transmission measurement. Users should consult with an instrument scientist to determine the best transmission reference for their experiment.



## 7.2 Calculate the Transmission

With the sample and reference transmission data sets collected, the zero-angle transmission can be calculated. After reading the data and preparing it as described in Section 3, above, the bad pixel mask is applied to the data. Then, the detector sensitivity is applied to the data. Finally, the data is normalized by the appropriate method and the dark current is subtracted. Denote the processed sample transmission data and uncertainties as  $I'_{sam}(x, y, \lambda)$  and  $\delta I'_{sam}(x, y, \lambda)$ , respectively. Similarly, denote the data for the reference as  $I'_{ref}(x, y, \lambda)$  and  $\delta I'_{ref}(x, y, \lambda)$ . Then, the zero-angle transmission  $T(\lambda, 2\theta = 0)$  and its uncertainty  $\delta T(\lambda, 2\theta = 0)$  are given by Equations 7.5 and 7.6.

$$T(\lambda, 2\theta = 0) = \frac{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{sam}(x,y,\lambda)}{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{ref}(x,y,\lambda)} \quad (7.5)$$

$$\delta T(\lambda, 2\theta = 0) = \frac{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{sam}(x,y,\lambda)}{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{ref}(x,y,\lambda)} \left[ \left( \frac{\left[ \sum_{x,y;\|\vec{r}_{x,y}\|<R} (\delta I'_{sam}(x,y,\lambda))^2 \right]^{1/2}}{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{sam}(x,y,\lambda)} \right)^2 + \left( \frac{\left[ \sum_{x,y;\|\vec{r}_{x,y}\|<R} (\delta I'_{ref}(x,y,\lambda))^2 \right]^{1/2}}{\sum_{x,y;\|\vec{r}_{x,y}\|<R} I'_{ref}(x,y,\lambda)} \right)^2 \right]^{1/2} \quad (7.6)$$

Here,  $R$  is the radius around the direct beam to integrate over and  $\|\vec{r}_{x,y}\|$  is the distance of a pixel  $(x,y)$  from the beam center  $\langle X, Y \rangle$  that has been determined as described in Section 3.1. At present, the default radius of integration for the transmission calculation in Mantid is 5 pixels. An alternative to consider using in the future is the beam radius,  $R_{beam}$ , from Equation 4.1.

To reduce the amount of noise introduced into the data reduction, the measured  $T(\lambda, 2\theta = 0)$  and  $\delta T(\lambda, 2\theta = 0)$  are fit using a linear function to provide  $T_{fit}(\lambda, 2\theta = 0)$  and its

uncertainty  $\delta T_{fit}(\lambda, 2\theta = 0)$ , which are the actual values used when correcting the measured data. A linear function is reasonable because of the relatively narrow wavelength band used by EQ-SANS. If the instrument is used in frame-skipping mode, which provides two wavelength bands separated by a  $\sim 3 \text{ \AA}$  gap, the bands are fit independently because a single linear function is not suitable for simultaneously fitting  $T(\lambda, 2\theta = 0)$  of both wavelength bands.

### 7.3 Correcting Scattering Data for Sample Transmission

Regardless of how the zero-angle transmission  $T(\lambda)$  is measured or the form of the path-length dependent transmission  $T(\lambda, x, y)$ , which has an uncertainty  $\delta T(\lambda, x, y)$ , a measured data set  $I_{sam}(x, y, \lambda)$  with uncertainty  $\delta I_{sam}(x, y, \lambda)$  is corrected for the sample transmission according to Equation 7.7. The uncertainty in the result is propagated according to Equation 7.8.

$$I'_{sam}(x, y, \lambda) = \frac{I_{sam}(x, y, \lambda)}{T(\lambda, x, y)} \quad (7.7)$$

$$\delta I'_{sam}(x, y, \lambda) = \frac{I_{sam}(x, y, \lambda)}{T(\lambda, x, y)} \left[ \left( \frac{\delta I_{sam}(x, y, \lambda)}{I_{sam}(x, y, \lambda)} \right)^2 + \left( \frac{\delta T(\lambda, x, y)}{T(\lambda, x, y)} \right)^2 \right]^{1/2} \quad (7.8)$$

It is possible to specify the transmission of a sample, which would generally only be done for diagnostic purposes. When the value is specified, the uncertainty is set to zero. Similarly, it is possible to disable the path-length corrections to the transmission. In such cases, the zero angle transmission  $T(\lambda)$  is applied to all pixels.

## 8. Subtract Background

An important operation performed is background subtraction, which is particularly when working with weakly scattering samples such as dilute particles in solution. Sources of background signal that are intrinsic to the measurement include the air along the beam path, imperfections in the apertures used for collimation, the sample environment and the sample cell. Similarly, the nuclear scattering signal from a sample can obscure magnetic scattering. Extrinsic sources of background signal include flaws in the shielding between the upstream instrument components and the detector, stray neutrons from neighboring instruments, and cosmic radiation. Ultimately, the signal of interest must be extracted from the measured data and the background signal seen during the measurement. All background subtractions are optional.

Several options for performing the background subtraction are provided that make it possible to suit the process to the needs of the user. The “right” approach for background subtraction is a matter of the experiment being performed and opinion, but best practices are identified below. For example, an empty cell measurement, like the sample measurement, contains background from the dark current, so one must be careful in the application of background subtraction so that a source of background is not removed multiple times. In the following, all data are assumed to have been appropriately corrected for the instrument geometry, detector sensitivity, transmission and have been normalized according to the procedures described above.

## 8.1 Correcting for sources of extrinsic background signal

The dark current or blocked beam represents background that does not inherently result from the passage of the neutron beam through the instrument. The choice of whether to use a dark current or blocked beam strongly depends on the instrument being used. In the case of the EQ-SANS instrument at SNS, the upstream shielding and the shielding between neighboring instruments is sufficient enough that there is no discernable difference between a dark current and a blocked beam. Therefore, a dark current measurement lasting a few hours and performed during machine physics is used instead of a blocked beam measurement. In contrast, the SANS instruments at HFIR, much like those at other reactor facilities, are not as well shielded, which results in a background that can vary with the instrument configuration in use and the measurements being performed on neighboring instruments. GP-SANS and Bio-SANS employ either a dark current measurement or a blocked beam, but the latter is used exclusively on the GP-SANS instrument. Subtracting the dark current or blocked beam is optional.

Best practices dictate that all data, which are denoted here as  $I(x, y, \lambda)$  with uncertainty  $\delta I(x, y, \lambda)$ , be corrected for the extrinsic background signal,  $I_{ex}(x, y, \lambda)$ , which has uncertainty  $\delta I_{ex}(x, y, \lambda)$ , regardless of whether it is a dark current or blocked beam measurement. Subtraction is performed with the 2D data. Note that for EQ-SANS, the dark current is normalized according to Equation 6.2, above. The subtraction and uncertainty propagations for dark current and blocked beam are given by Equations 8.1 and 8.2, respectively.

$$I'(x, y, \lambda) = I(x, y, \lambda) - I_{ex}(x, y, \lambda) \quad (8.1)$$

$$\delta I'(x, y, \lambda) = \left[ (\delta I(x, y, \lambda))^2 + (\delta I_{ex}(x, y, \lambda))^2 \right]^{1/2} \quad (8.2)$$

## 8.2 Correcting for sources of intrinsic background signal

The correction for the intrinsic background signal,  $I_{in}(x, y, \lambda)$  with uncertainty  $\delta I_{in}(x, y, \lambda)$ , accounts for undesirable signal that results from a beam traveling along the instrument. The intrinsic background is perhaps the most readily understood in a conceptual sense, but the exact measurement to perform to account for it can vary depending on the experiment. A measurement of the empty beam, lacking anything between the end of the collimation and the window on the detector tank, would account for air scattering and imperfections in the apertures and windows. A measurement of the empty cell or sample environment, which has everything in place except for the actual sample, would account for everything that an empty beam measurement does, as well as background signal coming from the sample environment equipment and the sample cell. When working with material dissolved in a solvent or dispersed in some other matrix, a measurement of the blank solvent or matrix would arguably be the most “complete” background to use for data reduction. An instrument scientist can help determine what the is the most appropriate background for data reduction. Correcting for the intrinsic background is performed according to Equations 8.3 and 8.4 after all signals have been properly treated and normalized.

$$I'(x, y, \lambda) = I(x, y, \lambda) - I_{in}(x, y, \lambda) \quad (8.3)$$

$$\delta I'(x, y, \lambda) = \left[ (\delta I(x, y, \lambda))^2 + (\delta I_{in}(x, y, \lambda))^2 \right]^{1/2} \quad (8.4)$$

For the HFIR SANS, it is also possible to correct data for a single value  $I_{in}$  with uncertainty  $\delta I_{in}$  that is determined from the average and standard deviation of  $I_{in}(x, y)$ .

### 8.3 Additional comments on background subtraction

The best resource for determining the best approach for correcting the background during a SANS experiment is the instrument staff. Background subtraction is often a nested process in which the sample signal and its background are first corrected using the extrinsic background (the dark current or blocked beam) prior to subtraction. The software for correcting for the background signals is modular to enable customization of background subtraction for an experiment. An example of a nested background subtraction for both intrinsic and extrinsic sources of background is shown in Equation 8.5.

$$I'(x, y, \lambda) = [I(x, y, \lambda) - I_{ex}(x, y, \lambda)] - [I_{in}(x, y, \lambda) - I_{ex}(x, y, \lambda)] \quad (8.5)$$

In such cases, the nested background corrections should be applied consistently. For example, if the sample signal is going to be corrected for the dark current rather than the blocked beam, then the background must be treated the same way. Failing to consistently apply nested background subtractions will result in an improper background subtraction.

Many researchers prefer to employ an “empty cell” measurement for the intrinsic background. This approach is often taken when the intrinsic background is not expected to display structure, such as when it is a solvent where the signal is dominated by incoherent scattering from hydrogen or deuterium. Others prefer to use a solvent or empty matrix measurement, i.e. if the sample is nanoparticles dispersed in solid polymer, then the empty matrix is the solid polymer for the intrinsic background data during data reduction. The latter form of background can present challenges that arise from concentration of the material that is dissolved in the solvent or dispersed in the matrix and the differences in hydrogen content between the material of interest and the solvent or matrix. The displacement of the solvent or

matrix by the material of interest results in a mis-match of the amount of solvent or matrix that must be accounted for. While scaling the background by a factor could account for the differences, the sample and intrinsic backgrounds must be corrected for the “empty cell” before doing so to ensure that the sample environment and cell are not inappropriately scaled by the volume fraction of the material of interest. Weakly scattering systems are particularly challenging. In such cases, background subtraction may be performed after data reduction, binning into 1D Q-space (section 11, below), and merging all data collected in various configurations into a single profile.

## 9. Polarization

As noted in Section 2, the use of half or full polarization in a SANS experiment requires parsing the data into states corresponding to the states of the flipper, and the analyzer that is located between the sample and the detector. The wavelength-dependent polarization of the incident beam that is defined by the polarizer is denoted  $P(\lambda)$ . Throughout this document, it is assumed that the polarizer produces the spin-up state, which would be true for a reflection polarizer. If a transmission polarizer, which produces the spin-down state, is employed, then the states derived from the measured flipper and analyzer conditions, described below, are reversed.

$P(\lambda)$  is related to the instantaneous populations of spin-up and spin-down neutrons in the beam, denoted  $N^\uparrow(\lambda)$  and  $N^\downarrow(\lambda)$ , respectively, through Equation 9.1.

$$P(\lambda) = \frac{N^\uparrow(\lambda) - N^\downarrow(\lambda)}{N^\uparrow(\lambda) + N^\downarrow(\lambda)} \quad (9.1)$$

Similarly, the wavelength-dependent polarization ratio  $R(\lambda)$ , often called the flipping ratio, is defined by Equation 9.2.

$$R(\lambda) = \frac{N^\uparrow(\lambda)}{N^\downarrow(\lambda)} \quad (9.2)$$

Then,  $R(\lambda)$  can be expressed in terms of  $P(\lambda)$  by Equation 9.3.

$$R(\lambda) = \frac{1+P(\lambda)}{1-P(\lambda)} \quad (9.3)$$

As is noted in Section 2.4, the event stream of data is parsed into workspaces representing the states of the polarizer and analyzer. Here, the notation used in the rest of this section is simplified by abbreviating any polarizer or analyzer state SANS data  $S(x, y, \lambda)$  as  $S(\lambda)$ . When half polarization is employed, the “off” and “on” states of the flipper are measured, and



the data are denoted  $S^0(\lambda)$  and  $S^1(\lambda)$ , respectively. The quantities of interest in a half polarization measurement are the spin-up and spin-down state SANS data sets, denoted  $S^\uparrow(\lambda)$  and  $S^\downarrow(\lambda)$ , respectively, which are related to the measured states. Using full polarization introduces two additional conditions for when the analyzer is set to either  $0^\circ$  or  $180^\circ$ . As a result, four SANS data sets are measured, which are denoted  $S^{00}(\lambda)$ ,  $S^{0\pi}(\lambda)$ ,  $S^{10}(\lambda)$ , and  $S^{1\pi}(\lambda)$ . By analogy with half polarization, there are four spin state SANS data sets of interest in a full polarization experiment, namely  $S^{\uparrow\uparrow}(\lambda)$ ,  $S^{\uparrow\downarrow}(\lambda)$ ,  $S^{\downarrow\uparrow}(\lambda)$  and  $S^{\downarrow\downarrow}(\lambda)$ , that are related to the four measured polarization states. Both cases are described below.

### 9.1 Half polarization – polarized beam without polarization analysis

Half polarization experiments use a spin flipper with a wavelength-dependent efficiency  $e(\lambda)$  between the polarizer and the sample to rotate the spins from up to down, or vice versa depending on the polarizer in use. Consider the case that the flipper is off. The fractions of the beam that are spin-up and spin down, denoted  $f_{up}$  and  $f_{down}$ , respectively, are given by Equations 9.4 and 9.5, respectively, and sample the spin-up scattering  $S^\uparrow(\lambda)$  and spin-down scattering  $S^\downarrow(\lambda)$ .

$$f_{up} = \frac{N^\uparrow}{N^\uparrow + N^\downarrow} = \frac{R(\lambda)}{1 + R(\lambda)} \quad (9.4)$$

$$f_{down} = \frac{N^\downarrow}{N^\uparrow + N^\downarrow} = \frac{1}{1 + R(\lambda)} \quad (9.5)$$

Then, the total scattering measured by the detector for the off state is given by Equation 9.6, and its uncertainty is given by Equation 9.7.

$$S^0(\lambda) = \frac{R(\lambda)}{1 + R(\lambda)} S^\uparrow(\lambda) + \frac{1}{1 + R(\lambda)} S^\downarrow(\lambda) \quad (9.6)$$

$$\left(\delta S^0(\lambda)\right)^2 = \left(\frac{R(\lambda)}{1+R(\lambda)}\right)^2 \left(\delta S^\uparrow(\lambda)\right)^2 + \left(\frac{1}{1+R(\lambda)}\right)^2 \left(\delta S^\downarrow(\lambda)\right)^2 \quad (9.7)$$

When the flipper is on, it rotates the spins in the beam produced by the polarizer by  $\pi$  with an efficiency  $e(\lambda)$ . The flipped spins are sensitive to the spin-down scattering state. The total scattering measured when the flipper is on, and its uncertainty are given by Equations 9.8 and 9.9, respectively.

$$S^1(\lambda) = \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right] S^\uparrow(\lambda) + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - e(\lambda)) \frac{1}{1+R(\lambda)} \right] S^\downarrow(\lambda) \quad (9.8)$$

$$\begin{aligned} (\delta S^1(\lambda))^2 &= \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right]^2 (\delta S^\uparrow(\lambda))^2 + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - \right. \\ &\left. e(\lambda)) \frac{1}{1+R(\lambda)} \right]^2 (\delta S^\downarrow(\lambda))^2 \end{aligned} \quad (9.9)$$

As can be seen in Equations 9.6 through 9.9, the measurements produce data that contain mixtures of both states. The spin-up and spin-down SANS data, along with the associated uncertainties, are obtained by performing two  $2 \times 2$  matrix inversions. The matrices are the coefficients in Equations 9.6 and 9.8, which is denoted  $M_1$ , and the coefficients in Equations 9.7 and 9.9, which is denoted  $M_2$ . The matrix equations are shown in Equations 9.10 and 9.11 and are functions of wavelength. These equations no longer abbreviate the x and y dimensions of the SANS data.

$$\begin{pmatrix} S^\uparrow(x, y, \lambda) \\ S^\downarrow(x, y, \lambda) \end{pmatrix} = M_1^{-1} \begin{pmatrix} S^0(x, y, \lambda) \\ S^1(x, y, \lambda) \end{pmatrix} \quad (9.10)$$

$$\begin{pmatrix} (\delta S^\uparrow(x, y, \lambda))^2 \\ (\delta S^\downarrow(x, y, \lambda))^2 \end{pmatrix} = M_2^{-1} \begin{pmatrix} (\delta S^0(x, y, \lambda))^2 \\ (\delta S^1(x, y, \lambda))^2 \end{pmatrix} \quad (9.11)$$

Solving for  $S^\uparrow(x, y, \lambda)$ ,  $S^\downarrow(x, y, \lambda)$  yields the following two equations.

$$S^\uparrow(x, y, \lambda) = S^0(x, y, \lambda) + \frac{1}{e(\lambda)[R(\lambda)-1]} [S^0(x, y, \lambda) - S^1(x, y, \lambda)] \quad (9.12)$$

$$S^\downarrow(x, y, \lambda) = S^0(x, y, \lambda) + \frac{1}{e(\lambda)[1-(1/R(\lambda))]} [S^0(x, y, \lambda) - S^1(x, y, \lambda)] \quad (9.13)$$

Similarly,  $\delta S^\uparrow(x, y, \lambda)$  and  $\delta S^\downarrow(x, y, \lambda)$  can be solved for if we let the uncertainty in the efficiency of the polarizer be  $\delta e(\lambda)$  and  $\delta R(\lambda) \sim \left( 2/(1 - P(\lambda))^2 \right) \delta P(\lambda)$ , where  $\delta P(\lambda)$  is the

uncertainty in the efficiency of the polarizer. The resulting expressions are shown in Equations 9.14 and 9.15.

$$\delta S^\uparrow(x, y, \lambda) = \left[ \left( 1 + \frac{1}{e(\lambda)(R(\lambda)-1)} \right)^2 (\delta S^0(x, y, \lambda))^2 + \left( \frac{1}{e(\lambda)(R(\lambda)-1)} \right)^2 (\delta S^1(x, y, \lambda))^2 + \left( \frac{S^0(x, y, \lambda) - S^1(x, y, \lambda)}{e(\lambda)(R(\lambda)-1)} \right)^2 \left( \left( \frac{\delta e(\lambda)}{e(\lambda)} \right)^2 + \left( \frac{\delta R(\lambda)}{R(\lambda)-1} \right)^2 \right) \right]^{1/2} \quad (9.14)$$

$$\delta S^\downarrow(x, y, \lambda) = \left[ \left( 1 - \frac{1}{e(\lambda)(1-(1/R(\lambda)))} \right)^2 (\delta S^0(x, y, \lambda))^2 + \left( \frac{1}{e(\lambda)(1-(1/R(\lambda)))} \right)^2 (\delta S^1(x, y, \lambda))^2 + \left( \frac{S^0(x, y, \lambda) - S^1(x, y, \lambda)}{e(\lambda)(1-(1/R(\lambda)))} \right)^2 \left( \left( \frac{\delta e(\lambda)}{e(\lambda)} \right)^2 + \frac{1}{(R(\lambda))^4} \left( \frac{\delta R(\lambda)}{1-(1/R(\lambda))} \right)^2 \right) \right]^{1/2} \quad (9.15)$$

## 9.2 Full polarization – polarized beam with polarization analysis

When full polarization is employed, a  $^3\text{He}$  spin filter is inserted between the sample and the detector. The filter is also a spin flipper that rotates the spin by  $0^\circ$  or  $180^\circ$  and the polarizations of the two states of the  $^3\text{He}$  spin filter are denoted by  $A^0(\lambda)$  and  $A^\pi(\lambda)$ . The  $0^\circ$  state is assumed to preferentially pass the spin-up state, but the software must allow for the reverse to be true. The four possible flipper states are related to the four spin states through Equations 9.16 through 9.19.

$$S^{00}(\lambda) = \frac{R(\lambda)}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} S^{\uparrow\uparrow}(\lambda) + \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} S^{\uparrow\downarrow}(\lambda) + \frac{1}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} S^{\downarrow\uparrow}(\lambda) + \frac{1}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} S^{\downarrow\downarrow}(\lambda) \quad (9.16)$$

$$S^{10}(\lambda) = \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right] \frac{A^0(\lambda)}{1+A^0(\lambda)} S^{\uparrow\uparrow}(\lambda) + \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right] \frac{1}{1+A^0(\lambda)} S^{\uparrow\downarrow}(\lambda) + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - e(\lambda)) \frac{1}{1+R(\lambda)} \right] \frac{A^0(\lambda)}{1+A^0(\lambda)} S^{\downarrow\uparrow}(\lambda) + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - e(\lambda)) \frac{1}{1+R(\lambda)} \right] \frac{1}{1+A^0(\lambda)} S^{\downarrow\downarrow}(\lambda) \quad (9.17)$$

$$S^{0\pi}(\lambda) = \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} S^{\uparrow\uparrow}(\lambda) + \frac{R(\lambda)}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} S^{\uparrow\downarrow}(\lambda) + \frac{1}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} S^{\downarrow\uparrow}(\lambda) + \frac{1}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} S^{\downarrow\downarrow}(\lambda) \quad (9.18)$$

$$S^{1\pi}(\lambda) = \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right] \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} S^{\uparrow\uparrow}(\lambda) + \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1 - e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \right] \frac{1}{1+A^\pi(\lambda)} S^{\uparrow\downarrow}(\lambda) + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - e(\lambda)) \frac{1}{1+R(\lambda)} \right] \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} S^{\downarrow\uparrow}(\lambda) + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1 - e(\lambda)) \frac{1}{1+R(\lambda)} \right] \frac{1}{1+A^\pi(\lambda)} S^{\downarrow\downarrow}(\lambda) \quad (9.19)$$

The associated uncertainties are provided in Equations 9.20 through 9.23.

$$\begin{aligned} (\delta S^{00}(\lambda))^2 &= \left[ \frac{R(\lambda)}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} \right]^2 (\delta S^{\uparrow\uparrow}(\lambda))^2 + \left[ \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} \right]^2 (\delta S^{\uparrow\downarrow}(\lambda))^2 + \\ &\left[ \frac{1}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} \right]^2 (\delta S^{\downarrow\uparrow}(\lambda))^2 + \left[ \frac{1}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} \right]^2 (\delta S^{\downarrow\downarrow}(\lambda))^2 \end{aligned} \quad (9.20)$$

$$\begin{aligned}
(\delta S^{10}(\lambda))^2 &= \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1-e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} \right]^2 (\delta S^{\uparrow\uparrow}(\lambda))^2 + \left[ e(\lambda) \frac{1}{1+R(\lambda)} + \right. \\
&\quad \left. (1-e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} \right]^2 (\delta S^{\uparrow\downarrow}(\lambda))^2 + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1-e(\lambda)) \frac{1}{1+R(\lambda)} \right. \\
&\quad \left. e(\lambda) \frac{1}{1+R(\lambda)} \frac{A^0(\lambda)}{1+A^0(\lambda)} \right]^2 (\delta S^{\downarrow\uparrow}(\lambda))^2 + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1-e(\lambda)) \frac{1}{1+R(\lambda)} \frac{1}{1+A^0(\lambda)} \right]^2 (\delta S^{\downarrow\downarrow}(\lambda))^2
\end{aligned} \tag{9.21}$$

$$\begin{aligned}
(\delta S^{0\pi}(\lambda))^2 &= \left[ \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\uparrow\uparrow}(\lambda))^2 + \left[ \frac{R(\lambda)}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\uparrow\downarrow}(\lambda))^2 + \\
&\quad \left[ \frac{1}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\downarrow\uparrow}(\lambda))^2 + \left[ \frac{1}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\downarrow\downarrow}(\lambda))^2
\end{aligned} \tag{9.22}$$

$$\begin{aligned}
(\delta S^{1\pi}(\lambda))^2 &= \left[ e(\lambda) \frac{1}{1+R(\lambda)} + (1-e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\uparrow\uparrow}(\lambda))^2 + \left[ e(\lambda) \frac{1}{1+R(\lambda)} + \right. \\
&\quad \left. (1-e(\lambda)) \frac{R(\lambda)}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\uparrow\downarrow}(\lambda))^2 + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1-e(\lambda)) \frac{1}{1+R(\lambda)} \right. \\
&\quad \left. e(\lambda) \frac{1}{1+R(\lambda)} \frac{1}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\downarrow\uparrow}(\lambda))^2 + \left[ e(\lambda) \frac{R(\lambda)}{1+R(\lambda)} + (1-e(\lambda)) \frac{1}{1+R(\lambda)} \frac{A^\pi(\lambda)}{1+A^\pi(\lambda)} \right]^2 (\delta S^{\downarrow\downarrow}(\lambda))^2
\end{aligned} \tag{9.23}$$

The coefficients of the set of Equations 9.16 through 9.19 make up a matrix  $M_3$ , being a  $4 \times 4$  matrix for the full polarization case, that can be inverted to derive the four spin states of SANS data through Equation 9.24. Similarly, Equations 9.20 through 9.23 also define a  $4 \times 4$  matrix  $M_4$  that can be inverted to obtain the uncertainties in the spin state SANS data sets, as shown in Equation 9.25. The x and y dimensions are no longer abbreviated below.

$$\begin{pmatrix} S^{\uparrow\uparrow}(x, y, \lambda) \\ S^{\uparrow\downarrow}(x, y, \lambda) \\ S^{\downarrow\uparrow}(x, y, \lambda) \\ S^{\downarrow\downarrow}(x, y, \lambda) \end{pmatrix} = M_3^{-1} \begin{pmatrix} S^{00}(x, y, \lambda) \\ S^{10}(x, y, \lambda) \\ S^{0\pi}(x, y, \lambda) \\ S^{1\pi}(x, y, \lambda) \end{pmatrix} \quad (9.24)$$

$$\begin{pmatrix} (\delta S^{\uparrow\uparrow}(x, y, \lambda))^2 \\ (\delta S^{\uparrow\downarrow}(x, y, \lambda))^2 \\ (\delta S^{\downarrow\uparrow}(x, y, \lambda))^2 \\ (\delta S^{\downarrow\downarrow}(x, y, \lambda))^2 \end{pmatrix} = M_4^{-1} \begin{pmatrix} (\delta S^{00}(x, y, \lambda))^2 \\ (\delta S^{10}(x, y, \lambda))^2 \\ (\delta S^{0\pi}(x, y, \lambda))^2 \\ (\delta S^{1\pi}(x, y, \lambda))^2 \end{pmatrix} \quad (9.25)$$

### 9.3 Application of corrections to polarized beam data

In the kinematic limit where the scattering is weak compared to the intensity of the neutron beam, the transmission of the incident beam through the sample is not spin dependent. Thus, the transmission correction should be applied before the polarization correction. In most cases, sources of intrinsic and extrinsic background are spin independent. Subtraction of background for these cases should also be performed before application of the polarization corrections (otherwise a false spin dependent asymmetry may be introduced into the background). It is especially important to apply the polarization corrections to the wavelength dependence of the data before the data are mapped to reciprocal space.

## 10. Calculate Q-resolution

The physical layout of a SANS instrument and the source of neutrons impart an uncertainty to the value of  $Q$ . These effects are collectively called the instrument resolution function, and they have a visible effect on the data that is termed “smearing”. The most common way to account for the instrument resolution is during data analysis, when a model intensity profile is smeared by the instrument resolution function prior to being compared against to measured data. The instrument resolution function is configuration-dependent and is output to the user with the reduced data for use during data analysis.

Most frequently, SANS measurements are considered to take place at scattering angles that make the use of the small-angle approximation, where  $\sin \theta \sim \tan \theta \sim \theta$ . This assumption is not reasonable for the SANS instruments at HFIR. Therefore, the uncertainty in  $Q$  is given by Equation 10.1.

$$\sigma_Q^2(\lambda) = [\sigma_Q^2(\lambda)]_{geo} + [\sigma_Q^2(\lambda)]_{wave} = \frac{(4*\pi*cos\theta)^2}{\lambda^2} \sigma_\theta^2(\lambda) + Q^2 \frac{\sigma_\lambda^2(\lambda)}{\lambda^2} \quad (10.1)$$

In Equation 10.1,  $\sigma_\theta$  is the uncertainty in the scattering angle  $2\theta$  and  $\sigma_\lambda$  is the uncertainty in the wavelength. It is more convenient to work in 2D because of the effect of gravity on the resolution function, and the data reduction software is implemented to do so. The uncertainty in  $Q$  can be written in terms of  $\sigma_{Qx}$  and  $\sigma_{Qy}$ , the uncertainties in  $Q_x$  and  $Q_y$ , respectively, which is given by Equation 10.2.

$$\left(\sigma_Q(x, y, \lambda)\right)^2 = \sigma_{Qx}^2 + \sigma_{Qy}^2 \quad (10.2)$$



Both  $\sigma_{Qx}$  and  $\sigma_{Qy}$  consist of terms that arise from the instrument geometry and those that arise from the wavelength. A full derivation is presented in Appendix 4. For the GP-SANS and Bio-SANS,  $\sigma_{Qx}$  and  $\sigma_{Qy}$  are given by Equations 10.3 and 10.4.

$$\sigma_{Qx}^2(\lambda) = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \right] + \frac{Q_x^2}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \quad (10.3)$$

$$\sigma_{Qy}^2(\lambda) = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta y_3}{2} \right)^2 + B^2 \lambda^4 \frac{2}{3} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right] + \frac{Q_y^2}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \quad (10.4)$$

Here,  $L_1$  and  $L_2$  are the source to sample and sample to detector distances, respectively.  $R_1$  and  $R_2$  are the source and sample aperture radii, respectively, while  $\Delta x_3$  and  $\Delta y_3$  are the height and width dimensions of the detector pixels. The term with the parameter  $B$  incorporates the effect of gravity on the resolution function.  $B = gm_N^2 L_2 (L_1 + L_2) / 2h^2$ , where  $g$  is the gravity of the Earth (small  $g$ , 9.8 m/s<sup>2</sup>),  $m_N$  is the mass of the neutron ( $1.674927471(21) \times 10^{-27}$  kg) and  $h$  is Planck's constant ( $6.62607015 \times 10^{-34}$  J·s). Together  $m_N/h = 3995 \text{ \AA m s}^{-1}$ . The width and wavelength of the triangular wavelength distribution defined by the velocity selector is  $\Delta\lambda/\lambda$ . Note at the small-angle approximation,  $\cos\theta \approx 1$ , equations 10.3 and 10.4 converge into equations derived in other literatures, such as in Hammouda and Mildner, J. Appl. Cryst. V40, pp250-259 (2007).

Much of  $\sigma_{Qx}$  and  $\sigma_{Qy}$  remain unchanged for EQ-SANS, with the primary difference being the contribution due to the wavelength, where there are contributions due to the binning by wavelength that is performed during the data reduction and the uncertainty in the emission time of a neutron from the moderator. Both also become functions of the wavelength. It is important to note that the contribution of gravity is rather small for EQ-SANS due to the shorter wavelengths typically used and the shorter  $L_1$  and  $L_2$  of the instrument. The contribution due

to the wavelength is described in Appendix 4, and the full form of  $\sigma_{Qx}$  and  $\sigma_{Qy}$  are presented in Equations 10.5 and 10.6.

$$\sigma_{Qx}^2(\lambda_i) = \left( \frac{2\pi \cos\theta \cos^2 2\theta}{\lambda_i L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1 + L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \right] + \frac{Q_x^2}{12} \left[ \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 + \left( \frac{3.9560 \sigma_{em}(\lambda)}{1000 \lambda_i (s2p + m2s)} \right)^2 \right] \quad (10.5)$$

$$\sigma_{Qy}^2(\lambda_i) = \left( \frac{2\pi \cos\theta \cos^2 2\theta}{\lambda_i L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1 + L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta y_3}{2} \right)^2 + B^2 \lambda_i^4 \frac{2}{3} \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 \right] + \frac{Q_y^2}{12} \left[ \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 + \left( \frac{3.9560 \sigma_{em}(\lambda_i)}{1000 \lambda_i (s2p + m2s)} \right)^2 \right] \quad (10.6)$$

In Equations 10.5 and 10.6, the  $\lambda_i$  and  $\Delta \lambda_i$  are the center and width of a wavelength bin. The root-mean-standard deviation (RMSD) of the uncertainty in the emission time of a neutron from the moderator is  $\sigma_{em}(\lambda_i)$ ,  $s2p$  is the distance from the sample to the pixel and  $m2s$  is the moderator to sample distance. Both distances are provided in meters. Note the factor of 12 on the second term rather than 6. It results from the use of square bins and the full RMSD for  $\sigma(\lambda_i)$ .

## 11. Bin into Q

In the majority of SANS experiments, data are binned into  $Q$ -space –  $I(Q)$  or  $I(Q_x, Q_y)$  – for analysis. It is one of the last steps, in the data reduction process prior to writing the final reduced data. Data should be separately binned into 2D  $\vec{Q}$ -space ( $I(Q_x, Q_y)$ ) and into a 1D  $Q$ -space ( $I(Q)$ ). Additionally, data may be binned into 1D in alternative ways, such as using wedges or annuli of 2D data, that make it easier to extract useful information from the results. A detailed description of the various binning processes and recommendations are provided below.

### 11.1 Binning into 2D

2D (i.e.  $I(Q_x, Q_y)$ ) binning of the data is useful for looking for anisotropic scattering patterns. Many SANS patterns lack anisotropic scattering features and therefore do not require the use of 2D data during the analysis. It is anticipated that a greater fraction of SANS experiments performed in the future will involve fitting the 2D intensity profile. Thus, a best practice is to make  $I(Q_x, Q_y)$  available for inspection during and after the data reduction by retaining the 2D workspace and saving the 2D data file in a suitable format before generating the 1D binned data.

At present, the only style of 2D binning provided for the SANS at ORNL has a linear spacing of the  $Q$ -bins. Regardless of the instrument being used, the range of  $Q$  sampled by an instrument configuration is defined by the instrument geometry and the wavelength, or range of wavelengths, being used. The natural spacings in  $Q_x$  and  $Q_y$  for the HFIR SANS are defined

by the real-space X and Y dimensions of the pixel and the wavelength. Similarly, the  $Q_x$  and  $Q_y$  spacings most reasonable for EQ-SANS derive from the pixel sizes and the wavelength band used. A set of  $I(Q_x, Q_y, \lambda)$  is not produced during data reduction for the EQ-SANS. All wavelengths are binned to a single  $I(Q_x, Q_y)$ . In practice, the number of bins,  $N$  is specified during data reduction. A square grid (i.e.  $Q_x = Q_y$ ) is used for the primary detector and a separate grid is used for the Bio-SANS wing detector, but the ability to use different  $Q$ -spacing for X and Y, and therefore number of  $Q$ -bins  $N_x$  and  $N_y$ , will be provided in the future. Below are recommendations for the number of  $Q$ -bins, which does not assume that the grid must be square. As present, the default value for  $N$  is 128. If Fast Fourier Transform methods are to be used during data analysis, then setting  $N = 2^m$ , where  $m$  is an integer, facilitates doing so.

#### **Number of 2D Q-bins ( $N_x$ and $N_y$ )**

- GP-SANS: integers up to 192 (horizontal) or 256 (vertical)
- Bio-SANS: integers up to 192 or 160 (horizontal main and wing, respectively) or 256 (vertical)
- EQ-SANS: integers up to 256 (reasonable for both horizontal and vertical due to TOF)

With  $N_x$  and  $N_y$  defined, the extent of the  $Q$ -range is determined using the corners of the detector and  $\lambda$ , or the minimum wavelength  $\lambda_{min}$  on EQ-SANS, which provides four values:  $Q_{x,min}$ ,  $Q_{x,max}$ ,  $Q_{y,min}$  and  $Q_{y,max}$ . The 2D  $Q$  is a vector, so some  $Q$ -values for the coordinates of the corners are negative. Then, the widths of the  $Q$ -bins,  $\partial Q_x$  and  $\partial Q_y$ , are given by Equations 11.1 and 11.2.

$$\partial Q_x = \frac{(Q_{x,max} - Q_{x,min})}{N_x} \quad (11.1)$$

$$\partial Q_y = \frac{(Q_{y,max} - Q_{y,min})}{N_y} \quad (11.2)$$

The  $\vec{Q}$  of the  $(m, n)$   $Q$ -pixel in the 2D data, where  $m$  ranges from 0 to  $N_x - 1$  and  $n$  ranges from 0 to  $N_y - 1$ , is given by Equation 11.3.

$$\vec{Q}(m, n) = \langle \partial Q_x(m + 1/2) + Q_{x,min}, \partial Q_y(n + 1/2) + Q_{y,min} \rangle \quad (11.3)$$

Similarly, the locations of the four corners of the  $(m, n)$   $Q$ -pixel are given by the vectors in Equations 11.4 to 11.7.

$$\vec{Q}_1(m, n) = \langle \partial Q_x(m) + Q_{x,min}, \partial Q_y(n) + Q_{y,min} \rangle \quad (11.4)$$

$$\vec{Q}_2(m, n) = \langle \partial Q_x(m + 1) + Q_{x,min}, \partial Q_y(n) + Q_{y,min} \rangle \quad (11.5)$$

$$\vec{Q}_3(m, n) = \langle \partial Q_x(m) + Q_{x,min}, \partial Q_y(n + 1) + Q_{y,min} \rangle \quad (11.6)$$

$$\vec{Q}_4(m, n) = \langle \partial Q_x(m + 1) + Q_{x,min}, \partial Q_y(n + 1) + Q_{y,min} \rangle \quad (11.7)$$

The next step is to assign the reduced data  $I(x, y, \lambda)$  to the appropriate  $Q$ -bin. It is important to note that the precise overlap between a pixel in  $Q$ -space and the bins is not being performed to fractionally assign regions of a reduced data pixel to the  $Q$ -bin that it overlaps. If this effect is desired during the binning, it can be approximated by using sub-pixel binning, described below. The appropriate  $Q$ -bin for assignment is determined from the beam center vector  $\vec{B} = \langle B_x, B_y, B_z \rangle$ , the location of the pixel center  $\vec{P} = \langle P_x, P_y, P_z \rangle$ , and the location of the sample  $\vec{S} = \langle S_x, S_y, S_z \rangle$ , all of which are in the coordinate system of the instrument. From these three vectors, vectors  $\vec{V}_1$  and  $\vec{V}_2$  are calculated using Equations 11.8 and 11.9.

$$\vec{V}_1 = \vec{B} - \vec{S} \quad (11.8)$$

$$\vec{V}_2 = \vec{P} - \vec{S} \quad (11.9)$$

The scattering angle  $2\theta$  is given by Equation 11.10 and  $\vec{Q}$  is calculated using Equation 11.11.

$$2\theta = \cos^{-1} \left( \frac{\vec{V}_1 \cdot \vec{V}_2}{|\vec{V}_1| |\vec{V}_2|} \right) \quad (11.10)$$

$$\vec{Q} = \left( \frac{4\pi \sin \theta}{\lambda} \right) \left( \frac{\vec{P} - \vec{B}}{|\vec{P} - \vec{B}|} \right) = \langle Q_x, Q_y \rangle \quad (11.11)$$

The intensity  $I(x, y, \lambda)$  is then assigned to the  $Q$ -bin that satisfies the criteria in Equations 11.12 and 11.13.

$$\partial Q_x(m) + Q_{x,min} < Q_x < \partial Q_x(m+1) + Q_{x,min} \quad (11.12)$$

$$\partial Q_y(n) + Q_{y,min} < Q_y < \partial Q_y(n+1) + Q_{y,min} \quad (11.13)$$

Call the number of  $I(x, y, \lambda)$  and the associated uncertainty  $\delta I(x, y, \lambda)$  that contribute to a bin  $N_{x,y,\lambda}$ , then the binned intensity, its uncertainty and uncertainty in  $Q$  are given by Equations 11.14, 11.15 and 11.16 when error-weighting is not used.

$$I(\vec{Q}) = \frac{\sum_{x,y,\lambda} I(x,y,\lambda)}{N_{x,y,\lambda}} \quad (11.14)$$

$$\delta I(\vec{Q}) = \frac{(\sum_{x,y,\lambda} (\delta I(x,y,\lambda))^2)^{1/2}}{N_{x,y,\lambda}} \quad (11.15)$$

$$\sigma_Q(\vec{Q}) = \frac{\sum_{x,y,\lambda} \sigma_Q(x,y,\lambda)}{N_{x,y,\lambda}} \quad (11.16)$$

If sub-pixel binning is employed, the actual pixel is divided into a grid with horizontal and vertical elements  $(N_A, N_B)$ . It is recommended that no more than 10 sub-pixels in either direction be employed. The height and width of the sub-pixels are given by the relations in Equation 11.17.

$$\partial Q_A = \frac{\partial Q_x}{N_A} \text{ and } \partial Q_B = \frac{\partial Q_y}{N_B} \quad (11.17)$$

The  $\vec{Q}$  of the sub-pixel is given by Equation 11.18, where  $A$  ranges from 0 to  $N_A - 1$  and  $B$  ranges from 0 to  $N_B - 1$ .

$$\vec{Q}(m, A, n, B) = \left\langle \begin{array}{l} \partial Q_x(m + 1/2) + \partial Q_A \left( A - \frac{N_A}{2} + \frac{1}{2} \right) + Q_{x,min}, \\ \partial Q_y(n + 1/2) + \partial Q_B \left( B - \frac{N_B}{2} + \frac{1}{2} \right) + Q_{y,min} \end{array} \right\rangle \quad (11.18)$$

The values, uncertainties and  $Q$ -value uncertainties of the sub-pixels are the same as the original values. The binning performed is shown in Equations 11.19, 11.20 and 11.21. The additional pixels are accounted for in the number shown in the denominator, which is now the number of subpixels that contribute to the bin.

$$I(\vec{Q}) = \frac{\sum_{x,A,y,B,\lambda} I(x,A,y,B,\lambda)}{N_{x,A,y,B,\lambda}} \quad (11.19)$$

$$\delta I(\vec{Q}) = \frac{\left( \sum_{x,A,y,B,\lambda} (\delta I(x,A,y,B,\lambda))^2 \right)^{1/2}}{N_{x,A,y,B,\lambda}} \quad (11.20)$$

$$\sigma_Q(\vec{Q}) = \frac{\sum_{x,A,y,B,\lambda} \sigma_Q(x,A,y,B,\lambda)}{N_{x,A,y,B,\lambda}} \quad (11.21)$$

When error-weighted averaging is used, the uncertainties in the intensities ( $\sigma(m, n)$  or  $\sigma(m, A, n, B)$ ) must be calculated first, then  $I(\vec{Q})$ ,  $\delta I(\vec{Q})$  and  $\sigma_Q(\vec{Q})$  are given by Equations 11.22, 11.23 and 11.24, which do not include subpixel binning.

$$I(\vec{Q}) = \frac{\sum_{x,y,\lambda \in Q_{bin}} \left( I(x,y,\lambda) / (\sigma(x,y,\lambda))^2 \right)}{\sum_{x,y,\lambda \in Q_{bin}} \left( 1 / (\sigma(x,y,\lambda))^2 \right)} \quad (11.22)$$

$$\delta I(\vec{Q}) = \frac{\left( \sum_{x,y,\lambda \in Q_{bin}} \left( \delta I(x,y,\lambda) / (\sigma(x,y,\lambda))^2 \right)^2 \right)^{1/2}}{\sum_{x,y,\lambda \in Q_{bin}} \left( 1 / (\sigma(x,y,\lambda))^2 \right)} \quad (11.23)$$

$$\sigma_Q(\vec{Q}) = \frac{\sum_{x,y,\lambda \in Q_{bin}} \left( \sigma_Q(x,y,\lambda) / (\sigma(x,y,\lambda))^2 \right)}{\sum_{x,y,\lambda \in Q_{bin}} \left( 1 / (\sigma(x,y,\lambda))^2 \right)} \quad (11.24)$$

If subpixel binning is used with error-weighted averaging, the summations shown in Equations 11.22-11.24 also span over the subpixel indices  $A$  and  $B$  in the manner shown in Equations 11.19-11.21.

## 11.2 Binning into 1D

The most common form of SANS data that are analyzed for structural information are data that have been binned into 1D  $Q$ -space. Below, two types of 1D binning are presented: (1) azimuthal, and variants thereof, where 1D data as a function of  $Q$  are produced; and (2) annular where the intensity at a value of  $Q$  as a function of the angle  $\Psi$  around the beam stop are produced.

### 11.2.1 Azimuthal Binning

Azimuthal binning produces a one dimensional  $I(Q)$  from  $I(P_x, P_y)$ . Users nearly always visualize and model SANS data as  $I(Q)$ . In addition to binning the entire detector, wedges of data may be binned when working with anisotropic data. Lines and boxes may be used to define the data to be binned from 2D to 1D, such as for quickly examining the position of a peak. In the case of binning data that lies in a box, the box need not include the origin. One possible application of this kind of 1D binning would be taking a 2D diffraction peak to 1D to quickly fit it. Regardless of exactly what data is to be included in the azimuthal binning, the user must decide if linear or logarithmic spacing of  $Q$  is desired, and how many points in the curve are desired. In the future, other variants of spacing in  $Q$  may be implemented.

#### 11.2.1.1 Linear $Q$ -spacing

When using linear binning, the number of bins is approximately determined by the number of pixels between the beam center and the furthest corner. Approximately for the GP-SANS this equates to ~230  $Q$ -bins counting from the middle of the length of a tube (128) and



between the two furthest tubes (192). In the case of EQ-SANS, the beam is almost always in the middle of the detector, making the natural maximum  $\sim 160$ , but the TOF allows the instrument to cover a broader  $Q$ -range in a smaller detector. Experience indicates that  $\sim 200$   $Q$ -bins is a reasonable high end for some configurations, such as 4 m/2.5 Å, with a strong scattering sample. However, for 1.3 m with 2 Å or less as the minimum setting, 300  $Q$ -bins is not unreasonable, but 400 is arguably a reasonable upper limit. We recommend that values up to 400 be allowed for the number of 1D  $Q$ -bins, but that the user be provided with a notification message if values above 300 are used. The use of multiple detectors on the Bio-SANS complicates how one might determine a reasonable number of  $Q$ -bins to use, and it will vary depending on whether the main or wing detector is being considered. Again, suggesting an upper limit of 300 for the instrument with warnings provided to the user for values over 200 is reasonable. For a lower limit, suggesting a value of 50  $Q$ -bins is reasonable. With too few  $Q$ -bins in any given data set, it becomes difficult to merge data from different configurations, even though the temptation exists to try to improve counting statistics for weakly scattering samples in this manner.

The specific points in  $Q$  to be output must be determined once  $N_{bins}$  has been specified. This process also determines the boundaries of the bins in  $Q$ . First, set the minimum  $Q$ ,  $Q_{min}$ , to  $0.0 \text{ Å}^{-1}$ . While the beam makes it impossible to get to this value, it is a natural reference. The number of  $Q$ -bins lost due to being in the beam or behind the beam stop will be minimal. The maximum  $Q$ ,  $Q_{max}$ , is set to the value of the furthest corner from the beam center, which is located at  $(0,0,L_2)$ , where  $L_2$  is the sample-to-detector distance (nomenclature from Chapter 10). Let this corner be the  $(i,j)$  pixel having real-space coordinates  $(x_{ij}, y_{ij}, z_{ij})$ , which has

dimensions  $(\Delta x_{3,ij}, \Delta y_{3,ij})$  (nomenclature from Chapter 10), then  $Q_{max}$  is given by Equation 11.25.

$$Q_{max} = \frac{4\pi}{\lambda} \sin \left[ 0.5 \tan^{-1} \left( \frac{\sqrt{\left(x_{ij} + \frac{\Delta x_{3,ij}}{2}\right)^2 + \left(y_{ij} + \frac{\Delta y_{3,ij}}{2}\right)^2 + (L_2 - z_{ij})^2}}{L_2} \right) \right] \quad (11.25)$$

Specifying the  $Q$ -range for the 1D binning in this manner for Bio-SANS, even with the two detectors, would provide the benefit of a single output profile with consistent  $Q$ -spacing in the final output. However, the relative normalization of the data may preclude doing so. With  $N_{bins}$  specified, the  $Q$ -value assigned to the  $K$ -th bin, where  $K \in (0, N_{bins} - 1)$ , that has a width  $\partial Q = Q_{max}/N_{bins}$  is given by Equation 11.26.

$$Q_K = \partial Q (K + 1/2) \quad (11.26)$$

The boundaries of this bin are determined according to Equation 11.27.

$$Q_{K,min} = (Q_K - 0.5 \cdot \partial Q) \text{ and } Q_{K,max} = (Q_K + 0.5 \cdot \partial Q) \quad (11.27)$$

#### 11.2.1.2 Logarithmic $Q$ -spacing

When logarithmic  $Q$ -binning is employed, the user has more options for determining the bins to be used, which are detailed below. First,  $Q_{min}$  must be selected from one of two possibilities.

- $Q_{min}$  is set to  $0.0001 \text{ \AA}^{-1}$ , which is below where the GP-SANS, Bio-SANS and EQ-SANS can measure, or
- $Q_{min}$  is set to a user-specified value, or

- $Q_{min}$  is set to  $\frac{2\pi}{\lambda} \sin \left[ 0.5 \tan^{-1} \left( \frac{\sqrt{\left(\frac{\Delta x_{3,ij}}{2}\right)^2 + \left(\frac{\Delta y_{3,ij}}{2}\right)^2}}{L_2} \right) \right]$ .

Then, the quantity  $C_{min} = \log Q_{min}$  is calculated. Next,  $Q_{max}$  is calculated according to Equation 11.25, or it is set to a user specified value, and the quantity  $C_{max}$  is calculated in one of the two following ways.

- $C_{max} = \text{ceil}(\log Q_{max})$ , where  $\text{ceil}(X)$  rounds  $X$  up to the nearest integer, or
- $C_{max} = \log Q_{max}$

Next, the number of  $Q$  bins,  $N$ , must be specified in one of the two following ways.

- The user specifies the number of bins per decade,  $N_{bin}$ , and the total number of  $Q$ -values is given by  $N = N_{bin}(C_{max} - C_{min})$ , or
- the user specifies the number of  $Q$ -bins,  $N$ .

Finally, the size of the step in  $\log Q$ ,  $\partial L$ , is given by Equation 11.28.

$$\partial L = \frac{(C_{max} - C_{min})}{N} \quad (11.28)$$

The  $K$ -th  $Q$ -value, where  $K \in (0, N - 1)$ , is given by Equation 11.29.

$$Q_K = 10^{\partial L \left(K + \frac{1}{2}\right) + C_{min}} \quad (11.29)$$

The boundaries of the  $K$ -th bin are determined according to Equation 11.30, except for the first and last bins.

$$Q_{K,min} = \frac{(Q_K + Q_{K-1})}{2} \text{ and } Q_{K,max} = \frac{(Q_{K+1} + Q_K)}{2} \quad (11.30)$$

For the first bin, set  $Q_{0,min} = Q_{min}$  and for the last bin,  $Q_{N-1,max} = Q_{max}$ .

In cases where the  $Q$ -range spans even decades and the user specifies the number of bins per decade to be used, one can instead specify that the first point lies on a decade using Equation 11.31.

$$Q_K = 10^{\partial L K + C_{min}} \quad (11.31)$$

The bin boundaries are calculated with Equation 11.30, except for the first and last bin, which are given by  $Q_{0,min} = 10^{C_{min} - \frac{\partial L}{2}}$  and  $Q_{N-1,max} = 10^{C_{max} + \frac{\partial L}{2}}$ . In this case, the total number of bins to be used should be  $N = N_{bin}(C_{max} - C_{min}) + 1$  to capture the final point on the decade.

#### 11.2.1.3 Binning into 1D

The final step in the azimuthal 1D-binning process is to perform the binning. Note that sub-pixel binning is not being considered here, but a description is provided at the end of this section. Equation 11.31 is used regardless of whether the azimuthal averaging is being applied to the full pattern, a wedge or wedges, a box, a line, or an arbitrary region of interest (ROI). While the full pattern average is most often used and conceptually simple to understand, the latter three types of azimuthal averages are rarely used but do not differ mathematically from the methods for the full pattern or for wedges. In contrast wedge averaging is frequently employed when dealing with anisotropic data, and methods for automatically choosing and outputting wedge-averaged data have been implemented. Therefore, a description of different ways how wedges can be defined follows.

#### *Defining wedges*

Let the wedge to be integrated span between two angles,  $\Psi_{min}$  and  $\Psi_{max}$ , in the pattern about the beam center into a 1D profile, as is shown in Figure 10.1. The angle of a pixel  $(m, n)$  is given by Equation 11.32. All rules related to the application of the inverse sine function in the plane apply to ensure that the angle is mapped to the appropriate quadrant for determining if it truly belongs in the  $Q$ -bin or not.

$$\Psi(m, n) = \sin^{-1} \frac{Q_y}{(Q_x^2 + Q_y^2)^{1/2}} \quad (11.32)$$

In addition to being able to specify a wedge by  $\Psi_{min}$  and  $\Psi_{max}$ , it is possible to specify the midpoint angle  $\Psi_{mid}$  and the angular width of the wedge to be integrated  $\psi$ . Then, the values of  $\Psi(m, n)$  included in the wedge are those between  $\Psi_{mid} - 0.5\psi$  and  $\Psi_{mid} + 0.5\psi$ .

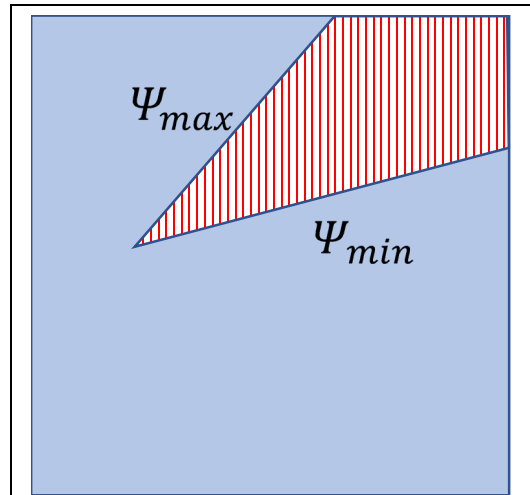
An expansion of this notion is useful for applying  $N$ -fold symmetry to automatically provide a set of  $0 \dots N - 1$  wedges of width  $\psi$  centered at  $\Psi_{mid,j}$ , as shown in Equation 11.33.

$$\Psi_{mid,j} = \Psi_0 + \frac{j360^\circ}{N} \quad (11.33)$$

Another method for defining wedges is to automatically find the peaks and valleys found around the beam center in angular space. This approach relies on the annual average, described below, and is an algorithm that can be called to return the wedges to be integrated.

#### *Performing the binning*

With the 1D array of  $Q$ -bins defined, the pixels to be included identified, and the desired



**Figure 11.1.** Schematic of a wedge average. The red, striped area is the region to be binned into 1D  $Q$ -space.

approach for the azimuthal average specified, the value assigned to each  $Q$ -bin can be determined. When  $Q_{K,min} \leq |\vec{Q}(m,n)| < Q_{K,max}$ , that 2D pixel is assigned to the  $K$ -th bin. If after processing the 2D  $Q$ -space map there are  $N_K$  pixels with values  $I(m,n)$ , uncertainties  $\delta I(m,n)$  and  $Q$ -value uncertainty  $\sigma_Q(m,n)$ , then the 1D intensity  $I(Q_K)$ , uncertainty in the intensity  $\delta I(Q_K)$  and the  $Q$ -value uncertainty  $\sigma_Q(Q_K)$  are given by Equations 11.34-11.36.

$$I(Q_K) = \frac{\sum_{m,n} I(m,n)}{N_K} \quad (11.34)$$

$$\delta I(Q_K) = \frac{(\sum_{m,n} \delta I(m,n)^2)^{1/2}}{N_K} \quad (11.35)$$

$$\sigma_Q(Q_K) = \frac{\sum_{m,n} \sigma_Q(m,n)}{N_K} \quad (11.36)$$

To perform sub-pixel binning, each pixel is divided into a grid having horizontal and vertical elements ( $N_A, N_B$ ). The height and width of the sub-pixels are then given by Equation 11.37.

$$\partial Q_A = \frac{\partial Q_x}{N_A} \text{ and } \partial Q_B = \frac{\partial Q_y}{N_B} \quad (11.37)$$

The  $Q$ -vector of the sub-pixel is given by Equation 11.38.

$$\vec{Q}(m, A, n, B) = \left\langle \begin{array}{l} \partial Q_x(m + 1/2) + \partial Q_A \left( A - \frac{N_A}{2} + \frac{1}{2} \right) + Q_{x,min}, \\ \partial Q_y(n + 1/2) + \partial Q_B \left( B - \frac{N_B}{2} + \frac{1}{2} \right) + Q_{y,min} \end{array} \right\rangle \quad (11.38)$$

The values, uncertainties and  $Q$ -value uncertainties of the sub-pixels the same as the original values. The binning performed is shown in Equations 11.39, 11.40 and 11.41. The additional pixels are accounted for in the number shown in the denominator, which is now the number of subpixels that contribute to the bin.

$$I(Q) = \frac{\sum_{x,A,y,B,\lambda} I(x,A,y,B,\lambda)}{N_{x,A,y,B,\lambda}} \quad (11.39)$$

$$\delta I(Q) = \frac{(\sum_{x,y,\lambda} (\delta I(x,A,y,B,\lambda))^2)^{1/2}}{N_{x,A,y,B,\lambda}} \quad (11.40)$$

$$\sigma_Q(Q) = \frac{\sum_{x,y,\lambda} \sigma_Q(x,A,y,B,\lambda)}{N_{x,A,y,B,\lambda}} \quad (11.41)$$

All of the operations for binning into a 1D  $I(Q)$  vs  $Q$  profile follow those employed for using full pixel shown in Equations 11.34-11.36, but the summations span all sub-pixels and  $N_K$  is the number of sub-pixels, rather than pixels.

When error-weighted averaging is used,  $I(Q_K)$  is given by Equation 11.42, which is readily expanded for sub-pixel binning as shown in Equation F. The calculation of  $\delta I(Q_K)$  and  $\sigma_Q(Q_K)$  do not change.

$$I(Q_K) = \frac{\sum_{m,n} \left( I(m,n) / (\sigma(m,n))^2 \right)}{\sum_{m,n} \left( 1 / (\sigma(m,n))^2 \right)} \quad (11.42)$$

At the end of the azimuthal binning process, all bins above  $Q_{max}$  (Equation 11.25) are removed from the array by resizing it.

### 11.2.2. Annular Binning

Annular binning differs from the 1D binning methods specified above in that it is not providing the data as a function of  $Q$ . Instead, the aim is to take the 2D  $I(Q_x, Q_y)$  and integrate over a limited range of scalar  $Q$  to produce the intensity as a function of the angle around the beam center,  $\Psi$ . In practice, the user specifies  $Q_{min}$  and  $Q_{max}$  as the start and end values of scalar  $Q$  to use for the average. In the case of an annular average, the only logical style of binning is linear spacing in  $\Psi$ . The number of bins,  $N_\Psi$ , should not exceed 200. Lower numbers of bins are more appropriate if the  $Q$ -range of the annulus for performing the binning is near the beam stop.  $\Psi$  is calculated using Equation 11.32. The  $Q_x$  axis is  $\Psi = 0.0$ , and the angle

increases in the clockwise direction. It is possible to specify an offset value  $\Psi'$  to shift a feature in the data, such as a peak or a valley, to  $\Psi = 0.0$ . Then, the spacing in  $\Psi$ ,  $\partial\Psi$ , and the centers and edges of the  $K$ -th bin ( $K \in (0 \dots N_\Psi - 1)$ ) are given by Equations 11.43-11.45.

$$\partial\Psi = \frac{360^\circ}{N_\Psi} \quad (11.43)$$

$$\Psi_K = (K + 1/2) \partial\Psi \quad (11.44)$$

$$\Psi_{K,min} = K\partial\Psi \text{ and } \Psi_{K,max} = (K + 1)\partial\Psi \quad (11.45)$$

First, consider the case where sub-pixel binning is not employed. A pixel contributes to the  $K$ -th bin if  $\Psi_{K,min} \leq \Psi(m, n) < \Psi_{K,max}$ . If the  $K$ -th bin has  $N_K$  pixels that contribute to it and have value  $I(m, n)$ , uncertainty  $\delta I(m, n)$  and  $Q$ -value uncertainty  $\sigma_Q(m, n)$ , then the intensity  $I(\Psi_K)$ , uncertainty in the intensity  $\delta I(\Psi_K)$  and the  $Q$ -value uncertainty in that bin,  $\sigma_Q(\Psi_K)$ , are given by Equations 11.46-11.48.

$$I(\Psi_K) = \frac{\sum_{m,n} I(m, n)}{N_K} \quad (11.46)$$

$$\delta I(\Psi_K) = \frac{(\sum_{m,n} \delta I(m, n)^2)^{1/2}}{N_K} \quad (11.47)$$

$$\sigma_Q(\Psi_K) = \frac{\sum_{m,n} \sigma_Q(m, n)}{N_K} \quad (11.48)$$

In the event that sub-pixel binning is employed, pixel subdivision into horizontal and vertical elements ( $N_A, N_B$ ) follows Equations 11.37 and 11.38. All of the operations for binning into a 1D  $I(Q)$  vs.  $Q$  profile follow those employed for using full pixel shown in Equations 11.34-11.36, but the summations span all sub-pixels and  $N_K$  is the number of sub-pixels, rather than pixels. Similarly, error-weighted averaging follows the description above, as shown in Equation 11.49.



$$I(\Psi_K) = \frac{\sum_{m,n} \left( I(m,n) / (\sigma(m,n))^2 \right)^2}{\sum_{m,n} \left( 1 / (\sigma(m,n))^2 \right)} \quad (11.49)$$

## 12. Final Normalization

A key step during data reduction is to place the data into absolute units of  $cm^{-1}$ , which takes the measured intensity  $I(Q)$  and scales it into the differential scattering cross section per unit volume,  $\partial\Sigma/\partial\Omega$ . Scaling data into absolute units makes it possible to determine the particle number density, molecular weight, volume fraction, specific surface area and other fitting parameters when using a model to describe the observed intensities. It is also possible to reveal the presence of measurement or sample artifacts. Performing the scaling makes it possible to complete studies using different instruments, which is critical given the limited availability of beam time.

The relationship between the measured intensity and the differential scattering cross section per unit volume is given by Equation 12.1, which is applicable to a monochromatic SANS instrument. Note that the data has already been reduced to  $I(Q)$ , and the scaling holds true for 2D data.

$$I(Q) = \phi A d T \frac{\partial\Sigma}{\partial\Omega} \Delta\Omega \varepsilon t \quad (12.1)$$

Here,  $\phi$  is the neutron flux on the sample,  $A$  is the illuminated area,  $d$  is the sample thickness, which assume a planar sample,  $T$  is the measured transmission,  $\Delta\Omega$  is the solid angle subtended by each pixel of the detector,  $\varepsilon$  is the detector's wavelength-dependent neutron collection efficiency,  $t$  is the counting time. An alternative expression is provided in Equation 12.2, which replaces the illuminated area and sample thickness in Equation 12.1 with the sample volume,  $V$ , which is a useful way to think of the relationship in cases with irregular sample geometries and conditions where the sample could be under-illuminated. Again, the scaling in Equation 12.2 holds for 2D data.

$$I(Q) = \phi VT \frac{\partial \Sigma}{\partial \Omega} \Delta \Omega \epsilon t \quad (12.2)$$

Two methods exist for scaling a measured intensity into absolute units: with an attenuated direct beam measurement, or with a measurement of a calibrated standard sample. The latter is sometimes referred to as the secondary standard method.

### 12.1 Normalization for illuminated volume

The first step in performing the final normalization of the measured data is to normalize for the illuminated volume. Traditionally, the normalization was performed for the sample thickness, which assumes that the sample is fully illuminated. As noted above, it is possible that a sample could be under-illuminated relative to an associated background measurement or the standard being used for the absolute normalization. Therefore, the data reduction for the instruments at ORNL normalize measured data  $I(Q)$  with uncertainty  $\delta I(Q)$  for the illuminated volume  $V$  according to Equations 12.3 and 12.4.

$$I'(Q) = \frac{I(Q)}{V} \quad (12.3)$$

$$\delta I'(Q) = \frac{\delta I(Q)}{V} \quad (12.4)$$

This scaling is also used for the 2D data.

### 12.2 Normalization with an attenuated empty beam

One approach for scaling the data into absolute units is with an empty beam measurement, which provides a measure of the neutron flux,  $\phi(\lambda)$ , in neutrons per second per unit area, that are incident on the sample. In practice, the attenuated, direct beam is measured

with the main detector of the instrument by using a calibrated attenuator upstream of the first collimating aperture. The measurements are made with the same instrument configuration as is used for the sample scattering measurements and require a few minutes per configuration. This method for determining the absolute scale factor does not require mounting a calibrated standard into any specific sample environment, which may be quite challenging for some sample environments, such as pressure cells, rheometers and magnets. This approach for determining the scale factor is most commonly used on GP-SANS. It could be used more regularly on Bio-SANS because the HFIR instruments are very similar. This method for determining the absolute scale factor has not been used on EQ-SANS, though it is not precluded.

The measurement of the direct beam for determining the absolute scale factor is no different than measuring a transmission measurement. However, unlike a transmission measurement that uses a reference measurement for the calculation of the value (see Section 7, above), the direct beam must be measured in an absolute sense. Therefore, the count rate must absolutely be within the linear response regime of the detector, the entire beam spot must be integrated, the beam monitor must be reliable, and most importantly, the wavelength-dependence of the attenuation coefficient of the attenuator used for the measurements must be known. The last requirement poses the greatest challenge for using this method, particularly on the EQ-SANS instrument.

For this approach, empty direct beam measurement is acquired in the same instrument configuration for sample and background measurements. The exceptions to the configuration are the positions of the beam stop and the attenuator. The beam stop is displaced downward

by a distance at least equal to the diameter of the beam stop used to allow the entire direct transmitted beam to impinge the detector. With the entire beam impinging the detector unhindered, an appropriate attenuator is used to ensure pixel counting rate is within the linear response regime of the detector for this measurement. The summed intensity of the beam spot,  $I_{BS}$  and its uncertainty  $\delta I_{BS}$  is given by equations 12.5 and 12.6, respectively.

$$I_{BS} = \sum_{x,y; \|\vec{r}_{x,y}\| < R_{beam}} I(x, y) \quad (12.5)$$

$$\delta I_{BS} = \sqrt{I_{BS}} \quad (12.6)$$

The pixels considered for producing the summed intensity fall within the beam spot of radius,  $R_{beam}$ . With an intensity scaling factor  $\chi$  and uncertainty  $\delta\chi$  for the attenuator, the total intensity and uncertainty is given by equations 12.7 and 12.8, respectively.

$$I_T = \chi I_{BS} \quad (12.7)$$

$$\delta I_T = I_{BS} \left[ \left( \frac{\delta\chi}{\chi} \right)^2 + \left( \frac{\delta I_{BS}}{I_{BS}} \right)^2 \right]^{1/2} \quad (12.8)$$

Ultimately, the data,  $I(Q_x, Q_y)$  are normalized to the intensity scalar similar to equations 12.3 and 12.4, above.

$$I'(Q_x, Q_y) = \frac{I(Q_x, Q_y)}{I_T} \quad (12.9)$$

$$\delta I'(Q_x, Q_y) = I'(Q_x, Q_y) \left[ \left( \frac{\delta I_T}{I_T} \right)^2 + \left( \frac{\delta I(Q_x, Q_y)}{I(Q_x, Q_y)} \right)^2 \right]^{1/2} \quad (12.10)$$

### 12.3 Normalization with a calibrated standard sample

The second method for obtaining the factor for scaling data into absolute units uses a calibrated standard sample, also known as a secondary standard. By measuring a material that

exhibits well-defined and reproducible scattering using the same configurations as are being used for the other measurements in the experiment, it is possible to obtain the absolute scale factor. Research at ORNL led to the proposal of a set of secondary standard samples, along with descriptions of the advantages and disadvantages of the materials [5]. Materials that have been used as secondary standards include water, vanadium, irradiated aluminum and porous silica. Presently, porous silica and water are the most commonly used secondary standards on the SANS at ORNL, with two porous silica secondary standards being maintained by the instrument staff.

During an experiment, the standard is measured, subjected to a full data reduction, and the results are examined. Rather than using the intensity directly, a feature of the data measured from the standard is measured to perform the scaling. For example, the scattering from water and vanadium are predominantly incoherent, which gives a flat  $I(Q)$ . The level of the flat scattering, denoted  $F$  and having an uncertainty  $\delta F$ , is determined from the data and can be scaled to the known value of the scattering level of the material, denoted  $F_{std}$ , which may also have an uncertainty  $\delta F_{std}$ . In the case of the porous silica, the data is reduced and fit as a Debye-Bueche curve ( $(I(Q))^{-1/2}$  vs.  $Q^2$ ) over a  $Q$ -range of  $0.01 - 0.03 \text{ \AA}^{-1}$  for the  $I(0) = F$  and its uncertainty  $\delta I(0) = \delta F$ . Then, a measured data set  $I(Q)$  with uncertainty  $\delta I(Q)$  can be related to its differential scattering cross section per unit volume by Equation 12.11, and the uncertainty is propagated according to Equation 12.12.

$$\frac{\partial \Sigma(Q)}{\partial \Omega} = \frac{F_{std} I(Q)}{F} \quad (12.11)$$

$$\delta \frac{\partial \Sigma(Q)}{\partial \Omega} = \left[ \left( \frac{F_{std}}{F} \left[ \left( \frac{\delta F}{F} \right)^2 + \left( \frac{\delta F_{std}}{F_{std}} \right)^2 \right]^{1/2} I(Q) \right)^2 + \left( \frac{F_{std}}{F} \delta I(Q) \right)^2 \right]^{1/2} \quad (12.12)$$

### 13. Write the reduced data file

After reduction, the data must be output in a useful format. At this step in the process, the data have been reduced to the 1D profile,  $I(Q)$ , and the 2D data,  $I(Q_x, Q_y)$ . Specific cuts through the data, such as a box, line, or wedge binning, which are also 1D profiles, may also be output at this point in the data reduction. The data reduction process also produces a log that contains input reduction parameters, metadata extracted from the data files, results derived from the data for use in the data reduction process, such as the sample and background transmissions, and results obtained by processing a region of interest in the data. Details regarding how the data and the logfiles, which are now information-rich repositories of information related to the data reduction that can be browsed after completion, are described below.

Three important details regarding the data that is output must be adhered to. First, all zero and negative values must be output, rather than being dropped for being unphysical or because graphing software cannot handle the values. The former should not preclude the data value being output, while the latter has rarely been observed to be true. In the case of 2D data, any pixels that were masked on the detector and remain unassigned at the time of the final output must be identified as having been masked because the entire Q-space map is to be output. Distinctive values, such as -inf or NaN, are acceptable. Finally, when a data file is re-reduced, the original results are over-written by the new results. If different file outputs are specified, then the files output by the original reduction are not over-written or deleted. The various file formats are stored in different subdirectories to make it easier to work the desired data format.



### 13.1 Output Formats for 1D $I(Q)$ data

SANS data analysis is most often performed using  $I(Q)$ , and the majority of data analysis programs read simple ASCII files containing the profile. However, other formats for reduced SANS and SAXS data exist, such as the CanSAS-defined formats that provide portability, as well as compatibility between facilities and instruments that use the standardized formats. The IgorPro data reduction software provided ASCII formatted data by default, while Mantid originally provided both ASCII data and CanSAS -formatted XML data file by default. Both of these formats continue to be available, and either or both can be saved according to user specifications. File names are based on the run number, but this default behavior can be overridden if a user provides descriptive file names.

#### 13.1.1 ASCII 1D $I(Q)$ data

Data to be output and header information to be included in the files are as follows.

Line 1: Name of the reduction logfile from the process that produced it

Line 2: Names of the columns in the file and units thereof

Line 3: Start of the data in the file saved as  $\langle Q, I(Q), \delta I(Q), \sigma_Q(Q) \rangle$

All data stored in the file are sorted by ascending  $Q$ -value, and the data lines are tab-separated float values, rather than being written in scientific notation. The standard file extension is “\_lq.txt”. An example of the first few lines of an ASCII 1D data file follow that illustrate the format.

```
# reduction log: run12345_reduction_log.hdf
```

# Q (1/Å)	I (1/cm)	dI (1/cm)	dQ (1/Å)
0.001000	100.000000	1.000000	0.000010
0.001500	99.000000	1.001000	0.000009

### 13.1.2 CanSAS XML 1D $I(Q)$ data

The CanSAS XML format uses tags to specify the information that a value contains. At a minimum, the CanSAS XML data files will contain the following header and data information.

Line 1: Name of the reduction logfile that produced it

Line 2: Start of the data in the file saved as  $\langle Q, I(Q), \delta I(Q), \sigma_Q(Q) \rangle$  with units specified, per the CanSAS standard.

All data stored in the file are sorted by ascending  $Q$ -value, and the data lines are formatted per the specification, and the standard file extension is “\_lq.xml”. The numbers are saved as float values, rather than being provided in scientific notation.

## 13.2 Output Formats for 2D $I(Q_x, Q_y)$ data

It is increasingly common for data analysis to be performed with the 2D reduced data  $I(Q_x, Q_y)$ , and occasionally the 2D data must be inspected to understand odd behavior in the 1D profile, such as when the sample was not well-aligned with respect to the beam. Here, an ASCII file format for 2D data is described. Further, a Nexus-formatted file is specified, which is a native format of Mantid and enables simple re-loading of 2D data from a file. Both of these formats can be saved according to user specifications after data reduction. File names are

based on the run number, but this default behavior can be over-written if a user provides descriptive file names.

### 13.2.1 ASCII 2D $I(Q_x, Q_y)$ data

Data to be output and header information to be included in the files are as follows.

Line 1: Name of the reduction logfile from the process that produced it

Line 2: Names of the columns in the file and units thereof

Line 3: Start of the data in the file saved as

$$\langle Q_x, Q_y, I(Q_x, Q_y), \delta I(Q_x, Q_y), \sigma_{Q_x}, \sigma_{Q_y} \rangle$$

Here,  $\langle \sigma_{Q_x}, \sigma_{Q_y} \rangle$  are coordinates of the  $\sigma(\vec{Q})$ . All data stored in the file are sorted by ascending  $Q$ -value, in  $Y$  for each value of  $X$ , and the data lines are tab-separated float values, rather than being written in scientific notation. The standard file extension is “\_lq.dat”. An example of the first few lines of an ASCII 2D data file follow that illustrate the format.

```
# reduction log: run12345_reduction_log.hdf

# Qx (1/A)   Qy (1/A)   I (1/cm)   dI (1/cm)   dQx (1/A)   dQy (1/A)
0.001000    0.001000    100.000000  1.000000    0.000010    0.000011
0.001000    0.001500    99.000000   1.001000    0.000009    0.000010
```

### 13.2.2 Nexus-formatted 2D $I(Q_x, Q_y)$ data

The Nexus format is the native format for EPICS, PyDAS and Mantid. At a minimum, the reduced data set saved in the Nexus format will contain the following header and data information.

Header: Name of the reduction logfile that produced it

Data: data saved as  $\langle Q_x, Q_y, I(Q_x, Q_y), \delta I(Q_x, Q_y), \sigma_{Q_x}, \sigma_{Q_y} \rangle$  with units specified according to the Nexus standard.

Here,  $\langle \sigma_{Q_x}, \sigma_{Q_y} \rangle$  are coordinates of the  $\sigma(\vec{Q})$ . All data stored in the file are sorted by ascending  $Q$ -value, in Y for each value of X, and the data lines are tab-separated float values, rather than being written in scientific notation. The standard file extension is “\_lq.nxs”. All information is saved in the native format of the Nexus standard.

### 13.3 The log file of the reduction

The log file of the reduction retains all relevant information from a reduction process. It serves as a record that the user can refer to in the event of errors, and it provides a place for derived values and results, such as sample transmission and the value of an integrated region of interest, to be stored for use in later data analysis. The goal of the logfile is to provide a user with the full provenance of the reduced data without adding a great deal of information to the actual files that would be used for data analysis. The information contained in the logfile is extensive, and as a result, it is stored in a hierarchical format using the HDF5 standard having a file extension “\_reduction\_log.hdf”. The following information is contained in the logfile

- Reduction\_log.hdf
  - Timestamp for the reduction process
  - User ID and computer
  - Version and build of Mantid
    - Mantid logging level

- Script for the reduction of a file
  - Script file name with full path
  - Algorithms called and version numbers
  - Parameters set by user in reduction script (absolute scale factor, for example)
  - Manual alterations to instrument parameters
    - Flag noting that a parameter was changed
    - Reason that it was changed (optional)
  - The actual script
- Metadata for the files used in reduction
  - Beam center file
  - Detector pixel map file
  - Sample files
  - Sample transmission files
  - Background files
  - Background transmission files
  - Dark current
  - Flood field
- Reduction operation output
  - Algorithm output
    - Console output
      - Errors

- Warnings
  - Errors highlighted independently
  - Derived parameters
    - Beam center, transmissions, etc.
    - Algorithm used
    - Notification of being over-ridden
  - Over-ride value(s)
    - Optional reason
- The reduced data
  - 1D data
  - 2D data

#### **13.4 Additional output from data reduction**

The primary output resulting from a data reduction process is the three aforementioned files. However, there are additional forms of output that can also be considered the result of the data reduction and add significant value to an experiment. The information may either be generated during the data reduction process, or after it completes using the 1D data, 2D data and the logfile. The specificity of many of the operations below require development of specific algorithms for producing the output from the information produced during data reduction. The lists below are likely to expand over time. Operations performed after the data reduction process should also be available for off-line post-processing of series of data by the user. Unless it is a 1D or 2D intensity profile, which is saved using one of the standard formats

described above, extracted information is saved in tab-separated ASCII format with a suitably descriptive filename or one provided by the user. Special files, such as a detector mask and the sensitivity file used for performing the correction, that are used during the data reduction but are not produced during it are not mentioned below.

#### *13.4.1 Additional output during the data reduction process*

- 1D data files produced from individual wavelength bins or time bins that are produced using TOF methods, such as from EQ-SANS or a HFIR SANS operating in event mode (and potential future TOF SANS instruments)
- Data files, in either 1D or 2D standard format, produced at intermediate stages of the data reduction process for diagnostic purposes.

#### *13.4.2 Output of transformations and extractions from reduced data*

- Wedge binned data:  $\langle Q, I(Q), \delta I(Q), \sigma_Q(Q) \rangle$  for each selected wedge center  $\Psi$  and width  $\psi$  using the standard 1D data format. The wedge specification is to be included in the header of the file and may be used to automatically produce the name of the file to be saved.
- Annular binned data:  $\langle \Psi, I(\Psi), \delta I(\Psi), \sigma_Q(\Psi) \rangle$  for selected ring centered at  $Q$  having a width  $\Delta Q$  using the standard 1D data format. The ring specification is to be included in the header of the file and may be used to automatically produce the name of the file to be saved.

- Line or box binned data:  $\langle Q, I(Q), \delta I(Q), \sigma_Q(Q) \rangle$  for each selected wedge center  $\Psi$  and width  $\psi$  using the standard 1D data format. The line or box specification is not included in either of the headers. Automatic file name generation is not performed for these forms of binning.
- 2D data transformed into “polar coordinates”:  $\langle |Q|, \Psi, I(Q), \delta I(Q), \sigma_Q(|Q|) \rangle$ . This is a non-standard representation of the data, but can be saved using a standard 2D data format in which  $\sigma_{Q_y} = 0.0$  to fulfill the specifications of the data file format or be omitted. The nature of the transformation can be included in the file header and may be used to automatically name the file.

#### 13.4.3 Output extracted from the log file(s) after the data reduction process

- EQ-SANS wavelength-dependent transmission  $\langle \lambda, T(\lambda), \delta T(\lambda), T_{fit}(\lambda), \delta T_{fit}(\lambda) \rangle$  in tab-separated column ASCII format. The header should include what information is in each column, and the file name may be generated automatically.
- DAS-collected metadata parameter values  $V(t)$  as a function of time (EPICS only)  $\langle t, V(t) \rangle$  in tab-separated column ASCII format. The header should include what information is in the column with units, if provided by DAS. The filename may be generated automatically.
- Metadata parameter values extracted from a series of log files. The user specifies the set of metadata to be extracted, such as temperature, ROI counts or transmission. The data are output as n-column tab-separated ASCII  $\langle \langle file \rangle, V_1, \dots, V_n \rangle$ . The result file name



is provided by the user. The header in the file should state which parameter is in each column with units, if available.

## 14. References

- [1]. Arnold, O., Bilheux, J.C., Borreguero, J.M., Buts, A., Campbell, S.I., Chapon, L., Doucet, M., Draper, N., Leal, R.F., Gigg, M.A., Lynch, V.E., Markyadsen, A., Mikkelsen, D.J., Mikkelsen, R.L., Miller, R., Palmen, K., Parker, P., Passos, G., Perring, T.G., Peterson, P.F., Ren, S., Reuter, M.A., Savici, A.T., Taylor, J.W., Taylor, R.J., Tolchenoy, R., Zhou, W., Zikowsky, J., 2014. "Mantid-Data analysis and visualization package for neutron scattering and mu SR experiments." *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Dect. Assoc. Equip.* **764**, 156–166.
- [2]. Dewhurst, C. Accessed 09/31/2018. <https://www.ill.eu/users/support-labs-infrastructure/software-scientific-tools/grasp/>
- [3]. Glinka, C. J., Barker, J. G., Hammouda, B., Krueger, S., Moyer, J. J., Orts, W. J. 1998. "The 30 m small-angle neutron scattering instruments at the National Institutes of Standards and Technology." *J. Appl. Crystallogr.* **31**, 430-445.
- [4]. He, L., Do, C., Qian, S., Wignall, G. D., Heller, W. T., Littrell, K., Smith, G. S. 2015. "Corrections for the geometric distortions of the tube detectors on SANS instruments at ORNL." *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Dect. Assoc. Equip.* **775**, 63-70.
- [5]. Wignall, G. D., Bates, F. S. 1987. "Absolute calibration of small-angle neutron scattering data." *J. Appl. Crystallogr.* **20**, 28-40.
- [6]. Mildner D. R. F., Carpenter, J. M. 1987. "Resolution of small-angle scattering with Soller collimation." *J. Appl. Crystallogr.* **20**, 419-424.

- [7]. Zhao, J. K., Gao, C. Y., Liu, D. 2010. "The extended q-range small-angle neutron scattering diffractometer at the SNS." *J. Appl. Crystallogr.* **43**, 1068-1077.
- [8]. Heller, W. T., Cuneo, M. J., Debeer-Schmitt, L. M., Do, C., He, L., Herous, L., Littrell, K. C., Pingali, S. V., Qian, S., Stanley, C. B., Urban, V. S., Wu, B., Bras, W. 2018. "The suite of small-angle neutron scattering instruments at Oak Ridge National Laboratory." *J. Appl. Crystallogr.* **51**, 242-248.

## Appendix 1. Metadata Contained in Measured Data Files

In addition to the neutron scattering data (i.e. counts in pixels), data files contain measurement-specific metadata. The metadata can be either user-specified parameters or hardware returned values, such as motor positions, which reflect the instrument configuration used for the measurement or sample temperature, which reflects the condition of the sample during the measurement. As a result, the information contained in the metadata has both value for the data reduction and for general experimental logging. Metadata contained in the data files are, and associated parameters of value to the logging or data reduction are listed in parenthesis. The metadata contained in the data files should be output into the data reduction log.

1. Start/stop of measurement time (duration of measurement)
2. Total detector count
3. Sample count rate
4. Sample thickness
5. Total beam monitor count
6. Proton charge
7. Instrument definition file
8. Source aperture size
9. Sample aperture size
10. Wavelength of neutron (velocity selector speed)
11. Wavelength spread (velocity selector tilt)
12. Sample to moderator distance

13. Sample to flange distance
14. Sample aperture to flange distance
15. Main detector to flange distance
16. SDD for the HFIR SANS, which is a sum of items 13 and 15, above
17. Main detector translational offset
18. Wing detector to sample distance
19. Wing detector angle
20. Wing detector radius
21. Wing detector rotational offset
22. Wing detector center internal offset
23. Pixel sizes along x and y direction
24. Number of pixels
25. Collimator positions (number of guides, this translates to source distance for the HFIR SANS)
26. Attenuator positions
27. Beam stop positions
28. Beam stop size
29. Flag of whether beam-stop is blocking the direct beam
30. Frequency of pulsed source
31. Choppers speed (frequency) and phase
32. DSP Trigger delay
33. Scaling factor for absolute intensity

34. Directory of "Main Detector Sensitivity Correction File"
35. Directory of "Main Detector Beam Center File"
36. Directory of "Main Detector Dark Current File"
37. Directory of "Main Detector Bar Scan File"
38. Directory of "Wing Detector Sensitivity Correction File"
39. Directory of "Wing Detector Beam Center File"
40. Directory of "Wing Detector Dark Current File"
41. Directory of "Wing Detector Bar Scan File" Experiment title
42. Experiment identifier (IPTS #)
43. Filename
44. Run number
45. Sample position
46. Sample name
47. Sample changer
48. Sample info (coming from IPTS, e.g. a Barcode, Density, Formula, etc.)
49. Temperature
50. Shutter position
51. Mask (bar) position

## Appendix 2. Determining the Pixel Locations

To properly convert the data from real space to  $Q$ -space, it is necessary to know the sizes and locations of each pixel in the detector. Presently, the SANS detectors at ORNL are arrays of 8mm diameter, meter long LPSD tubes arranged in two layers, with the detectors in the back plane nominally centered in the gaps between the detectors on the front plane, with a lateral 5.5 mm between tube centers. The known arrangement of the tubes in the array can be used to place the horizontal positions of the centers of the pixels within the detector. While the relative positions of the tubes are well-defined, the positions of the pixels are not physically defined. The positional encoding along the length of the tube is performed by charge division followed by binning into electronically-defined pixels. As a result, the positional encoding along the length of the tube can vary from tube-to-tube due as a result of manufacture and tuning of the associated electronics. The encoding is not fixed. Instead, it can change over time and can change dramatically when the low voltage power to the detectors is cycled.

There are two methods for defining where the vertical positions of the pixels lie in the space of the detector. In the first method, the length of the LPSD is simply divided by the number of pixels into which it is sub-divided, and their locations in the space of the detector are defined using this calculated size. The instrument definition files used by Mantid employ this method for defining the geometry of the instrument. This approach requires that the tune of the detectors, which can be checked by a diffraction standard, such as silver behenate, is acceptable. Mantid can use the information contained in the instrument definition file to position the pixels on the detector in real space for performing data reduction.

The second approach for establishing the vertical positions of the pixels in a given tube is to calibrate them using what is known as “the bar scan” to create a map of the position and size of the pixels in the tubes. All three SANS instruments are equipped with a neutron-absorbing bar that spans the width of the detector, which is also described below. The bar is mounted on a translation stage that moves the bar vertically and the travel covers nearly the entire length of the LPSDs. The translation stage has an absolute encoder, which provides a good measure of the relative position of the bar in front of the detector. To perform the bar scan, a series of flood-field measurements are performed with the bar at different positions. At least seven measurements are made, but more are made in practice. The bar scan measurements should be made using a sample and with an instrument configuration suitable for a flood field measurement. Additionally, a flood field measurement is used to determine the apparent width of the tubes and is combined with the known layout of the tubes in the array to complete the map of the locations of the pixels in the detector that can be used during data reduction.

## **Appendix 2.1 Measuring the vertical pixel position and size**

For a given location of the bar, the bottom edge is found. The height of the bar is not consistent along the length. The bottom edge of the bar is at a consistent height with respect to the LPSDs, but the bar is thicker at one end than the other to provide mechanical strength. The edge is found in a bar scan file  $B_i(m, n)$  in the following manner. Note that no corrections are applied to the  $B_i(m, n)$ , but bad tubes that are identified according to Section 2, above, are ignored.



1. Determine the average number of counts in a tube,  $T_{ave}$ .
2. From the bottom pixel in a tube, step up through the pixels until a pixel having a number of counts greater than  $0.2T_{ave}$  is found. If the next 3 pixels also have more counts than the threshold, then define this pixel as the “bottom” pixel of the tube,  $n_{bot}$ , and proceed to step 3, below.
3. From the top pixel in a tube, step down through the pixels until a pixel having a number of counts greater than  $0.2T_{ave}$  is found. If the next 3 pixels also have more counts than the threshold, then define this pixel as the “top” pixel of the tube,  $n_{top}$ , and proceed to step 4, below.
4. Starting from  $n_{bot} + 4$ , step up until the counts in a pixel is less than  $0.3T_{ave}$ . The pixel that satisfies this condition is the position of the bottom edge of the bar in tube  $m$  and is retained as  $n_{m,i}$  if the counts in the pixel at  $n_{m,i} + 4$  is also less than  $0.3T_{ave}$ .
5. The real space y-coordinate of this position of the bar is given by  $y_{m,i} = Y_{bar} + 565mm$ , which has been calibrated for the motor that move the bar. The bottom edge of the bar has been verified to be level.
6. The  $m$  arrays of  $n_{m,i}(y_{m,i})$  are then fit using a 5<sup>th</sup> order polynomial shown in Equation A2.1 to obtain the coefficients  $A_{m,j}$ , which are also used to provide the y-widths of the pixels according to Equation A2.2.

$$y_m(n) = A_0 + A_1n + A_2n^2 + A_3n^3 + A_4n^4 + A_5n^5 \quad (A2.1)$$

$$w_{y,m}(n) = A_1 + 2A_2n + 3A_3n^2 + 4A_4n^3 + 5A_5n^4 \quad (A2.2)$$

## Appendix 2.2 Measuring the apparent tube width and defining its position

Once the vertical positions of the pixels and their heights have been determined as described in Appendix 2.1, the apparent widths of the tubes must be determined. The location of the tubes relative to one another in the array is already predetermined. The term “apparent width” refers to the width that is most efficient for detecting neutrons. The measurement requires a flood field measurement,  $I(x, y)$ , that is collected at a long sample-to-detector distance and a mask,  $M(x, y)$ , that covers the beam stop, the surrounding region that may be impacted by parasitic scattering, and the top and bottom edges of the detector tubes where their diminished response is observed. The measurements of the pixel widths also require the pixel heights,  $H(x, y)$ . The counts per length in a tube  $x$ , which is referred to as  $C(x)$ , is determined according to Equation A2.3.

$$C(x) = \sum_{(x,y) \notin M(x,y)} \left[ \frac{I(x,y)}{H(x,y)} \right] \quad (\text{A2.3})$$

Then, if  $N_{tubes}$  is the number of tubes in the detector array, the average counts per unit length,  $C_{ave}$ , is calculated using Equation A2.4. The average values in the front and back banks of tubes,  $C_{front}$  and  $C_{back}$ , are determined according to Equations A2.5 and A2.6.

$$C_{ave} = \frac{\sum N_{tubes} C(x)}{N_{tubes}} \quad (\text{A2.4})$$

$$C_{front} = \frac{\sum_{N_{tubes} \in front} C(x)}{N_{tubes, front}} \quad (\text{A2.5})$$

$$C_{back} = \frac{\sum_{N_{tubes} \in back} C(x)}{N_{tubes, back}} \quad (\text{A2.6})$$

In Equations A2.5 and A2.6,  $N_{tubes, front}$  and  $N_{tubes, back}$  are the number of detector tubes in the front and back banks of the detector, respectively.

The apparent widths of the tubes are calculated for the front and back banks of detector tubes, resulting in only two values for the apparent widths, namely  $W_{front}$  and  $W_{back}$ , that are calculated according to Equations A2.7 and A2.8. The detector arrays are manufactured such that the spacing of the tubes in the plane of a given 8-pack of tubes is 5.5 mm. The nominal width,  $W_{nominal}$ , is set to this value.

$$W_{front} = \frac{c_{front}}{c_{ave}} W_{nominal} \quad (A2.7)$$

$$W_{back} = \frac{c_{back}}{c_{ave}} W_{nominal} \quad (A2.8)$$

## **Appendix 3. Preparation of a Detector Sensitivity File**

Detector tubes and individual pixels do not respond in an identical manner. The tube-to-tube variation in response, as well as the variation between 8-packs in the LPSD array are the largest source of differences in detector response, which is called sensitivity. To correct for this variation, a detector sensitivity file is prepared for use during data reduction. The process is described below.

### **Appendix 3.1 Obtaining the Flood Field Data for the Detector Sensitivity File**

The sensitivity correction involves using the results of a measurement or set of measurements, called a flood field. The flood field is collected from a sample that scatters uniformly in angle to ensure that every pixel on the detector has the potential to see the same number of neutrons. Different materials are currently used at the beamlines to measure the flood field. At present, the standard samples used for measuring the flood field are H<sub>2</sub>O in a 1 mm path length cell for GP-SANS and Bio-SANS, and a 1 mm sheet of PMMA for EQ-SANS.

The best material to use for the flood field measurement is currently being investigated by the Bio-SANS team. The investigation was prompted by distortions observed when reducing data from a sample containing mostly D<sub>2</sub>O by using a sensitivity calibration derived from a flood field measurement of H<sub>2</sub>O. The distortions exist in both the main and wing detectors and are therefore not likely to arise from the high- $q$  water peak. Bio-SANS studies indicate that the distortions result from inelastic scattering events in the sample that result in higher energy (lower wavelength) neutrons. The LPSDs that make up the detector are less efficient for higher energy neutrons, which creates distortions in the flood field data, and therefore in the

sensitivity correction. Two possible approaches to address these distortions are being investigated. One possibility is to use a thinner ( $< 1$  mm) piece of PMMA to minimize inelastic events in the flood field sample. Another possibility is to match the hydrogen content of the material used for the sensitivity measurement with that of the samples being measured in order to best match the level of incoherent and inelastic scattering, thereby matching the detector response during the sample measurement to that of the flood field. The matter remains under investigation.

The following configurations are measured on the three beamlines during instrument calibration. At least one configuration is measured that has minimal distortions due to the differences in the solid angle subtended by pixels at the largest scattering angles relative to the smallest ones sampled. For the measurements, the brightest wavelength, or band of wavelengths, is used for the sake of minimizing measurement time. The thin samples necessitate the use of the beam stop. Data collection times that are used provide 1000 counts per pixel on average, which provides counting statistics of  $\sim 3\%$ .

<b>Instrument</b>	<b>SDD (and wing angle)</b>
EQ-SANS	4m (used for $\geq 4$ m), 2.5m, 1.3m
GP-SANS	6.8m
Bio-SANS	7m (main) and 1.0, 1.4, 3.2, 12.2° (wing)

The SANS at HFIR have the ability to translate the main detector perpendicular to the beam, while the EQ-SANS does not. As a result, the HFIR SANS may measure multiple flood field data sets using different detector offsets, which makes it possible to patch over the gap in the data due to the beam stop. EQ-SANS cannot translate the detector sufficiently far to make

this possible. In either case, accounting for the gap in the scattering pattern that results is addressed during processing into a sensitivity file, described below. Additional measurements, such as a transmission, a background and its associated transmission are also required for processing a sensitivity file and are described below. In the case of the GP-SANS, there are the three flood fields measurements with transmission measurements and three associated backgrounds with transmission measurements. Three additional empty beam reference measurements for the transmissions are also required, making for a total of fifteen data sets. Although the specific measurements differ, EQ-SANS also requires fifteen data sets.

## **Appendix 3.2 Processing the Flood Field Data into a Sensitivity File**

Prior to use during data reduction, the flood field data sets must be processed and saved as a sensitivity file. In fact, all of the operations up to the binning into  $Q$  are very much like a typical data reduction, with specific exceptions being noted in the descriptions below.

### *Appendix 3.2.1 Read the file and prepare the flood field and background data*

The flood field data and the background files are read in as described in Section 2 and 3, above. The transmission measurements and the empty beam reference are also read in the same manner. The beam center is determined, the coordinate system of the instrument is established, and the solid angle correction is applied. Next, the mask is applied to the data, as described in Section 4, above. When multiple files are being used to create the sensitivity for GP-SANS and Bio-SANS, each file will have a mask for the beam stop. The masks will also include the top and bottom edges of the detector. The sensitivity is not applied (Section 5,

above) because the goal of this process is to determine the sensitivity that must be applied. Then, the flood field data and the background data are normalized using the desired method (Section 6, above). Next, the sample and background transmissions are determined and the scattering data are corrected for it (Section 7, above). The background is then subtracted from the scattering data, as described in Section 8, above.

At this point, the process deviates from a typical data reduction. If the data was collected using GP-SANS or Bio-SANS, it is ready to be used to prepare the detector sensitivity file. However, flood field data collected on EQ-SANS requires one additional step prior to use. The wavelength bins are summed together with appropriate propagation of the uncertainty to remove the time-of-flight nature from the data according to Equation A3.1 and A3.2. A 2D data set results that is similar to that produced by GP-SANS and Bio-SANS.

$$I(m, n) = \sum_{\lambda_i} I(m, n, \lambda_i) \quad (\text{A3.1})$$

$$\delta I(m, n) = \left[ \sum_{\lambda_i} (\delta I(m, n, \lambda_i))^2 \right]^{1/2} \quad (\text{A3.2})$$

Finally, a mask is drawn on the result and applied to the data. The purpose of this mask is to remove regions of the detector, such as the top and bottom edges or the area around the beamstop, from the calculation of the sensitivity. Pixels that are masked at this point in the process may be “patched”, as described below.

### *Appendix 3.2.2 Calculate the average and normalize*

The next step in the process is to determine the average value per pixel. This is done for as many flood fields as exist, when patching by means of a series of detector offsets is

employed. If there are  $N$  pixels in the patched pattern that are not masked out with mask  $M(m, n)$ , then the average and error in the average are provided by Equations A3.3 and A3.4.

$$F = \frac{\sum_{m,n} [I(m,n) \notin M(m,n) \vee \neq -\text{inf}, \text{NaN}]}{N} \quad (\text{A3.3})$$

$$\delta F = \frac{(\sum_{m,n} [\delta I(m,n) \notin M(m,n) \vee \neq -\text{inf}, \text{NaN}]^2)^{1/2}}{N} \quad (\text{A3.4})$$

Then, the first cut at the sensitivity  $S_1$  and its uncertainty  $\delta S_1$  are produced using Equations A3.5 and A3.6.

$$S_1(m, n) = \frac{I(m, n)}{F} \quad (\text{A3.5})$$

$$\delta S_1(m, n) = \frac{I(m, n)}{F} \left[ \left( \frac{\delta I(m, n)}{I(m, n)} \right)^2 + \left( \frac{\delta F}{F} \right)^2 \right]^{1/2} \quad (\text{A3.6})$$

### *Appendix 3.2.3 Apply a bad pixel threshold*

In the event that a pixel responds considerably different than the average, which suggests that it is not performing correctly, the pixel can be removed from use during data reduction. Such pixels are identified by applying thresholds to  $S_1(m, n)$ . The minimum and maximum thresholds,  $S_{min}$  and  $S_{max}$ , respectively, are values between 0.0 and 2.0, but have historically varied depending on the beamline. For example, default values are  $S_{min} = 0.5$  and  $S_{max} = 2.0$ , but other values are employed. Each pixel is checked against these threshold values. If it lies within the range, it is retained. If not, the value is set to “-inf” or “NaN” to flag it as locally masked, but it is treated differently than masked pixels are during the remainder of the sensitivity file preparation. These “bad” pixels must also be taken into account as part of the mask applied during reduction (section 4, above) and are therefore clearly identified in the sensitivity file that is written for use in data reduction.



#### Appendix 3.2.4 Patch the sensitivity

In order to fill in the missing information in the sensitivity that results from the beam stop, the masked regions must be patched. There are two approaches for doing so. The first is to make a series of measurements with different detector offsets, as is noted above. The second approach to patching the beam stop mask is to create the missing values by means of a functional fit to the unmasked portions of the tube and interpolating the missing values. Note that no patching can be performed in cases where an entire tube has been masked out. Such masked tubes will be part of the primary mask defined during the data reduction, so the results of the data reduction are not affected. Regardless of the method employed, if a pixel was marked as non-performing (i.e.  $S_1(m, n) = -\text{inf or NaN}$ ), it is ignored during the calculations performed during the patching and its value is not overwritten.

##### Appendix 3.2.4.1 Patching by merging multiple sensitivity measurements

If there is a set of  $k$  sensitivity measurements that have been processed through section Appendix 3.2.3, above, to provide  $S_{1,k}(m, n)$ , and the masks that remove the beam stop are  $M_k(m, n)$ , then it is possible to use the set to produce an average sensitivity that does not have any gaps due to the beam stop. The number of sensitivity measurements for which there is good information in the sensitivity measurements, i.e. not masked or marked as bad as described above, is  $N_G(m, n)$ . Then, the patched sensitivity  $S_2(m, n)$  is given by Equations A3.7 and A3.8.

$$S_2(m, n) = \frac{\sum_k [S_{1,k}(m, n) \notin M_k(m, n) \vee \neq -\text{inf, NaN}]}{N_G(m, n)} \quad (\text{A3.7})$$

$$\delta S_2(m, n) = \frac{\left( \sum_k [\delta S_{1,k}(m, n) \notin M_k(m, n) \vee \neq -inf, NaN]^2 \right)^{1/2}}{N_G(m, n)} \quad (A3.8)$$

#### Appendix 3.2.4.2 Patching by functional interpolation

When it is not possible to perform a series of flood field measurements using different offsets of the detector, the beam stop mask in the sensitivity file can be patched by means of a functional fit to the intensity along a tube and then interpolating the value for the masked pixels. In practice, all masked pixels in all tubes are patched. The only pixels left unaffected are those marked as non-performing (i.e.  $S_1(m, n) = -inf$  or  $NaN$ ). Fitting is performed with either a first or second order polynomial, which is selected by the user creating the sensitivity, and the values for the masked pixels in tube  $m$  are given by Equation A3.9. The uncertainties in the coefficients of the polynomial are used to estimate the uncertainties in the sensitivity for the pixels in tube  $m$  using Equation A3.10.

$$S_1(m, n) = A_0 + A_1 m + A_2 m^2 \quad (A3.9)$$

$$\delta S_1(m, n) = \left( \delta A_0^2 + (\delta A_1 m)^2 + (\delta A_2 m^2)^2 \right)^{1/2} \quad (A3.10)$$

#### Appendix 3.2.5 Perform final normalization and write the sensitivity file

The final step in the process is to determine the average value per pixel of the patched pattern according to Equations A3.11 and A3.12. If there are  $N_2$  pixels in the patched pattern that are not masked out as a result of having their response be outside the specified range,

$$F_2 = \frac{\sum_{m,n} [S_2(m, n) \notin M(m, n) \vee \neq -inf]}{N_2} \quad (A3.11)$$

$$\delta F_2 = \frac{(\sum_{m,n} [\delta S_2(m,n) \notin M(m,n) \vee \neq -inf]^2)^{1/2}}{N_2} \quad (\text{A3.12})$$

Then, the final sensitivity  $S$  and its uncertainty  $\delta S$  are produced using Equations A3.13 and A3.14.

$$S(m, n) = \frac{S_2(m, n)}{F_2} \quad (\text{A3.13})$$

$$\delta S(m, n) = \frac{S_2(m, n)}{F_2} \left[ \left( \frac{\delta S_2(m, n)}{S_2(m, n)} \right)^2 + \left( \frac{\delta F_2}{F_2} \right)^2 \right]^{1/2} \quad (\text{A3.14})$$

### 3.2.6 Write the sensitivity file

After the final sensitivity is produced, the file must be output in a manner that is easily read by Mantid. The standard ASCII format for 2D data, without the columns providing  $\sigma_{Q_x}$  and  $\sigma_{Q_y}$ , to save  $\langle m, n, S(m, n), \delta S(m, n) \rangle$  is suitable and it is well-suited to direct inspection of specific values in the sensitivity for diagnostic purposes. The pixels identified as performing outside the desire response range are included in the file with  $S(m, n) = -inf$ . A dedicated reader will be required for importing a sensitivity file into the reduction that also identifies the non-performing pixels that should be appended to the masks that are also loaded for use in the data reduction.

## Appendix 4. Derivation of the Q-resolution

The instrument resolution function is required for data analysis because model profiles need to be convoluted with it before being compared against the data. The SANS resolution function has two contributions that result from the instrument geometry and the wavelength distribution, as shown in Equation A4.1, which does not employ the small-angle approximation. For convenience, the wavelength dependence of the geometric contribution is only shown explicitly in Equation A4.1 and as required when showing the effect of gravity later in this Appendix.

$$\sigma_Q^2(\lambda) = [\sigma_Q^2(\lambda)]_{geo} + [\sigma_Q^2(\lambda)]_{wave} = \frac{(4*\pi*cos\theta)^2}{\lambda^2} \sigma_\theta^2(\lambda) + Q^2 \frac{\sigma_\lambda^2(\lambda)}{\lambda^2} \quad A4.1$$

### Appendix 4.1. The instrument geometry contribution

The instrument geometric contribution to the resolution function can be expressed in terms of the radial variance  $\sigma_r$  and the sample-to-detector distance  $L_2$  according to Equation A4.2.

$$\sigma_\theta^2 = \frac{\sigma_r^2 \cos^4 2\theta}{4L_2^2} \quad (A4.2)$$

Rather than working in 1D, the variances in the x-direction and y-direction,  $\sigma_x$  and  $\sigma_y$ , respectively, are treated independently, which provides smearing information for modeling data in 2D and makes it possible to account for the effects of gravity on the resolution. First, only the geometric effects are considered, which results in the variances having the same form. Therefore, only  $\sigma_x$  is presented below. Deriving it begins with Equation A4.3.

$$\sigma_x^2 = \left(\frac{L_2}{L_1}\right)^2 \langle x^2 \rangle_1 + \left(\frac{L_1+L_2}{L_1}\right)^2 \langle x^2 \rangle_2 + \langle x^2 \rangle_3 \quad (\text{A4.3})$$

Here,  $L_1$  is the source-to-sample distance,  $\langle x^2 \rangle_1$  denotes an average over the source aperture,  $\langle x^2 \rangle_2$  is the average over the sample aperture, and  $\langle x^2 \rangle_3$  is the average over a detector pixel. Assuming that circular source and sample apertures of radii  $R_1$  and  $R_2$  and referring the user to the work of Mildner and Carpenter for the origin of the  $L_2/L_1$  and  $(L_1 + L_2)/L_1$  factors [6], the value of  $\langle x^2 \rangle_1$  is calculated as shown in Equation A4.4.

$$\langle x^2 \rangle_1 = \frac{\int_0^{2\pi} \int_0^{R_1} r^2 \cos^2 \varphi r dr d\varphi}{\int_0^{2\pi} \int_0^{R_1} r dr d\varphi} = \frac{\int_0^{R_1} r^3 dr \int_0^{2\pi} \cos^2 \varphi d\varphi}{\int_0^{2\pi} \int_0^{R_1} r dr d\varphi} = \frac{R_1^2}{4} \quad (\text{A4.4})$$

Similarly for the sample aperture,  $\langle x^2 \rangle_2 = R_2^2/4$ . The pixels of the detector are rectangular with horizontal and vertical dimensions  $\Delta x_3$  and  $\Delta y_3$ , respectively, so the average takes the form shown in Equation A4.5.

$$\langle x^2 \rangle_3 = \frac{\int_{-\frac{\Delta x_3}{2}}^{\frac{\Delta x_3}{2}} x^2 dx}{\int_{-\frac{\Delta x_3}{2}}^{\frac{\Delta x_3}{2}} dx} = \frac{\Delta x_3^2}{12} = \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \quad (\text{A4.5})$$

To compute the detector pixel contribution to the resolution function, we must average of the rectangular detector cell of sides  $\Delta x_3$  and  $\Delta y_3$ , as shown in Equation A4.6.

$$\langle x^2 \rangle_3 = \frac{\int_{-\frac{\Delta x_3}{2}}^{\frac{\Delta x_3}{2}} x^2 dx}{\int_{-\frac{\Delta x_3}{2}}^{\frac{\Delta x_3}{2}} dx} = \frac{\Delta x_3^2}{12} = \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \quad (\text{A4.6})$$

Assembling the results of the calculations above,  $\sigma_{Qx}$  is given by Equation A4.7.

$$[\sigma_{Qx}^2]_{geo} = \left( \frac{2\pi \cos \theta \cos^2 2\theta}{\lambda L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \right] \quad (\text{A4.7})$$

The form for  $[\sigma_{Qy}^2]_{geo}$ , being the geometric contribution to the y-direction resolution function, follows directly by substituting  $\Delta y_3$  for  $\Delta x_3$ .

The contribution of gravity to the resolution function of a SANS instrument,  $[\sigma_{Qy}]_{grav}$  can be added to that of the geometric contribution to give  $[\sigma_{Qy}^2]_{geo+grav} = [\sigma_{Qy}^2]_{geo} + [\sigma_{Qy}^2]_{grav}$ . Gravity has its largest effect data collected with the GP-SANS and Bio-SANS instruments, which use a velocity selector to define the wavelength. Let the wavelength  $\lambda$  have a distribution provided by the velocity selectors with relative width  $\Delta\lambda/\lambda$ . During the course of traveling from the source to the detector, being  $L_1 + L_2$ , a neutron will fall the distance shown in Equation A4.8.

$$\Delta y = \lambda^2 g m_N^2 L_2 (L_1 + L_2) / 2h^2 \quad (A4.8)$$

where  $g$  is the gravity of earth (or small  $g$ , 9.8 m/s<sup>2</sup>),  $h$  is Planck's constant (6.63x10<sup>-34</sup> m<sup>2</sup>kg/s), and  $m_N$  is the mass of the neutron (1.674927471x10<sup>-27</sup> kg). Defining a parameter  $B = g m_N^2 L_2 (L_1 + L_2) / 2h^2$ , the variance in  $\Delta y$  arises from the wavelength variance and results in  $[\sigma_{Qy}]_{grav}$  having the form shown in Equation A4.9.

$$[\sigma_y^2]_{grav} = B^2 (\langle \lambda^4 \rangle - \langle \lambda^2 \rangle^2) \quad (A4.9)$$

The double averages of over the triangular wavelength distributions are shown in Equations A4.10 and A4.11.

$$\langle \lambda^2 \rangle = \lambda^2 \left[ 1 + \frac{1}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right] \quad (A4.10)$$

$$\langle \lambda^4 \rangle = \lambda^4 \left[ 1 + \left( \frac{\Delta\lambda}{\lambda} \right)^2 + \frac{1}{15} \left( \frac{\Delta\lambda}{\lambda} \right)^4 \right] \quad (A4.11)$$

In practice, the fourth order term in Equation A4.11 is not used for calculating  $[\sigma_{Qy}]_{grav}$ , which results in the fully-assembled terms of  $\sigma_{Qy}$  that do not result from the wavelength distribution being given by Equation A4.12.

$$[\sigma_{Qy}^2]_{geo+grav} = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta y}{2} \right)^2 + B^2 \lambda^4 \frac{2}{3} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right] \quad (A4.12)$$

#### Appendix 4.2. Contribution of the wavelength distribution

For reactor-based SANS a velocity selector is how the beam is monochromatic. This means the velocity selector geometry determines the wavelength uncertainty by Equation A4.13.

$$\frac{\Delta\lambda}{\lambda} = \frac{d}{L \tan(\alpha)} \quad (A4.13)$$

Here  $d$  is the slit opening,  $L$  is the length of the selector, and  $\alpha$  is the tilt angle of the slit. Velocity selectors produce neutrons having a triangular distribution that peaks at  $\lambda$  and has a full-width at half maximum (FWHM) of  $\Delta\lambda$ . As a result, Equation A4.14 provides  $\langle\lambda^2\rangle$ .

$$\langle\lambda^2\rangle = \lambda^2 \left[ 1 + \frac{1}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right] \quad (A4.14)$$

Then, the wavelength variance is given by Equation A4.15.

$$\frac{\sigma_\lambda^2}{\lambda^2} = \left( \frac{\langle\lambda^2\rangle - \langle\lambda\rangle^2}{\lambda^2} \right) = \frac{1}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \quad (A4.15)$$

As a result, the wavelength spread contributions to the GP-SANS and Bio-SANS resolution functions are given by Equations A4.16 and A4.17.

$$[\sigma_{qx}^2]_{wave} = q_x^2 \left( \frac{1}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right) \quad (A4.16)$$

$$[\sigma_{Qy}^2]_{wave} = q_y^2 \left( \frac{1}{6} \left( \frac{\Delta\lambda}{\lambda} \right)^2 \right) \quad (A4.17)$$

In contrast to a reactor-based SANS, the wavelength resolution on a spallation TOF SANS instrument arises from the uncertainty in the emission time from the moderator and the size of the wavelength bins used during the data reduction, which is done for the sake of speed. The x and y components of the resolution function are calculated in the same manner as it is for a velocity selector above, but the relative wavelength variance  $\sigma_\lambda/\lambda$  is given by Equation A4.18.

$$\left(\frac{\sigma_\lambda}{\lambda}\right)^2 = \left(\frac{3.9560 \sigma_{em}(\lambda)}{1000 \lambda (s2p+m2s)}\right)^2 + \left(\frac{\Delta\lambda_{bin}}{\lambda_{bin}}\right)^2 \quad (\text{A4.18})$$

In Equation A4.18,  $s2p$  is the sample-to-pixel distance (in meters), while  $m2s$  is the moderator to sample distance (in meters).  $\sigma_{em}(\lambda)$  is the RMSD of the uncertainty in the emission time of a neutron from the moderator that was obtained by simulation and was provided by Erik Iverson. The simulation is consistent with measurements of  $\sigma_{em}(\lambda)$  for the cold moderator at SNS that serves the other side of the instrument hall. The value used during the data reduction is a sixth-order polynomial fit to the simulation results. The width of the binning currently employed on EQ-SANS, which factors into the  $\sigma_{bin}$  term, is described in Section 3, above. The factor of 1/6, which results from the shape of the distribution of wavelengths provided by a velocity selector, becomes 1/12 because square bins and the RMSD of the uncertainty in the time-of-flight are used.

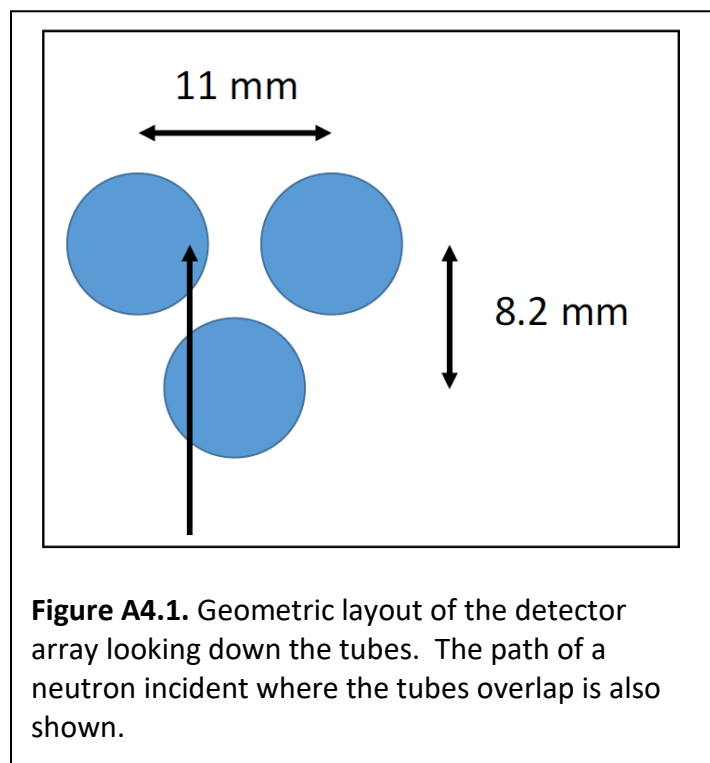
#### **Appendix 4.3. Comments regarding $\Delta x_3$ and $\Delta y_3$**

There is a discrepancy regarding the size of the pixel used in the data reduction for calculating the resolution function information between the instruments at HFIR and the EQ-SANS that was, at least at one point in time, implemented in the original Mantid code. The HFIR instruments have always used the nominal pixel size, or that determined from the bar



scan calibration as described in Appendix 2. In contrast, the EQ-SANS and Bio-SANS team investigated the possibility that the uncertainty in the position of a detected event needs to be accounted for in determining the resolution function. The effect of the positional encoding uncertainty for the purposes of calculating the resolution function is used during EQ-SANS data reduction.

A detected neutron has an inherent position uncertainty within a given tube, which is a well-known behavior for gas detectors. In the original instrument paper [7], the FWHM of a detected event was estimated to be 7 mm along the length of the tube, which is in the y-direction of the instrument geometry. The nominal pixel size used to determine  $Q$  is  $\sim 4.3$  mm,



and the latter value is employed in the calculation of the resolution function in the original implementation of the Mantid code. The measurements made by the detector group suggest that 7 mm is a more appropriate value to use for  $\Delta y_3$  when calculating  $\sigma_Q$ .

In the case of the x-direction, the detector geometry provides a contiguous gas volume by having 2

planes of tubes that overlap, shown schematically in Figure A4.1. The overlap makes it possible for a neutron incident at a scattering angle where tubes overlap to be detected in either plane of detectors. Measurements using a mask having 1 mm-wide slits verified that neutrons

incident in the region of the gas overlap can be detected in both the front and back banks of tubes. The result for a detected neutron is a positional uncertainty of  $\pm 1$  tube, which equates to an effective uncertainty of 11 mm in the x-position. The following values are used for the resolution function calculation for EQ-SANS.

$$\Delta x_3 = 11 \text{ mm}$$

$$\Delta y_3 = 7 \text{ mm}$$

#### Appendix 4.4. The full resolution function formulas

Taken together, the full equations for  $\sigma_{Qx}$  and  $\sigma_{Qy}$  for the GP-SANS and Bio-SANS are given by Equations A4.19 and A4.20, which are also presented in Section 9, above.

$$\sigma_{Qx}^2 = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \right] + \frac{Q_x^2}{6} \left( \frac{\Delta \lambda}{\lambda} \right)^2 \quad (\text{A4.19})$$

$$\sigma_{Qy}^2 = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta y_3}{2} \right)^2 + B^2 \lambda^4 \frac{2}{3} \left( \frac{\Delta \lambda}{\lambda} \right)^2 \right] + \frac{Q_y^2}{6} \left( \frac{\Delta \lambda}{\lambda} \right)^2 \quad (\text{A4.20})$$

Similarly for EQ-SANS, for  $\sigma_{Qx}(\lambda_i)$  and  $\sigma_{Qy}(\lambda_i)$  are given by Equations A4.21 and A4.22.

$$\sigma_{Qx}^2(\lambda_i) = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta x_3}{2} \right)^2 \right] + \frac{Q_x^2}{12} \left[ \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 + \left( \frac{3.9560 \sigma_{em}(\lambda)}{1000 \lambda_i (s2p+m2s)} \right)^2 \right] \quad (\text{A4.21})$$

$$\sigma_{Qy}^2(\lambda_i) = \left( \frac{2\pi \cos\theta}{\lambda} \frac{\cos^2 2\theta}{L_2} \right)^2 \left[ \left( \frac{L_2}{L_1} \right)^2 \frac{R_1^2}{4} + \left( \frac{L_1+L_2}{L_1} \right)^2 \frac{R_2^2}{4} + \frac{1}{3} \left( \frac{\Delta y_3}{2} \right)^2 + B^2 \lambda_i^4 \frac{2}{3} \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 \right] + \frac{Q_y^2}{12} \left[ \left( \frac{\Delta \lambda_i}{\lambda_i} \right)^2 + \left( \frac{3.9560 \sigma_{em}(\lambda_i)}{1000 \lambda_i (s2p+m2s)} \right)^2 \right] \quad (\text{A4.22})$$