Robust approaches for model-free small-angle scattering data analysis

Philipp Bender,a Dirk Honecker,b Mathias Bersweiler,c Rocio Costo,d Tamara Kahmann,e Frank Ludwig,e Jon Leiner,a and Johanna K. Jochum*a

a Heinz Maier-Leibnitz Zentrum (MLZ), Technische Universität München, D-85748 Garching, Germany, b ISIS Neutron and Muon Facility, Rutherford Appleton Laboratory, Chilton, OX11 0QX, United Kingdom, c Physics and Materials Science Research Unit, University of Luxembourg, 162A avenue de la Faïencerie, L-1511 Luxembourg, Grand Duchy of Luxembourg, d Instituto de Ciencia de Materiales de Madrid, ICMM/CSIC, C/Sor Juana Inés de la Cruz 3, 28049 Madrid, Spain, and *Institute of Electrical Measurement Science and Fundamental Electrical Engineering and Laboratory for Emerging Nanometrology (LENA), Technische Universität Braunschweig, 38106 Braunschweig, Germany. *Correspondence e-mail: johanna.jochum@frm2.tum.de

The small-angle neutron scattering data of nanostructured magnetic samples contain information regarding their chemical and magnetic properties. Often, the first step to access characteristic magnetic and structural length scales is a model-free investigation. However, due to measurement uncertainties and a restricted $q$ range, a direct Fourier transform usually fails and results in ambiguous distributions. To circumvent these problems, different methods have been introduced to derive regularized, more stable correlation functions, with the indirect Fourier transform being the most prominent approach. Here, the indirect Fourier transform is compared with the singular value decomposition and an iterative algorithm. These approaches are used to determine the correlation function from magnetic small-angle neutron scattering data of a powder sample of iron oxide nanoparticles; it is shown that with all three methods, in principle, the same correlation function can be derived. Each method has certain advantages and disadvantages, and thus the recommendation is to combine these three approaches to obtain robust results.

1. Introduction

Small-angle neutron scattering (SANS) probes chemical and magnetic structure on the mesoscale (1–500 nm) (Jeffries et al., 2021), which makes SANS an ideal tool to investigate nanostructured magnetic materials such as bulk ferromagnets or magnetic nanoparticle systems (Mühlbauer et al., 2019). In SANS data analysis, it is good practice to perform a Fourier transform to obtain starting parameters for the characteristic magnetic and structural length scales that are relevant for a system (Feigin & Svergun, 1987).

In the case of pure nuclear scattering the extracted correlation function corresponds to the autocorrelation function of the scattering length density profile (Li et al., 2016). This is not the case for magnetic neutron scattering due to the anisotropic nature of the dipole–dipole interaction (Mettus & Michels, 2015). However, the derived correlation functions still contain important information that reflects the real-space magnetization over the mesoscale (Bender et al., 2021). Thus, the Fourier transform of reciprocal SANS data is an easy and straightforward approach to obtain model-independent information regarding the chemical and magnetic nanostructure of the sample.

Real experimental data usually have measurement uncertainties and a restricted $q$ range, which can lead to ambiguous correlation functions when performing a direct Fourier

---

*This article is part of a virtual special issue on Magnetic small-angle neutron scattering – from nanoscale magnetism to long-range magnetic structures.

Keywords: small-angle scattering; correlation functions; Fourier transform; magnetic nanoparticles; modulation of intensity with zero effort; MIEZE; RESEDA.
transform of the data. To circumvent these issues the indirect Fourier transform (IFT) was introduced in the 1970s (Glatter, 1977). In this case, the correlation function is essentially obtained by a fit of the experimental data using a Tikhonov regularization to force smooth distributions. To find the optimal value for the smoothing degree a Bayesian analysis can be applied (Hansen, 2000). This approach allows for the inclusion of the maximum size of the scattering, which defines the size range for which the correlation function is computed, as fit parameter (Hansen, 2012). Such a model-free analysis of scattering data has proven to be a powerful approach to study several categories of materials, e.g. polymer solutions (Hammond, 2010), protein conformation (Sanchez-Fernandez et al., 2017) and the structure of colloidal particles (Fritz et al., 2000).

The IFT can further be used to derive the 2D correlation functions from the complete scattering pattern (Fritz-Popovski, 2013). However, this approach is computationally intensive for large data sets. Therefore, two faster numerical approaches were recently introduced to determine the 2D correlation functions from scattering patterns, namely the singular value decomposition (SVD) (Bender et al., 2019) and the iterative Kaczmarz algorithm (KA) (Bender et al., 2021).

Here, we analyze 1D scattering data and apply all three model-free approaches, namely the IFT, SVD and KA. We use these methods to determine the correlation functions of a powder sample of magnetic iron oxide nanoparticles from the magnetic-field-dependent SANS intensities measured at the multipurpose neutron spin-echo spectrometer RESEDA in modulation of intensity with zero effort (MIEZE)-SANS mode (Franz, Säubert et al., 2019; Franz, Soltwedel et al., 2019; Jochum et al., 2019).

2. Methods

2.1. Pre-characterization of the sample

The sample was a dense powder of ~200 mg freeze-dried iron oxide nanoparticles. The spherical single-crystalline particles were synthesized by thermal decomposition and had an average diameter of around 10 nm with a very narrow size distribution (σ < 0.1). Fig. 1(a) shows a transmission electron microscopy (TEM) image of the nanoparticles.

The macroscopic magnetic properties of the powder at 300 K were determined by quasistatic direct-current magnetometry (DCM) and alternating-current susceptibility (ACS) [see Figs. 1(b) and 1(c)]. The DCM measurement, measured from 6 T → −6 T → 6 T, exhibits a Langevin-type magnetization behavior with vanishing coercivity and remanence, indicating a superparamagnetic behavior of the sample. This is verified by the ACS measurement whose imaginary part is close to zero at low frequencies (Ludwig et al., 2017). This means that the magnetic moments of the nanoparticles can freely follow the external magnetic field via Néel-type relaxation. With increasing frequency the imaginary part slightly increases and the real part decreases accordingly. However, within the accessible frequency range (i.e. 10–10⁶ Hz) no relaxation peak in the imaginary part is observed, which indicates that the characteristic relaxation times of the particles are significantly below τ = 1/ω = 1/(2πf) < 1.6 × 10⁻¹⁰ s.

2.2. Magnetic SANS measurements

For the SANS measurements the particle powder was placed into a quartz glass cuvette with an optical path length of 1 mm. The scattering patterns were measured with the neutron spin-echo spectrometer RESEDA using the MIEZE-SANS arm (Franz, Soltwedel et al., 2019). The 2D detector was positioned off-center of the direct neutron beam as shown in Fig. 2. The neutron wavelength was 6 Å with a wavelength spread Δλ/λ of 11.7%. The experimental setup included a
magnet with the field applied in the horizontal plane. From the scattering patterns we determined the sector average in the horizontal direction, i.e. along the magnetic field direction, with a sector width of ±10°. In the following, we use three different approaches to derive the underlying correlation functions from the 1D sector averages.

3. Data analysis
3.1. Introduction of the three model-free approaches

From the 1D scattering intensity shown in Fig. 2, the underlying correlation function \( P(r) \) was determined by three different approaches. First the IFT was applied. In this case the \( N \)-dimensional vector \( P(r) \) is determined by minimizing the functional

\[
\frac{1}{2\sigma^2} \left\| AP(r) - I(q) \right\|^2 + \alpha \left\| LP(r) \right\|^2. \tag{1}
\]

Here, \( \sigma = \sigma(q) \) is the standard deviation of each data point and \( I(q) \) is the measured scattering intensity with \( M \) data points. The matrix \( A \) in equation (1) is the \( M \times N \) data transfer matrix with \( A_{ij} = 4\pi i [(\sin(q,jr))]/(q,jr) \Delta r_j \). The matrix \( L \) is an \( N \times N \) regularization matrix, which is multiplied with the regularization parameter \( \alpha \). To penalize oscillations within the extracted distributions, the non-singular approximation of the discrete second-order derivative operator can be used for \( L \):

\[
L = \frac{1}{2} \begin{pmatrix}
2 & 0 & \cdots & 0 \\
-1 & 2 & \cdots & 0 \\
& \ddots & \ddots & \ddots \\
& & -1 & 2 & -1 \\
& & & 0 & 2
\end{pmatrix}.
\tag{2}
\]

For numerical computation, equation (1) is inconvenient and the least-squares solution of

\[
\left( \frac{1}{\sigma^2} A^T \right) P(r) = \frac{1}{\sigma^2} A^T I(q) - \frac{1}{\sigma^2} 0_{N,1}
\]

is determined, where \( 0_{N,1} \) is a zero vector of length \( N \). To perform the IFT the size range \( 0 - D_{\text{max}} \) of \( P(r) \) has to be defined \textit{a priori}. However, \( D_{\text{max}} \) can be included as an additional fit parameter. Here, we performed the IFT for a total of 251 \( D_{\text{max}} \) values (between 10 and 30 nm). Furthermore, we varied \( \alpha \) over six orders of magnitude (from \( 10^0 \) to \( 10^5 \)) and determined for each set of \( (D_{\text{max}}, \alpha) \) the evidence according to Hansen (2000) and as described in detail by Bender et al. (2017). The \( D_{\text{max}} \) value for which the highest evidence was found was then used for the SVD and the KA as well.

With the SVD the data transfer matrix \( A \) is decomposed according to

\[
A = USV^T,
\]

where \( U \) and \( V \) are orthogonal \( N \times N \) and \( M \times M \) matrices, respectively, and \( S \) is an \( N \times M \) matrix whose main diagonal elements are the singular values \( s_i \). The correlation function \( P(r) \) is then calculated using

\[
P(r) = VS^+U^TI(q). \tag{4}
\]

The diagonal of \( S^+ \) contains the reciprocal values \( 1/s_i \), and all off-diagonal entries of \( S^+ \) are zero. In the case of ill-posed problems, many of the singular values will be very small which amplifies measurement uncertainties. Thus, using all singular values usually results in large unphysical oscillations in the derived distributions. The singular values of \( S \) are usually given in descending order, and thus a smoothing can be accomplished by reducing the number of singular values \( N_s \) which are considered for the reconstruction (Berkov et al., 2000).

Using the KA, \( P(r) \) is calculated by updating the elements \( P(r_j) \) after each iteration according to

\[
P_{k+1}(r_j) = P_k(r_j) + \frac{I(q_i) - \frac{A_j P_k(r_i)}{|A_i|}}{\sigma |A_i|^2} A_i,
\tag{5}
\]

where \( A_i \) is the \( i \)-th row of the matrix \( A \), \( A_i^T \) is its transpose, \( k \) is the iteration number, and one iteration contains a sweep over all rows \( i \). Here we shuffle randomly through all rows \( A_i \) and normalize the residuals \( I(q_i) - A_i P_k(r_i) \) to \( \sigma(q_i) \), similar to a weighted least-squares fit.

3.2. Results of the model-free data analysis

Fig. 3 shows the results for the IFT. The highest evidence was found for \( D_{\text{max}} = 22 \) nm. The fit of the data is shown in Fig. 3(a). In Fig. 3(b) the correlation functions for all \( \alpha \) values are plotted and the corresponding evidences are shown in Fig. 3(c). The \( P(r) \) with the highest evidence is highlighted in red. The correlation function exhibits one oscillation with the first zero crossing at around 10 nm. This value agrees well with the physical particle size. The oscillations can be attributed to a structure factor due to inter-particle interference.

---

![Figure 3](image-url)

\( a \) Fit of the 1D sector average at zero field by IFT. \( b \) Correlation functions \( P(r) \) determined for all \( \alpha \) values at \( D_{\text{max}} = 22 \) nm. The \( P(r) \) with highest evidence is marked red. \( c \) The evidences computed for all \( \alpha \) values.
Fritz-Popovski et al. (2011), which leads to a shift of the zero crossing towards smaller sizes and concerns nuclear and magnetic contributions as discussed at the end of this section.

Fig. 4 shows the results for the SVD. As can be seen, by increasing the number of singular values that are considered for the reconstruction of \( P(r) \) to \( N_r = 6 \), the total weighted error \( \chi^2 \) decreases [Fig. 4(c)] before leveling out. But for \( N_r > 6 \), large unphysical oscillations are obtained in the resultant correlation functions \( P(r) \), as shown in Fig. 4(b). The \( P(r) \) determined for \( N_r = 6 \) is highlighted in blue and the corresponding fit of \( I(q) \) is plotted in Fig. 4(a).

Fig. 5 shows the results for the KA. As plotted in Fig. 5(c), by increasing the iterations of the KA the total error \( \chi^2 \) tends to decrease until approaching \( \chi^2 \approx 3 \), similar to the SVD. The correlation function obtained after \( k = 100 \) iteration steps is plotted in green in Fig. 5(b) and the corresponding fit is shown in Fig. 5(a).

Fig. 6(a) shows the correlation functions determined by the IFT, the SVD and the KA. All three curves are basically identical. However, when \( D_{\text{max}} \) is not determined by the IFT but randomly chosen, this is not the case anymore. Fig. 6(b) shows as an example what happens when \( D_{\text{max}} \) is fixed to 50 nm. The correlation functions determined by the SVD and KA are basically the same as before, whereas the IFT results in a completely different one. This inability to automatically determine the appropriate maximum accessible sizes for the system and the need for a sophisticated guess by an informed user are significant issues for most of the common available tools to estimate \( P(r) \).

To better understand the physical meaning of \( P(r) \), we will now analyze the field-dependent data.

Fig. 7(a) shows the sector averages at 3 and 0 T [as shown in Fig. 2(b)], as well as the difference between the two. The sectors are parallel to the applied field and thus the scattering intensity \( I(q) \) at 3 T is dominated by the nuclear scattering intensity as the sample is nearly completely magnetically saturated at this field [see Fig. 1(b)]. Therefore, the correlation functions derived from \( I(q) \) are essentially the autocorrelation function of the nuclear scattering profile, and thus the oscillations in \( P(r) \) shown in Fig. 7(b) can be attributed to the nuclear structure factor (Weyerich et al., 1999). Negative values of \( P(r) \) are associated with distances that connect particle volumes with a scattering length density below the average. The difference between the scattering intensities measured at zero field and 3 T, on the other hand, is of purely magnetic origin (Mühlbauer et al., 2019). Thus, the derived correlation function \( P(r) \) contains information regarding the magnetic origin (Mu¨ hlbauer et al., 2019). Therefore, the correlation functions determined after all \( k = 100 \) iteration steps. The \( P(r) \) reconstructed after \( k = 100 \) is plotted in green. (c) The total weighted error \( \chi^2 \) computed after each iteration step \( k \).

![Figure 4](image4.png)

**Figure 4**

(a) Fit of the 1D sector average at zero field by the SVD. (b) Correlation functions \( P(r) \) determined for a varying number of singular values \( (N_r = 1–8) \). The \( P(r) \) reconstructed with \( N_r = 6 \) is marked in blue. (c) The total weighted error \( \chi^2 \) as a function of \( N_r \).

![Figure 5](image5.png)

**Figure 5**

(a) Fit of the 1D sector average at zero field by the KA. (b) Correlation functions \( P(r) \) determined after all \( k = 100 \) iteration steps. The \( P(r) \) reconstructed after \( k = 100 \) is plotted in green. (c) The total weighted error \( \chi^2 \) computed after each iteration step \( k \).

![Figure 6](image6.png)

**Figure 6**

(a) Comparison of the correlation functions determined by the IFT [red, from Fig. 3(b)], SVD [blue, from Fig. 4(b)] and KA [green, from Fig. 5(b)] at zero field. \( D_{\text{max}} = 22 \) nm was determined by the IFT. (b) Comparison of the correlation functions determined by the IFT (red), SVD (blue) and KA (green) when \( D_{\text{max}} \) is fixed to 50 nm.
moment correlations between neighboring particles at zero field. The observed negative values at $r > 10$ nm indicate antiferromagnetic-like moment correlations similar to what was observed by Bender et al. (2018). A comparison of the magnetic correlation functions with the distributions from Fig. 7(a) shows pronounced differences in the range 10–15 nm. This demonstrates that the nuclear and magnetic structure factors are not the same (Honecker et al., 2020). Furthermore, the magnetic correlation functions have no second positive peak in the range 17–22 nm, which indicates that magnetic order only exists between nearest neighbors. For larger distance, no correlation exists due to thermal fluctuations of the particle moments. Regarding the comparison of the three approaches for the determination of $P(r)$ we reiterate that all three approaches result in very similar correlation functions [Figs. 7(b) and 7(c)].

4. Discussion and summary

Here, we have analyzed the magnetic-field-dependent SANS pattern of a powder sample of 10 nm iron oxide nanoparticles measured with the neutron resonant spin-echo spectrometer RESEDA. Our analysis of the 1D sector averages shows that the IFT, SVD and KA all result in identical correlation functions $P(r)$. IFT is the standard approach for such problems and well established. In comparison with the IFT, both SVD and KA are less sensitive to the size range chosen for the reconstruction of $P(r)$. This means that, for the IFT, the $D_{\text{max}}$ value has to be included as a fit parameter, but this is not the case for the SVD and KA, which also significantly reduces the computational costs. In general, an advantage of the SVD and KA is the faster computation times. While this is negligible when handling 1D data, it is a huge advantage for 2D data analysis as shown e.g. by Bender et al. (2019, 2021). Regarding an automated data analysis, the KA in particular has great potential as also discussed in the context of other measurement techniques (Karpavičius et al., 2021). This approach could also be applied when simultaneously analyzing data sets of complementary characterization techniques. Such a global fit could be pursued for dilute spherical nanoparticle systems, e.g. to determine the functional form of the size distribution of the magnetic volume without skewing the data to a given, potentially false, distribution (e.g. log-normal) form.

Regarding small-angle scattering data analysis, we recommend applying the SVD and KA approaches in addition to the IFT, as a consistent result with all three methods significantly strengthens the confidence in the obtained correlation functions. The corresponding Python scripts can be found at https://github.com/PBenderLux/Data-analysis and are free to use. Furthermore, the SVD and KA will be implemented in the open-source SasView software package to provide easy access for data analysis.

Acknowledgements

The authors thank the Heinz Maier-Leibnitz Zentrum (MLZ) for providing beamtime at the instrument RESEDA. Furthermore, we grateful acknowledge the support of the Braunschweig International Graduate School of Metrology B-IGSM and the DFG Research Training Group 1952 Metrology for Complex Nanosystems. Open access funding enabled and organized by Projekt DEAL.

Funding information

Funding for this research was provided by Deutsche Forschungsgemeinschaft; Bundesministerium für Bildung und Forschung (grant No. 05K16WO6). This work benefited from the use of the SasView application, originally developed under NSF award No. DMR-0520547. SasView contains code developed with funding from the European Union’s Horizon 2020 research and innovation programme under the SINE2020 project, grant agreement No. 654000.

References


Bender, P., Wetterskog, E., Honecker, D., Fock, J., Frandsen, C., Moerland, C., Bogart, L. K., Posth, O., Szczerba, W., Gavilán, H.,