# sasPDF: pair distribution function analysis of nanoparticle assemblies from small-angle-scattering data

Chia-Hao Liu,[a] Eric Janke,[b] Ruipen Li,[c] Pavol Juhás,[d] Oleg Gang,[a,e,f] Dimitri V. Talapin[b] and Simon J. L. Billinge[a,g]

[a]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA, [b]Department of Chemistry, University of Chicago, Chicago, IL 60637, USA, [c]National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, New York 11973, USA, [d]Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA, [e]Center for Functional Nanomaterials, Energy and Photon Sciences Directorate, Brookhaven National Laboratory, Upton, New York 11973, USA, [f]Department of Chemical Engineering, Columbia University, New York, NY 10027, USA, and [g]Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, New York 11973, USA

## 1. Introduction

With the advent of high degrees of control over nanoparticle synthesis (Murray *et al.*, 1993; Hyeon *et al.*, 2001; de Mello Donegá *et al.*, 2005) attention is turning to assembling superlattices of them as metamaterials (Boles *et al.*, 2016; Choi *et al.*, 2016) and applications of nanoparticle assemblies (NPA) based devices such as solar cells and field effect transistors have been demonstrated (Talapin & Murray, 2005; Sargent, 2008; Talapin *et al.*, 2010). It is crucial to study the structures of these NPAs if

their properties are to be optimized. For example, it has been shown that the mechanical (Akcora *et al.*, 2009), optical (Young *et al.*, 2014), electrical (Vanmaekelbergh & Liljeroth, 2005) and magnetic (Sun *et al.*, 2000) properties can be further engineered by controlling the spatial arrangement of the constituents in the NPA.

Getting detailed quantitative structural information from NPAs, especially in 3D, is a challenging and largely unsolved problem. Small angle scattering and electron microscopy (EM) have been the major techniques for studying the structure of NPAs (Murray *et al.*, 2000; Talapin *et al.*, 2009). The technique of TEM yields high-resolution images of NPAs. To obtain quantitative structural information it is necessary to either analyze the images manually (Wang, 2000) or match observed images with patterns that are algorithmically generated from known structures (Shevchenko *et al.*, 2006). This approach can yield the structure types (Zhuang *et al.*, 2008) but does not typically result in the kind of quantitative 3D structural information we are used to obtaining for atomic structures of crystals, including accurate inter-particle vectors and distributions of inter-particle distances, or the range of structural coherence of the packing order. It is desirable to explore scattering approaches that can yield that kind of information.

The technique of small-angle x-ray or neutron scattering (SAS) has been an important tool to study objects that have sizes from nano- to micrometer length-scales (Turkevich & Hubbell, 1951; Glatter, 1977; Guinier, 1994; Koch *et al.*, 2003), such as large nanocrystals (Polte *et al.*, 2010) and biological molecules (Koch *et al.*, 2003), yielding information about the intrinsic shape, size distributions and scattering density of objects on these scales (Glatter, 1977; Beaucage, 1995; Pedersen, 1997; Volkov & Svergun, 2003; Beaucage *et al.*, 2004).

When these nanoscale objects aggregate, correlation peaks appear in the SAS data resembling atomic-scale interference peaks (Diffuse scattering and Bragg peaks), but

encoding information about particle packing rather than atomic packing (Murray *et al.*, 2000; Nykypanchuk *et al.*, 2008). Obtaining structural information about the NPAs from these correlation peaks appears to be a promising approach. Although the recent developments in SAS modeling demonstrates the ability to account for phase, morphology and orientations of NPs in a lattice (Yager *et al.*, 2014; Lu *et al.*, 2019), fitting the SAS data with robust structural models to obtain quantitative information about the structure has barely been explored (Macfarlane *et al.*, 2011)

On the other hand, the atomic pair distribution function (PDF) analysis of x-ray and neutron powder diffraction has proven to be a powerful tool for characterizing local order in materials, and for extracting quantitative structural information (Proffen *et al.*, 2005; Egami & Billinge, 2012; Zobel *et al.*, 2015; Keen & Goodwin, 2015) when the atoms are not long-range ordered, as is the case in nanoparticles. Here we extend PDF analysis to handle correlation peaks in the small angle scattering data, allowing us to study the arrangement of particles in nanoparticle assemblies using the same quantitative modeling tools that are available for studying the atomic arrangements in nanoparticles themselves. We describe the extension of the PDF equations in the small-angle scattering (SAS) regime and describe the data collection protocol for optimum data quality. We also present the PDFGETS3 software package that can be readily used to extract the PDF from small-angle scattering data. We then apply the SASPDF method to investigate structures of some representative NPA samples with different levels of structural order.

## 2. Samples

To test the method we obtained SAS data from the samples listed in Table 1. Synthesis details of these NPA samples can be found in the references listed in the table.

Table 1: Nanoparticle assemblies (NPA) considered in this study. Building block indicates the NP and surfactant linkers used to build the assemblies. $D$ is the particle diameter (one standard deviation in parentheses) estimated from TEM images and reported in the original publications listed in the Ref. column. Beamline is the x-ray beamline where the SAXS data were measured (see text for details). PMA is Poly(methyl acrylate) and DDT is dodecanethiol.

| Sample | Building block | $D$ (nm) | Beamline | Ref. |
|--------|----------------|----------|----------|------|
| Au NPA | DNA-capped Au NP | 11.4(1.0) | X21 | (Nykypanchuk *et al.*, 2008) |
| $Cu_2S$ NPA | DDT-capped $Cu_2S$ NP | 16.1(1.3) | 11-BM | (Han *et al.*, 2008) |
| $SiO_2$ NPA | PMA-capped $SiO_2$ NP | 14(4) | 11-BM | (Bilchak *et al.*, 2017) |

### 3. sasPDF method

The data were collected using a standard SAXS setup at an x-ray synchrotron source, with a 2D area detector mounted perpendicular to the beam in transmission geometry. Both the $Cu_2S$ NPA and the $SiO_2$ NPA samples were measured at beamline 11-BM at the National Synchrotron Light Source-II (NSLS-II). The $Cu_2S$ NPA powders were sealed between two rectangular Kapton tapes with a circular deposited area of diameter about 3 mm and thickness about 0.2 mm. The $SiO_2$ NPA formed a circular, free-standing stable film of diameter about 5 mm and thickness about 1 mm which was mounted perpendicular to the beam and no further sealing was carried out. The scattering data of the $Cu_2S$ NPA and $SiO_2$ NPA samples were collected with a Pilatus 2M (Dectris, Switzerland) detector with a sample-detector distance 2.02 m using an x-ray wavelength of 0.918 Å. An example of the diffraction image from the $Cu_2S$ NPA is shown in the inset of Fig. 1.
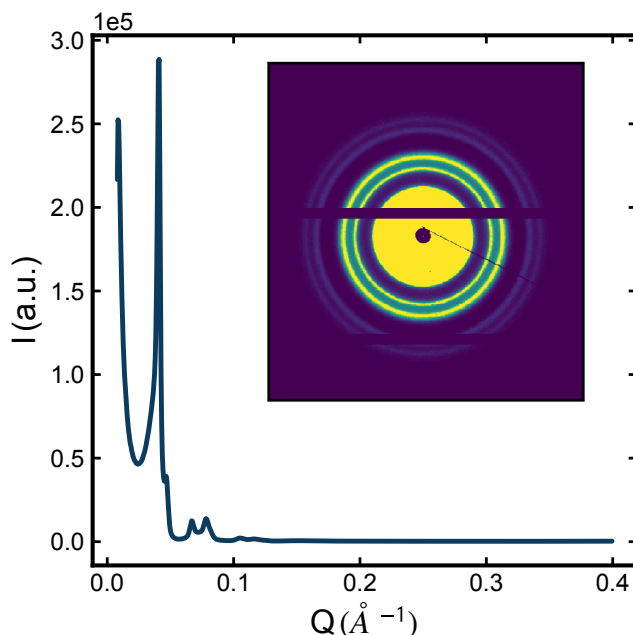
Fig. 1. Example of the 1D diffraction pattern $I_m(Q)$ from the Cu$_2$S NPA sample. The data were collected with the spot exposure time and scan exposure time reported in the text. The inset shows the corresponding 2D diffraction image. The horizontal stripes in the image are from the dead zone between panels of the detector. The diagonal line is the beam-stop holder.

The scattering from these samples is isotropic as the sample consists of powders of randomly oriented NPA crystallites, and the 2D diffraction images can be reduced to a 1D diffraction pattern, $I_m(Q)$, by performing an azimuthal integration around rings of constant scattering angle on the detector. This was done using PYFAI (Kieffer *et al.*, 2015). This requires a calibration of the experiment geometry described below, but the integrated 1D pattern from the 2D diffraction image is shown in Fig. 1. The relative positions and intensities of sharp peaks in the $I_m(Q)$ originate from the Debye-Scherrer rings in the 2D image.

We need to use a data acquisition strategy that mitigates effects of x-ray beam-damage to the sample. The linkers that connect nanoparticles in the assemblies play a crucial role for the NPA structure formed but are susceptible to degradation in the

intense x-ray beam that may result in changes in the NPA structure. To describe the strategy we separate the concepts of a "spot exposure time" and the "scan exposure time". The latter is the total integrated exposure time to obtain a dataset with sufficient statistics. The former is the length of time that any spot on the sample is exposed. The scan exposure then consists of multiple spot exposures, where the sample is translated after each spot exposure so that a fresh region of sample is exposed. For ease of experimentation we would like to determine a spot exposure time that is as long as possible whilst ensuring that the sample has not degraded significantly during that exposure. We have found that the maximum safe spot exposure time depends on the nature of the NPA sample, as well as experimental conditions such as x-ray energy, flux and sample temperature. It therefore requires a trial-and-error approach to determine it. To choose the optimal spot exposure time we locate the beam on a fixed spot of the sample and take a sequence of short exposures, monitoring for significant changes in the intensity of the strongest correlation peak in $I_m(Q)$. The spot exposure time determined this way for our experimental setup was 30 s for both $Cu_2S$ NPA and $SiO_2$ NPA samples and the scan exposure time was 5 minutes (30 s, 10 spots) for the $Cu_2S$ NPA sample and 10 minutes (30 s, 20 spots) for the $SiO_2$ NPA sample.

The scan exposure time is estimated based on an assessment of noise in the PDF given a desired $Q_{max}$, but it depends sensitively on the counting statistics in the high-$Q$ region of the diffraction pattern, which is easiest to assess by looking in the high-$Q$ region of the reduced structure function $F(Q)$. For illustration purposes, the effect of scan exposure time on the $F(Q)$ (and the resulting PDF) is illustrated in Fig. A1 of Appendix section.

For the calibration of the experimental geometry, such as sample-detector distance and detector tilting we use the calibration capability in the Python pack-

age PYFAI (Kieffer *et al.*, 2015). We first measured silver behenate (AgBh) (Gilles *et al.*, 1998) as a well characterized calibration sample. The *d*-spacing of the calibration sample, the x-ray wavelength and the pixel dimensions of the detector are known, which allows the geometric parameters to be refined in PYFAI. We found that selecting the strongest few rings (even just two or three work well) in the image allowed a stable refinement of the calibration parameters.

Finally, in this study we also consider legacy data from measurements carried out previously (Nykypanchuk *et al.*, 2008). The data of the Au NPA sample were collected at beamline X21 at the National Synchrotron Light Source (NSLS) from a sample loaded into a wax-sealed 1 mm diameter quartz capillary. The scattering data were collected with a MarCCD (Rayonix, USA) area detector using an x-ray wavelength of 1.55 Å. Details of the measurements are reported in (Nykypanchuk *et al.*, 2008).

The PDF, denoted $G(r)$, is a truncated sine Fourier transform of the reduced structure function $F(Q) = Q\left[(S(Q) - 1)\right]$ (Egami & Billinge, 2012)

$$G(r) = \frac{2}{\pi} \int_{Q_{\min}}^{Q_{\max}} F(Q) \sin(Qr)\, \mathrm{d}Q. \tag{1}$$

Since $F(Q)$ can be easily computed once $S(Q)$ is available, we will first focus on describing the precise definition of $S(Q)$ and its relation to the measured diffraction pattern $I_m(Q)$. The measured intensity, $I_m(Q)$, depends on experimental details such as the flux, and beam size of the x-ray source, the data collection time and the sample density. From the point of developing the SASPDF formalism, we will focus on the coherent scattering intensity $I_c(Q)$ (Egami & Billinge, 2012) which is obtained after correcting $I_m(Q)$ for the experimental factors as we describe below.

The coherent scattering intensity $I_c(Q)$ from a unit cell with $N_s$ atoms is (Egami & Billinge, 2012; Guinier, 1963)

$$I_c(\mathbf{Q}) = \sum_{m=1}^{N_s} \sum_{n=1}^{N_s} f_m^*(\mathbf{Q}) f_n(\mathbf{Q}) \exp\left[i\mathbf{Q} \cdot (\mathbf{r}_m - \mathbf{r}_n)\right], \tag{2}$$

where $\mathbf{Q}$ is the scattering vector, $f_m(\mathbf{Q})$ and $\mathbf{r}_m$ are the atomic form factor amplitude and position of $m$-th atom in the unit cell, respectively.

If the scattering from a sample is isotropic, for example, it is an untextured powder or a liquid with no anisotropy, the observed scattering intensity will depend only on the magnitude of $\mathbf{Q}$, $|\mathbf{Q}| = Q$ and not its direction in space. The observed scattering intensity in this case will depend on the orientationally averaged $I_c(\mathbf{Q})$,

$$I_c(Q) = \left\langle \sum_{m=1}^{N_s} \sum_{n=1}^{N_s} f_m^*(\mathbf{Q}) f_n(\mathbf{Q}) \exp\left[i\mathbf{Q} \cdot (\mathbf{r}_m - \mathbf{r}_n)\right] \right\rangle, \tag{3}$$

where $\langle \cdot \rangle$ denotes the orientational average.

This formalism is readily extended to the case where the scattering objects are not atoms, but are some other finite-sized object, for example, nanoparticles. In this case, the atomic form-factor would be replaced with the form-factor for the scattering objects in question. The form factor $f(\mathbf{Q})$ for a generalized scatterer, with volume $V$ and its electron density as a function of position $\rho(\mathbf{r})$ is (Guinier, 1963)

$$f(\mathbf{Q}) = \int_V \left[\rho(\mathbf{r}) - \rho_0\right] \exp\left(i\mathbf{Q} \cdot \mathbf{r}\right) \mathrm{d}\mathbf{r}, \tag{4}$$

where $\rho_0$ is the average electron density of the ambient environment of the scatterers.

In situations where there is only one type of scatterer we pull the form factors out of the sum **(Egami & Billinge, 2012)**. **Further, if the form factors and the structure factors are separable** Eq. 3 may be further simplified to

$$I_c(Q) = N_s \left\langle f^2(Q) \right\rangle + \langle f(Q) \rangle^2 \left\langle \sum_{m=1}^{N_s} \sum_{n \neq m}^{N_s} \exp\left[i\mathbf{Q} \cdot (\mathbf{r}_m - \mathbf{r}_n)\right] \right\rangle. \tag{5}$$

**This is valid for spherical or nearly spherical shaped particles, and may be more broadly true (Guinier *et al.*, 1955; Kotlarchyk & Chen, 1983; Li *et al.*, 2016) , though if the shape of the particle results in an orientation that depends on the packing, for example a long axis lies along a particular crystallographic direction in the nanoparticle assembly, this approximation**

will not be ideal and Eq. 3 would have to be used. We have not tested how badly the approximation breaks down in these circumstances.

Following the Faber-Ziman formalism (Faber & Ziman, 1965),

$$S(Q) = \frac{I_c(Q)}{N_s\langle f(Q)\rangle^2} - \frac{\langle f^2(Q)\rangle - \langle f(Q)\rangle^2}{\langle f(Q)\rangle^2},$$

(6)

we plug in $\langle f^2(Q)\rangle = \langle f(Q)\rangle^2$ and Eq. 6 becomes

$$I_c(Q) = N_s\langle f^2(Q)\rangle S(Q).$$

(7)

This expression is equivalent to representing the scatterers as points at the position of their scattering center, convoluted with their electron distributions. The resulting structure function, $S(Q)$, yields the arrangement of scatterers in the sample. This expression is often expressed in the SAS literature as (Guinier, 1963)

$$S(Q) = \frac{I_c(Q)}{N_s P(Q)}.$$

(8)

Where $P(Q)$ is equivalent to $\langle f^2(Q)\rangle$ (Guinier, 1963), the orientational average of the square of the form-factor. We note that, as with the atomic PDF, the above analysis can be generalized to the cases of scattering from multiple types of scatterers (Kotlarchyk & Chen, 1983; Yager *et al.*, 2014; Senesi & Lee, 2015) and in the SAS case approximate corrections for asphericity of the electron density (Jones *et al.*, 2010; Ross *et al.*, 2015; Zhang *et al.*, 2015), may be applied.

To determine $S(Q)$ we need to have $P(Q)$. $P(Q)$ can be computed from a given electron density, or directly measured. For the case of a NPA sample, the precise scattering properties of the NP ensemble in the sample, including any polydispersity or distribution of geometric shapes, are not always known, therefore it is best to measure the form factor directly, as described below. In general we do not know $N_s$ and all of the experimental factors (for example, the incident flux, multiple scattering and so on). The algorithm (Billinge & Farrow, 2013) that is widely used for carrying

out corrections for these effects in the atomic PDF literature (Juhás *et al.*, 2013) is also suitable for the SAS data. It takes advantage of our knowledge of the asymptotic behavior of the $S(Q)$ function to obtain an *ad hoc* but robust estimation of $S(Q)$ from the measured $I_m(Q)$. This is described in detail in (Juhás *et al.*, 2013). The resulting scale of the PDF is not well determined, but when fitting models to the data this is not a problem (Peterson *et al.*, 2003), and in practice it gives close to a correct scale for high quality measurements. Here we show that we can take the same approach to obtain the PDF from the measured SAS data here.

In the test experiments we describe here, in each case the form factor of the nanoparticles was obtained from a measurement. The NPs are suspended in solvent at a sufficient dilution to avoid significant inter-particle correlations. **If the NPs start aggregating in the solution, a plateau appears in the low-$Q$ region of the scattering intensity, and observation of such a plateau may indicate a problem with the form factor sample.** The SAS signal of the dilute NP solution is measured with good statistics over the same range of $Q$ as the measurement of the nanoparticle assemblies themselves, and ideally on the same instrument. The signal of the solvent and its holder is also measured and then subtracted from the SAS signal of the dilute NP solution to obtain the correct particle form factor signal. We emphasize that it is important to measure exactly the same batch of NPs to have an accurate form factor for the NPA sample considered.

A form factor measured with high statistics is crucial as the signal in $I_c(Q)$ is weak in the high-$Q$ region and noise from the $P(Q)$ measurement can be significant in this region. Fig. A2 shows the effect on $F(Q)$ (and the resulting PDF), when processed using $P(Q)$ from different scan exposure times. It is clear that the statistics of the form-factor measurement has a significant effect on the results. In cases where any signal in $P(Q)$ does not change rapidly it may be smoothed to reduce the effects of

limited statistics, at the cost of possibly introducing bias if the smoothing is not done ideally. This will be particularly relevant when the nanoparticles are not monodisperse, as is somewhat common.

The experimental PDF $G(r)$ is then obtained via the Fourier transformation, Eq. 1. The success of the SASPDF method depends heavily on the good statistics (high signal-to-noise ratio) throughout the entire diffraction pattern $I_c(Q)$ and the form factor $P(Q)$, as important information about the structure may reside in the high-$Q$ region where the signal intensity is weak. It is recommended to use intense radiation sources such as synchrotrons. A comparison in data quality from an in-house instrument and a synchrotron source is shown in Fig. A3 of Appendix section.

## 4. Software

To facilitate the SASPDF method, we implemented a PDFGETS3 software program for extracting the sasPDF from experimental data. Information about obtaining the software is on the DIFFPY organization website (https://www.diffpy.org). The software is currently supported in Python 2 (2.7) and Python 3 (3.4 and above). It requires a license and is free for researchers conducting open academic research, but other uses require a paid license.

The PDFGETS3 program takes in a measured diffraction pattern $I_m(Q)$ and a form factor, $P(Q)$, as the inputs and applies a series of operations such as subtraction of experimental effects and form factor normalization and outputs the PDF, $G(r)$. If the square of the orientationally averaged form-factor $\langle f(Q)\rangle^2$ is available, both $P(Q)$ and $\langle f(Q)\rangle^2$ can be specified in the program, and the $S(Q)$ will be computed based on Eq. 6 which accounts for the anisotropy of scatterers in the material. Processing parameters used in PDFGETS3 operations, such as the form-factor file, the $Q$-range of the Fourier transformation on $F(Q)$ and the $r$-grid of the output $G(r)$, can be set

in a configuration file in the same way detailed in (Juhás *et al.*, 2013). Similar to
PDFgetX3, an interactive window for tuning these processing parameters, is also
available in PDFgetS3. An illustration of such interactive interface is shown in the
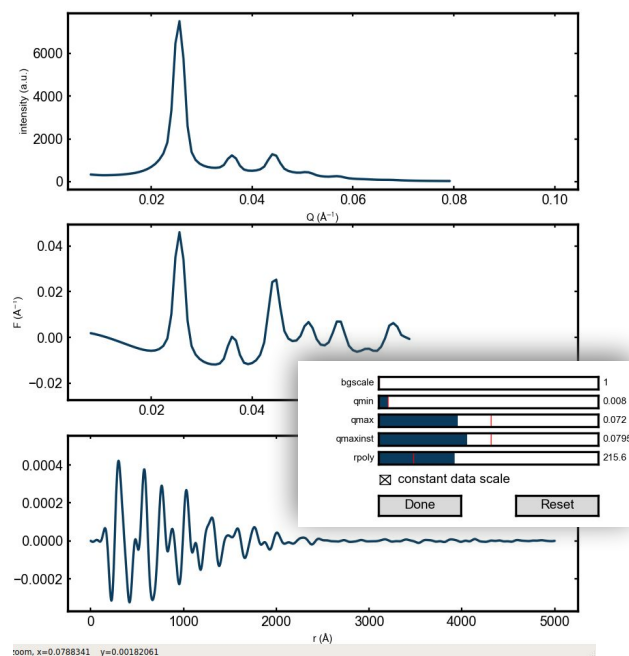Fig. 2.



Fig. 2. Illustration of the interactive interface for tuning the process parameters in the
PDFgetS3 program.

Sliders for each processing parameter allow the user to inspect the effect on the
output PDF data immediately.

Once the optimal processing parameters are determined based on the quality of
the PDF, those parameter values will be stored as part of the metadata in the out-
put $G(r)$ file. The final values of $Q_{min}$ and $Q_{max}$ should be used when calculating
PDF from a structure model, as these parameters contribute to the ripples in the
PDF (Peterson *et al.*, 2003). Full details on how to use the program is available on

the DIFFPY organization website.

## 5. PDF method

The PDF gives the scaled probability of finding two scatterers in a material a distance $r$ apart (Egami & Billinge, 2012). For a macroscopic object with $N$ scatterers, the atomic pair density, $\rho(r)$, and $G(r)$ can be calculated from a known structure model using

$$\rho(r) = \frac{1}{4\pi r^2 N} \sum_m \sum_{n \neq m} \frac{f_m(Q) f_n^*(Q)}{\langle f(Q) \rangle_{s.a.}^2} \delta(r - r_{mn}), \tag{9}$$

and

$$G(r) = 4\pi r \left[ \rho(r) - \rho_0 \right]. \tag{10}$$

Here, $\rho_0$ is the number density of scatters in the object. $f_m(Q) = \langle f_m(\mathbf{Q}) \rangle$ is the orientationally averaged form-factor of the $m$-th scatterer. $\langle f(Q) \rangle_{s.a.} = \sum_{m=1}^N (\frac{N_m}{N}) f_m(Q)$ denotes the sample average of $f(Q)$ over all scatterers in the material, where $N_m$ is the number of scatterers that are of the same kind as the $m$-th scatter. Finally, $r_{mn}$ is the distance between the $m$-th and $n$-th scatterer. We use Eq. 10 to fit the PDF generated from a structure model to a PDF determined from experiment.

PDF modeling, where it is carried out, is performed by adjusting parameters of the structure model, such as the lattice constants, positions of scatterers and particle displacement parameters (PDPs), to maximize the agreement between the theoretical and an experimental PDF. In practice, the delta functions in Eq. 10 are Gaussian-broadened to account for thermal motion of the scatterers and the equation is modified with a damping factor to account for instrument resolution effects. The modeling of SASPDF can be done seamlessly with tools developed in the atomic PDF field, with parameter values scaled accordingly. We outline the modeling procedure using PDFGUI (Farrow *et al.*, 2007), which is widely used to model the atomic PDF. In PDFGUI, the nanoparticle arrangements can simply be treated analogously as atomic

structures, with a unit cell and fractional coordinates, but the lattice constants reflect the size of the NPA, which is usually at the order of 100 nm = 1000 Å. The atomic displacement parameters (ADPs) defined in PDFGUI can be directly mapped to the particle displacement parameters (PDPs) in the sASPDF case and, empirically, we find the PDP values are roughly four to five orders of magnitude larger than the values of its counterpart on the atomic scale, therefore starting values of 500 Å$^2$ are reasonable. These will be adjusted to the best-fit values during the refinement.

The measured sASPDF signal falls off with increasing $r$. The damping may originate from various factors, for example, the instrumental $Q$-space resolution (Egami & Billinge, 2012) and finite range of order in the superlattice assembly. In PDFGUI there is a a Gaussian damping function $B(r)$,

$$B(r) = \exp\left[-\frac{(rQ_{damp})^2}{2}\right].$$ (11)

We define a $r_{damp}$ parameter

$$r_{damp} = \frac{1}{Q_{damp}},$$ (12)

which is the distance where about one third of the sASPDF signal disappears completely. It is also possible to generalize the modeling process to the case of a customized damping function and non-crystallographic structure with Diffpy-CMI (Juhás et al., 2015), which is a highly flexible PDF modeling program. In the following section, we use PDFGUI for modeling data from more ordered samples (Au NPA and Cu$_2$S NPA) and Diffpy-CMI for modeling data from a disordered sample (SiO$_2$ NPA).

## 6. Application to representative structures

To illustrate the sASPDF method we have applied it to some representative nanoparticle assemblies from the literature (Nykypanchuk et al., 2008; Han et al., 2008; Bilchak et al., 2017). The first example is from DNA templated gold nanoparticle superlattices,

originally reported in (Nykypanchuk *et al.*, 2008). The measured intensity, $I_m(Q)$, the
reduced total structure function $F(Q) = Q[S(Q) - 1]$, and the PDF $G(r)$ are shown
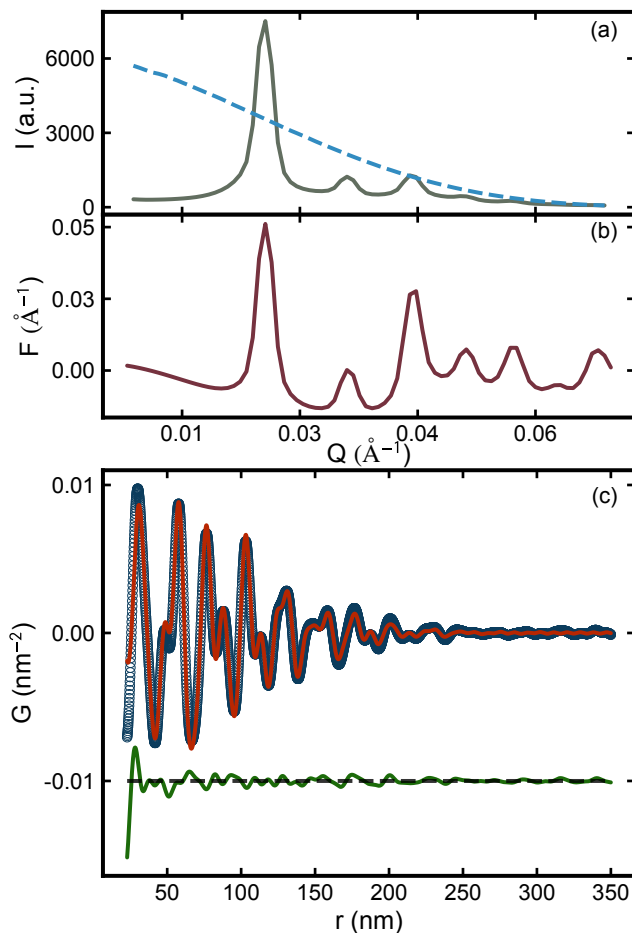in Fig. 3(a), (b) and (c), respectively.



Fig. 3. Measured (a) scattering intensity $I_m(Q)$ (grey) and form factor $P(Q)$ (blue),
(b) reduced total structure function $F(Q)$ (red) and (c) PDF (open circle) of Au
NPA. In (c), the PDF calculated from body-center cubic (bcc) model is shown in
red and the difference between the measured PDF and the bcc model is plotted in
green with an offset.

It is clear that the data corrections and normalizations to get $F(Q)$ result in a more
prominent signal in the high-$Q$ regime of the scattering data, and a highly structured
PDF after the Fourier transform (Fig. 3(c)).

Table 2: Refined parameters for NPA samples. Model column specifies the structural model used to fit the measured PDF. $a$ is the lattice constant of the unit cell, PDP stands for particle displacement parameters, which is an indication of the uncertainty in position of the nanoparticles. $r_{damp}$ is the standard deviation of the Gaussian damping function defined in Eq. 12. Scale is a constant factor being multiplied to the calculated PDF. $R_w$ is the residual-function, commonly used as a measure for the goodness of fit.

|  | Au NPA | Cu$_2$S NPA |
| --- | --- | --- |
| Model | bcc | fcc |
| $a$ (nm) | 34.73 | 26.55 |
| PDPs (nm$^2$) | 4.78 | 0.253 |
| $r_{damp}$ (nm) | 83.3 | 61.4 |
| Scale | 0.537 | 0.361 |
| $R_w$ | 0.172 | 0.221 |

The PDF signal dies off around 350 nm, which puts a lower bound on the size of the NPA. The first peak in the PDF is located at 30.07 nm which corresponds to the nearest inter-particle distance in the assembly. This distance is expected because the shortest inter-particle distance can be approximated as the average size of Au NPs (11.4 nm) plus the average surface-to-surface distance ($d_{ss}$) between nearest neighbor NPs (18 nm) (Nykypanchuk $et$ $al.$, 2008). Peaks beyond the nearest neighbor give an indication of characteristic inter-particle distances in the assembly and codify the 3D arrangement of the nanoparticles in space.

A semi-quantitative interpretation of conventional powder diffraction data suggested the Au NPA forms a body-centered cubic (bcc) structure (Nykypanchuk $et$ $al.$, 2008). We therefore test the bcc model against the measured PDF. The fit is shown in Fig. 3(c) and the refined parameters are reproduced in Table 2. The agreement between the bcc model and the measured data is good. We refine a lattice parameter that is ∼3 % smaller than the value reported from the semi-quantitative analysis. Additionally, the PDF gives information about the disorder in the system in the form of the crystallite size (∼350 nm) and the particle displacement parameter (PDP), the nanoparticle assembly version of the atomic displacement parameter (ADP) in

atomic systems. The PDF derived crystallite size is drastically smaller than the value ($\sim$500 nm) estimated from the FWHM of the first correlation peak (Nykypanchuk *et al.*, 2008) and it is clear by visual inspection of the PDF that the $\sim$500 nm value is an overestimate. These results suggest that even in the case where it is straightforward to infer the geometry of underlying assembly using qualitative and semi-quantitative means there is an advantage to carrying out a more fully quantitative sasPDF analysis.

Next we consider a dataset from a dodecanethiol (DDT)-capped $Cu_2S$ NPA (Han *et al.*, 2008). In this case the form factor is measured on an in-house Cu $K_\alpha$ instrument. This was necessary in the current case because the instability of the nanoparticles in suspension prevented a good measurement to be made at the synchrotron. As a result the form factor measurement was somewhat noisy (Fig. A4(a), blue curve) and we elected to smooth it by applying a Savitzky-Golay filter (Orfanidis, 1996). The smoothing parameters of window size and polynomial order were selected as 13 and 2, respectively, based on a trial and error approach optimized to result in a good smoothing without changing the shape of the signal. The smoothed curve is shown in Fig. A4(a). It is worth noting that in general, a smoothing process may start failing when the signal-to-ratio in the data exceeds a certain threshold, and so good starting data is always desirable. A conventional semi-quantitative analysis on diffraction data from the sample collected on an in-house Cu $K_\alpha$ instrument is shown in Fig. A5. It suggests the NPA forms a face-centered cubic (fcc) structure with an inter-particle distance of 18.8 nm. The SAS PDF obtained from the same NPA sample is shown in Fig. 4. It clearly shows that peaks die out at around 300 nm, which again signifies the crystallite size of the assembly. The first peak of the measured PDF is at 18.5 nm, corresponding to the inter-particle distance in the NPA. This value is about 1.6 % smaller than the value estimated from the semiquantitative analysis.

The best-fit PDF of a close-packed face-centered cubic (fcc) structural model is shown in red in the figure and refined structural parameters are presented in Table 2.
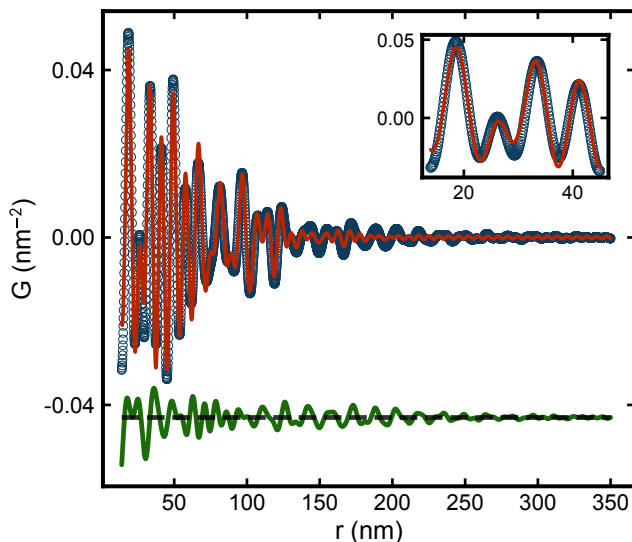


Fig. 4. Measured PDF (open circle) of a $Cu_2S$ NPA sample with the best fit PDF from the fcc model (red line). The Difference curve between the data aAs a result, nd model is plotted offset below in green. The inset shows the region of the first four nearest neighbor peaks of the PDF along with the best-fit fcc model.

The fcc model yields a rather good agreement with the measured PDF of $Cu_2S$ NPA in the short-range (up to ∼130 nm). Interestingly, the refined lattice parameter of this cubic model is 26.55 nm, from which we can calculate an average inter-particle spacing of 18.78 nm, which is much closer to the value estimated from the in-house data than directly extracting the position of the first peak in the PDF. The first peak in the PDF calculated from the model lines up with that from the data at 18.5 nm, which means that the position of the peak, as extracted from the peak maximum, underestimates the actual inter-particle distance by ∼ 1.5%, which may be due to the sloping background in the $G(r)$ function (Egami & Billinge, 2012). Quatntiative modeling is always preferred for obtaining the most precise determination of inter-particle distance.

The region of the first four nearest-neighbor peaks in the PDF, together with the fit, is shown in the inset to Fig. 4. A close investigation of this region shows subtle shifts in peak positions between the measured PDF and the refined fcc model. At around 26 nm (second peak), the peak from refined model is shifted to higher-$r$ compared to the measured data, while at around 33 nm (third peak), the relative shift in peak position is towards the low-$r$ direction. These discrepancies suggest the NPA structure is more complicated than a simple fcc structure and may reflect the presence of internal twined defects, for example (Banerjee *et al.*, 2019). Furthermore, it is clear that signal persists in the measured PDF in the high-$r$ region that is not captured by the single-phase damped fcc model. There is clearly more to learn about the structure of the NPA by finding improved structural models and fitting them to the PDF, though this is beyond the scope of the current paper.

It is worth noting that the refined PDP value of DDT-capped $Cu_2S$ NPA is significantly smaller than that of the DNA-templated Au NPA described above. A small PDP means the positional disorder of the NPs is small which would be expected with shorter, more rigid, linkers between the particles. The inter-particle distance (18.8 nm) can be decomposed into the sum of the average particle diameter (16.1 nm) and the particle-surface to particle-surface distance $d_{ss} = 2.7$ nm. Based on the chemistry the linker would have length 1.7 nm in the fully stretched out state, which would result in a maximal $d_{ss} = 3.4$ nm if the linkers were stretched out and oriented radially. Half the observed surface-surface distance, $d_{ss}/2 = 1.4$ nm. This result is reasonable, suggesting the linkers are either not straight, or not radial, or possibly partially interleaved. Nonetheless, this shorter linker would be expected to be more rigid and therefore consistent with our observation of a smaller PDP value from the sasPDF analysis.

Finally we consider a dataset from a **more disordered system**, poly(methyl acry-

late) (PMA) capped $SiO_2$ NPA sample. The molecular weight and density of the capping polymers can be varied and in the current sample was 0.47 chains/nm$^2$ and 132 kDa, respectively. Studies had suggested that similar NPA samples exhibit no structural order, based on an empirical metric using the height of the first peak in the measured $S(Q)$ (Bilchak *et al.*, 2017). Here we apply the sasPDF method to obtain a more **complete** understanding of the structure of this NPA.

To start we want to verify whether there is any evidence for close-packing of the NPs so we start with face-centered cubic (fcc), hexagonal close-packed (hcp) and icosahedral models (Baus, 1983) to see if any good agreement between the structural model and the data can be achieved. The results are shown in Fig. 5.

Fig. 5. Measured PDF (open circle) of the SiO$_2$ NPA sample and calculated PDFs (solid lines) from (a) fcc, (b) hcp, (c) icosahedral (d) damped sine-wave models. In each panel, the line in red is the PDF calculated from the corresponding model with optimum parameters. From (a) to (c), the line in grey is the PDF calculated from the same model but with small PDPs. In (d), the line in grey is the PDF calculated from the undamped sine-wave model. Dashed lines indicate maxima of the sharper PDFs in each panel.

A highly broadened version of the fcc model yields a reasonable agreement with the data (Fig. 5(a)). The PDF of the fcc structure is shown in grey in the figure, and after significant broadening it yields the red curve. Other close-packed cluster models, hcp and icosahedral, were also **tried** (Fig. 5(b) and (c)). Finally, a damped-sine wave model that is appropriate for highly disordered systems where on average the packing around a central atom is completely isotropic was tried (Fig. 5(d)) (Cargill, 1975; Konnert & Karle, 1973; Doan-Nguyen *et al.*, 2014). **The fcc and hcp structures would be expected for close-packed hard spheres. Interestingly, in the current case, the very simple damped sine-wave model with only 3 parameters yields a more satisfactory fit to the data than the more complicated (with 4-5 parameter) close-packed models, suggesting that we have soft-sphere like packing in the current case. This finding is being explored in more detail in another publication (Liu *et al.*, in press).** This example shows how interparticle packing may be examined using SASPDF method even when there is not long or intermediate range order.

## 7. Acknowledgements

# Appendix A
# Illustration of of data acquisition strategy

In this section, important effects related to the data quality are illustrated. In general, for a successful sasPDF experiment, it is crucial to achieve a high signal-to-noise ratio throughout the entire $Q$-range for both the form factor and sample measurements. Figs. A1 and A2 show the effect of insufficient counting statistics in the sample and form factor measurements, respectively. Fig. A3 compares the data quality from an in-house instrument and a synchrotron source. Finally, Fig. A4 shows the remedial effect of smoothing data from in-house measured form factor with insufficient statistics. The proper remedy is to measure with sufficient statistics in the first place.
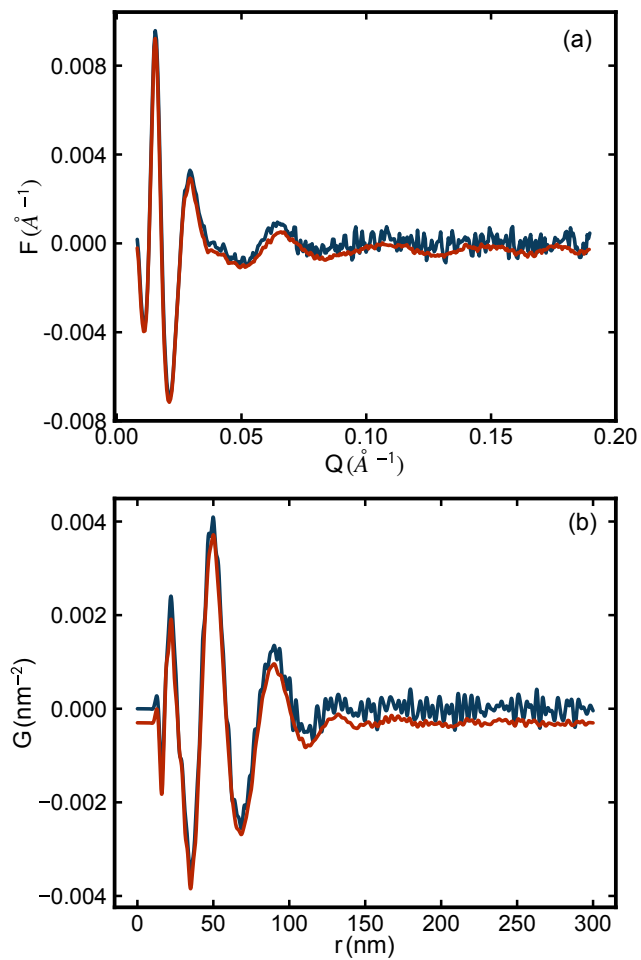
Fig. A1. (a) Reduced structure functions $F(Q)$ and (b) PDFs $G(r)$ of the $SiO_2$ NPA sample with different scan exposure times. Blue is from data with 1 s scan exposure time and red is from data with 30 s scan exposure time. In both panels, data are plotted with a small offset for ease of viewing. In both cases the form factor was measured with an scan exposure time of 600 s.
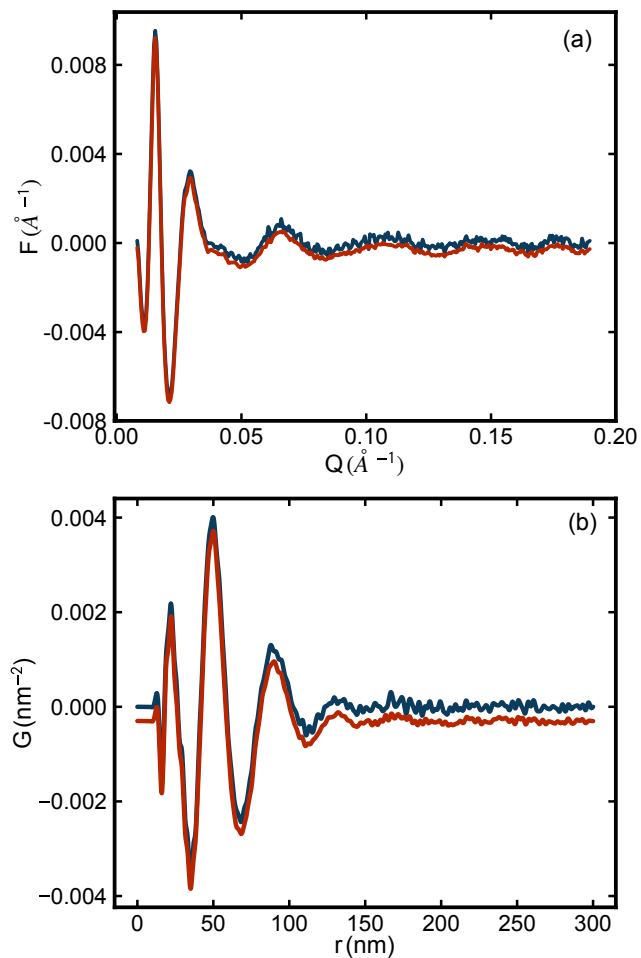
Fig. A2. (a) Reduced structure functions $F(Q)$ and (b) PDFs $G(r)$ of the SiO$_2$ NPA sample processed with form factor $P(Q)$ from different scan exposure times. Blue is made with a form-factor measured for 30 s and red is with a form factor collected for 600 s. In both cases the scan exposure time for the NPA sample was 600 s. In both panels, data are plotted with a small offset for ease of viewing.
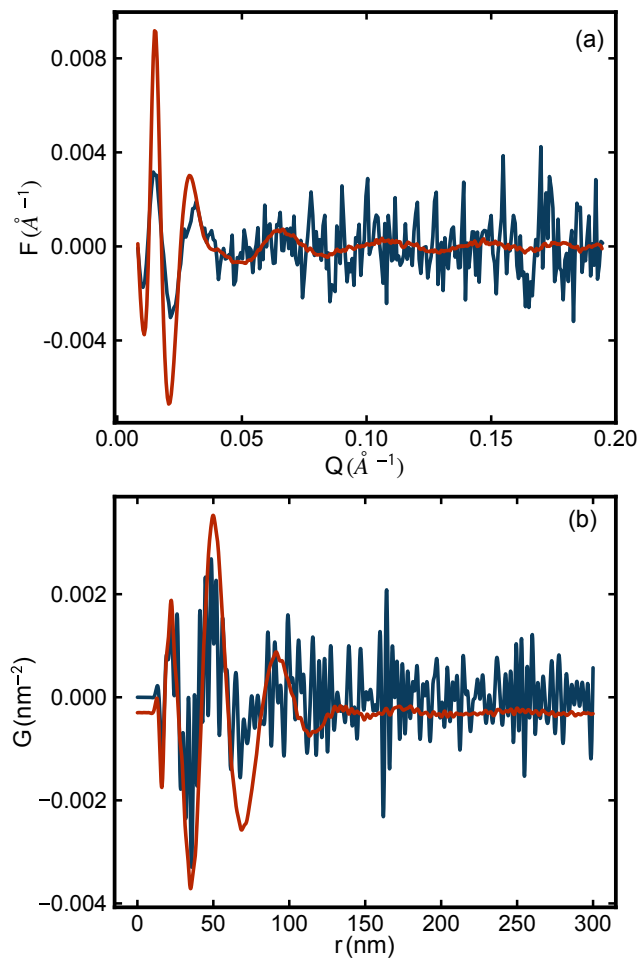
Fig. A3. (a) Reduced structure functions $F(Q)$ and (b) PDFs $G(r)$ of the $SiO_2$ NPA sample. Blue is from data collected at Columbia University using a SAXSLAB (Amherst, MA) instrument with a 2-hour (7200 s) scan exposure time for both $I(Q)$ and $P(Q)$ measurements. Red is from data collected at beamline 11-BM, NSLS-II with 30 s scan exposure time for both $I_m(Q)$ and $P(Q)$ measurements.
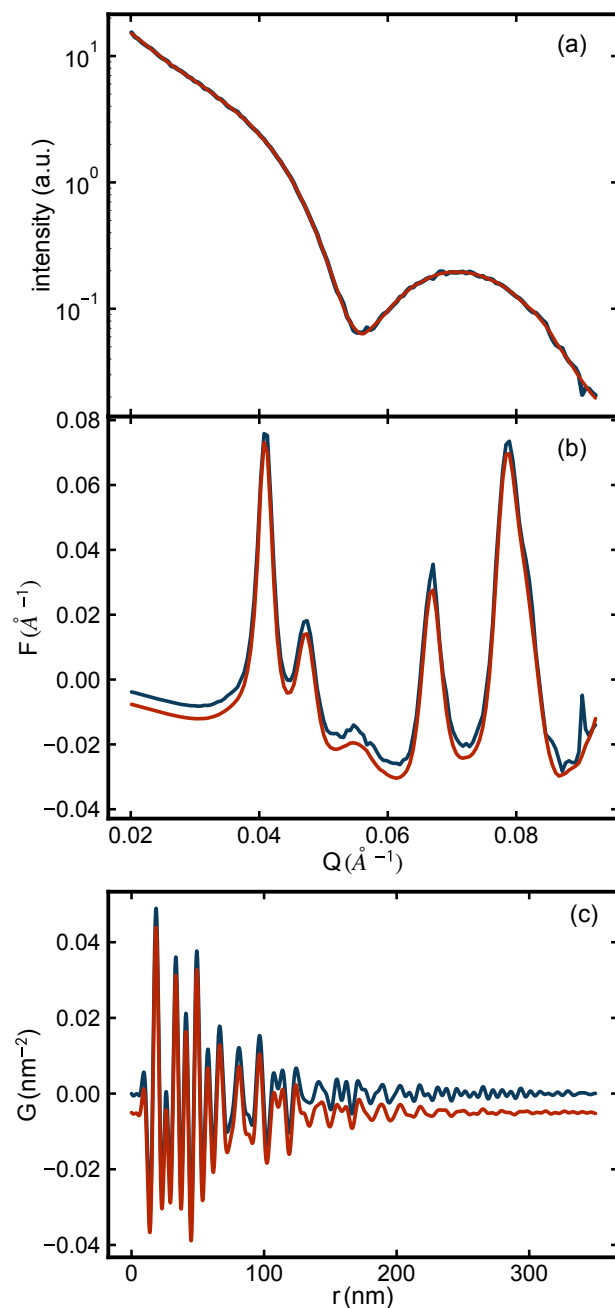
Fig. A4. (a) Form factor signal from Cu$_2$S NPs. Blue is the raw data collected at an in-house instrument and red is the data smoothed by applying a Savitzky-Golay filter with window size 13 and fitted polymer order 2. (b) reduced structure functions, $F(Q)$, and (c) PDFs, $G(r)$ from the Cu$_2$S NPA sample. In both panel, blue represents the data processed with raw form factor signal and red represents the data processed with smoothed form factor signal. Curves are offset from each other slightly for ease of view.
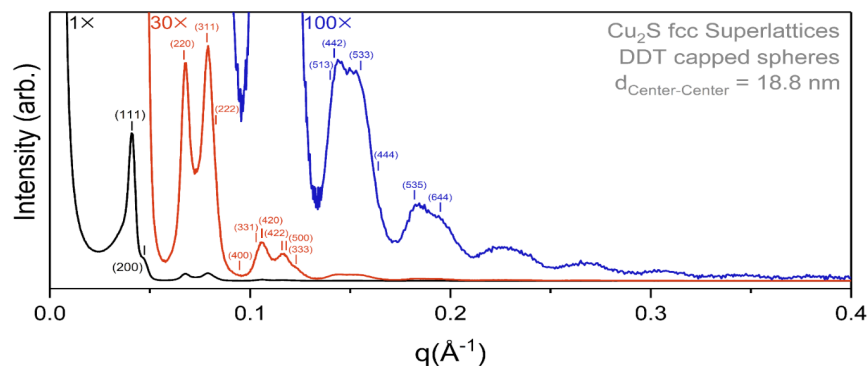
Fig. A5. Semi-quantitative structural analysis on Cu$_2$S NPA sample.

# References

Akcora, P., Liu, H., Kumar, S. K., Moll, J., Li, Y., Benicewicz, B. C., Schadler, L. S., Acehan, D., Panagiotopoulos, A. Z., Pryamitsyn, V., Ganesan, V., Ilavsky, J., Thiyagarajan, P., Colby, R. H. & Douglas, J. F. (2009). *Nat. Mater.* **8**(4), 354–359.

Banerjee, S., Liu, C.-H., Jensen, K. M. O., Juhás, P., Lee, J. D., Ackerson, C. J., Murray, C. B. & Billinge, S. J. L. (2019). *Acta Crystallogr. A.*

Baus, M. (1983). *Mol. Phys.* **50**(3), 543–565.

Beaucage, G. (1995). *J. Appl. Crystallogr.* **28**(6), 717–728.

Beaucage, G., Kammler, H. K. & Pratsinis, S. E. (2004). *J Appl Cryst*, **37**(4), 523–535.

Bilchak, C. R., Buenning, E., Asai, M., Zhang, K., Durning, C. J., Kumar, S. K., Huang, Y., Benicewicz, B. C., Gidley, D. W., Cheng, S., Sokolov, A. P., Minelli, M. & Doghieri, F. (2017). *Macromolecules*, **50**(18), 7111–7120.

Billinge, S. J. L. & Farrow, C. L. (2013). *J. Phys: Condens. Mat.* **25**, 454202.

Boles, M. A., Engel, M. & Talapin, D. V. (2016). *Chem. Rev.* **116**(18), 11220–11289.

Cargill, G. S. (1975). In *Solid State Physics*, edited by H. Ehrenreich, F. Seitz & D. Turnbull, vol. 30, pp. 227–320. Academic Press.

Choi, J.-H., Wang, H., Oh, S. J., Paik, T., Sung, P., Sung, J., Ye, X., Zhao, T., Diroll, B. T., Murray, C. B. & Kagan, C. R. (2016). *Science*, **352**(6282), 205–208.

de Mello Donegá, C., Liljeroth, P. & Vanmaekelbergh, D. (2005). *Small*, **1**(12), 1152–1162.

Doan-Nguyen, V. V. T., Kimber, S. A. J., Pontoni, D., Hickey, D. R., Diroll, B. T., Yang, X., Miglierini, M., Murray, C. B. & Billinge, S. J. L. (2014). *ACS Nano.* **8**(6), 6163–6170.

Egami, T. & Billinge, S. J. L. (2012). *Underneath the Bragg peaks: structural analysis of complex materials.* Amsterdam: Elsevier, 2nd ed.

Faber, T. E. & Ziman, J. M. (1965). *Philos. Mag.* **11**, 153–157.

Farrow, C. L., Juhás, P., Liu, J., Bryndin, D., Božin, E. S., Bloch, J., Proffen, T. & Billinge, S. J. L. (2007). *J. Phys: Condens. Mat.* **19**, 335219.

Gilles, R., Keiderling, U. & Wiedenmann, A. (1998). *J Appl Cryst, J Appl Crystallogr*, **31**(6), 957–959.

Glatter, O. (1977). *J Appl Cryst, J Appl Crystallogr*, **10**(5), 415–421.

Guinier, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies.* San Francisco: W.H. Freeman.

Guinier, A. (1994). *X-Ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies.* Courier Corporation.

Guinier, A., Fournet, G., Walker, C. & Yudowitch, K. (1955). *Small-angle scattering of x-rays.* New York: John Wiley & Sons, Inc.

Han, W., Yi, L., Zhao, N., Tang, A., Gao, M. & Tang, Z. (2008). *J. Am. Chem. Soc.* **130**(39), 13152–13161.

Hyeon, T., Lee, S. S., Park, J., Chung, Y. & Na, H. B. (2001). *J. Am. Chem. Soc.* **123**(51), 12798–12801.

Jones, M. R., Macfarlane, R. J., Lee, B., Zhang, J., Young, K. L., Senesi, A. J. & Mirkin, C. A. (2010). *Nat. Mater.* **9**(11), 913–917.

Juhás, P., Davis, T., Farrow, C. L. & Billinge, S. J. L. (2013). *J. Appl. Crystallogr.* **46**, 560–566.

Juhás, P., Farrow, C. L., Yang, X., Knox, K. R. & Billinge, S. J. L. (2015). *Acta Crystallogr. A*, **71**(6), 562–568.

Keen, D. A. & Goodwin, A. L. (2015). *Nature,* **521**(7552), 303–309.

Kieffer, J., Ashiotis, G., Deschildre, A., Nawaz, Z., Wright, J. P., Karkoulis, D. & Picca, F. E. (2015). *J. Appl. Crystallogr.* **48**, 510–519.

Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). *Q. Rev. Biophys.* **36**(2), 147–227.

Konnert, J. H. & Karle, J. (1973). *Acta Cryst A*, **29**(6), 702–710.

Kotlarchyk, M. & Chen, S.-H. (1983). *The Journal of Chemical Physics*, **79**(5), 2461–2469.

Li, T., Senesi, A. J. & Lee, B. (2016). *Chem. Rev.* **116**(18), 11128–11180.

Liu, C.-H., Buenning, E., Jhalaria, M., Brunelli, M., Durning, C. J., Huang, Y., Benicewicz, B. C., Kumar, S. K. & Billinge, S. J. L. (in press). A structural signature for jamming in soft colloids.

Lu, F., Vo, T., Zhang, Y., Frenkel, A., Yager, K. G., Kumar, S. & Gang, O. (2019). *Sci. Adv.* **5**(5), eaaw2399.

Macfarlane, R. J., Lee, B., Jones, M. R., Harris, N., Schatz, G. C. & Mirkin, C. A. (2011). *Science*, **334**(6053), 204–208.

Murray, C. B., Kagan, C. R. & Bawendi, M. G. (2000). *Annu. Rev. Mater. Sci.* **30**(1), 545–610.

Murray, C. B., Norris, D. J. & Bawendi, M. G. (1993). *J. Am. Chem. Soc.* **115**(19), 8706–8715.

Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. (2008). *Nature,* **451**(7178), 549–552.

Orfanidis, S. J. (1996). *Introduction to Signal Processing.* Prentice Hall.

Pedersen, J. S. (1997). *Advances in Colloid and Interface Science*, **70**, 171–210.

Peterson, P. F., E. S. Božin, Proffen, T. & Billinge, S. J. L. (2003). *J. Appl. Crystallogr.* **36**, 53.

Polte, J., Erler, R., Thünemann, A. F., Sokolov, S., Ahner, T. T., Rademann, K., Emmerling, F. & Kraehnert, R. (2010). *ACS Nano,* **4**(2), 1076–1082.

Proffen, T., Page, K. L., McLain, S. E., Clausen, B., Darling, T. W., TenCate, J. A., Lee, S.-Y. & Ustundag, E. (2005). *Z. Kristallogr.* **220**(12), 1002–1008.

Ross, M. B., Ku, J. C., Vaccarezza, V. M., Schatz, G. C. & Mirkin, C. A. (2015). *Nat. Nanotech.* **10**(5), 453–458.

Sargent, E. H. (2008). *Adv. Mater.* **20**(20), 3958–3964.

Senesi, A. J. & Lee, B. (2015). *J Appl Cryst, J Appl Crystallogr,* **48**(4), 1172–1182.

Shevchenko, E. V., Talapin, D. V., Kotov, N. A., O'Brien, S. & Murray, C. B. (2006). *Nature,* **439**(7072), 55–59.

Sun, S., Murray, C. B., Weller, D., Folks, L. & Moser, A. (2000). *Science,* **287**(5460), 1989–1992.

Talapin, D. V., Lee, J.-S., Kovalenko, M. V. & Shevchenko, E. V. (2010). *Chem. Rev.* **110**(1), 389–458.

Talapin, D. V. & Murray, C. B. (2005). *Science*, **310**(5745), 86–89.

Talapin, D. V., Shevchenko, E. V., Bodnarchuk, M. I., Ye, X., Chen, J. & Murray, C. B. (2009). *Nature,* **461**(7266), 964–967.

Turkevich, J. & Hubbell, H. H. (1951). *J. Am. Chem. Soc.* **73**(1), 1–7.

Vanmaekelbergh, D. & Liljeroth, P. (2005). *Chem. Soc. Rev.* **34**(4), 299–312.

Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Crystallogr.* **36**(3-1), 860–864.

Wang, Z. L. (2000). *J. Phys. Chem. B*, **104**(6), 1153–1175.

Yager, K. G., Zhang, Y., Lu, F. & Gang, O. (2014). *J Appl Cryst, J Appl Crystallogr*, **47**(1), 118–129.

Young, K. L., Ross, M. B., Blaber, M. G., Rycenga, M., Jones, M. R., Zhang, C., Senesi, A. J., Lee, B., Schatz, G. C. & Mirkin, C. A. (2014). *Adv. Mater.* **26**(4), 653–659.

Zhang, Y., Pal, S., Srinivasan, B., Vo, T., Kumar, S. & Gang, O. (2015). *Nat. Mater.* **14**(8), 840–847.

Zhuang, Z., Peng, Q., Zhang, B. & Li, Y. (2008). *J. Am. Chem. Soc.* **130**(32), 10482–10483.

Zobel, M., Neder, R. B. & Kimber, S. A. J. (2015). *Science*, **347**(6219), 292–294.