

# Combined fitting of small- and wide-angle X-ray total scattering data from nanoparticles: benefits and issues

Anton Gagin, Andrew J. Allen and Igor Levin\*

Materials Measurement Science Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. Correspondence e-mail: igor.levin@nist.gov

Simultaneous fitting of small- (SAS) and wide-angle (WAS) X-ray total scattering data for nanoparticles has been explored using both simulated and experimental signals. The nanoparticle types included core/shell metal and quantum-dot CdSe systems. Various combinations of reciprocal- and real-space representations of the scattering data have been considered. Incorporating SAS data into the fit consistently returned more accurate particle-size distribution parameters than those obtained by fitting the WAS data alone. A popular method for fitting the Fourier transform of the WAS data (*i.e.* a pair-distribution function), in which the omitted SAS part is represented using a parametric function, typically yielded significantly incorrect results. The Pareto optimization method combined with a genetic algorithm proved to be effective for simultaneous SAS/WAS analyses. An approach for identifying the most optimal solution from the Pareto set of solutions has been proposed.

© 2014 International Union of Crystallography

## 1. Introduction

X-ray diffraction is the primary probe of atomic order in nanoparticles. In nanocrystals, significant broadening of Bragg reflections precludes conventional crystallographic refinements (Billinge & Levin, 2007). This problem is addressed by fitting a structural model either to the total (*i.e.* Bragg peaks plus diffuse background) scattering function (Vogel *et al.*, 1997, 2000; Hall, 2000; Zanchet *et al.*, 2000; Kumpf *et al.*, 2005; Kumpf, 2006; Cervellino *et al.*, 2006) or to its Fourier transform, which represents an atomic pair-distribution function (PDF) (Warren, 1969; Dmowski *et al.*, 2002; Hwang *et al.*, 2002; Korsounski *et al.*, 2003; Proffen *et al.*, 2003; Cannas *et al.*, 2004; Billinge & Kanatzidis, 2004; Petkov, 2005; Petkov *et al.*, 2005; Michel *et al.*, 2005; Chaudhuri *et al.*, 2006; Gilbert *et al.*, 2006). The total scattering function,  $S(Q)$ , is related to interatomic distances through the Debye formula:

$$S(Q) - 1 = \left( \frac{1}{N \langle f \rangle^2} \right) \sum_i \sum_{j \neq i} f_i f_j \frac{\sin Q r_{ij}}{Q r_{ij}}, \quad (1)$$

where  $Q$  is the scattering vector magnitude [ $Q = (4\pi/\lambda) \sin \theta$ ,  $\theta$  is half the scattering angle and  $\lambda$  is the wavelength of the incident radiation],  $r_{ij}$  is the distance between atoms  $i$  and  $j$ ,  $N$  is the number of atoms in the system,  $f_i(Q)$  is the scattering factor for atom  $i$ , and  $\langle f \rangle^2$  is the average scattering factor (Egami & Billinge, 2003). The Fourier transform of  $S(Q)$  yields a PDF according to

$$\begin{aligned} G(r) &= (2/\pi) \int_{Q_{\min}}^{Q_{\max}} Q[S(Q) - 1] \sin(Qr) dQ \\ &= 4\pi r [\rho(r) - \rho_0 \gamma_0(r)], \end{aligned} \quad (2)$$

where  $\rho(r)$  is the atomic pair density distribution,  $\gamma_0(r)$  is a factor that accounts for the usually omitted (because of instrumental limitations) small-angle scattering (SAS) at  $Q < Q_{\min}$  and  $\rho_0$  is the spatially averaged atomic number density of the material (Cargill, 1971; Farrow & Billinge, 2009). The atomic scattering factors are constant for neutrons but  $Q$  dependent for X-rays. For neutron scattering,  $\rho(r)$  can be calculated as

$$\rho(r) = \frac{1}{4\pi N r^2 \langle f \rangle^2} \sum_{i \neq j} f_i f_j \delta(r - r_{ij}). \quad (3)$$

In real-space calculations of X-ray PDFs, the atomic scattering factors are assumed to be constant so that equation (3) can be used; commonly, these scattering factors are equated with atomic numbers.

X-ray/neutron scattering over small ( $Q < 1 \text{ \AA}^{-1}$ ) and wide ( $Q > 1 \text{ \AA}^{-1}$ ) angles is measured separately, although some recent neutron instruments, such as NIMROD at the ISIS Facility in the UK (Bowron *et al.*, 2010), cover the entire angular range. In reciprocal space, the Debye equation enables explicit modeling of  $S(Q)$  over any arbitrary  $Q$  range. Modeling of  $G(r)$  obtained using only the wide-angle scattering (WAS) part requires knowledge of  $\gamma_0(r)$ , which represents the Fourier transform of the omitted SAS. In the absence of SAS data,  $\gamma_0(r)$  can be approximated using an analytical function, which has become a common practice in PDF analyses of nanoparticles (Rayleigh, 1914; Glatter & Kratky, 1982; Neder & Korsunskiy, 2005; Korsunskiy & Neder, 2005; Howell *et al.*, 2006; Korsunskiy *et al.*, 2007; Gilbert, 2008). However, the analytical forms of  $\gamma_0(r)$  can only be obtained

for relatively simple particle shape models and, even then, these approximations become inaccurate for very small (a few nanometres) particle sizes (Howell *et al.*, 2006; Mullen & Levin, 2011).

The idea of combining SAS and WAS is attractive because of the overlapping information encoded in these two complementary parts of  $S(Q)$ . While SAS is mostly sensitive to mesoscale nanoparticle characteristics (particle sizes/shapes, intraparticle substructure, agglomeration), WAS is determined by both atomic arrangements and the mesoscale features. Thus, SAS could potentially constrain those mesoscale parameters that become correlated with structural variables (*e.g.* Debye–Waller factors) while fitting the WAS data. One case of combined fitting of SAS and WAS, if measured separately, would involve using the Fourier transform of SAS as  $\gamma_0(r)$  to calculate  $G(r)$ . This is equivalent to fitting  $G^*(r)$ , which is defined as

$$G^*(r) = 4\pi r \rho(r). \quad (4)$$

Alternatively, one could either fit  $S(Q)$  directly over both SAS and WAS ranges or perform joint real- and reciprocal-space fits. Surprisingly, the actual benefits of such combined SAS/WAS refinements of nanoparticle structures have not been evaluated.

In the present study, we compared the accuracy of structural parameters of nanoparticles determined using both separate and combined analyses of SAS and WAS data. In the absence of nanoparticle standards for structure determination, we employed simulated error-free data for neutron scattering generated for several realistic nanoparticle models. In all cases, simultaneous fitting of the SAS and WAS data yielded more accurate estimates of the mesoscale parameters than those obtained using either the WAS data alone or the parametric forms of  $\gamma_0(r)$ . Subsequently, we applied the combined SAS/WAS analyses to the experimental X-ray scattering data collected on CdSe quantum-dot nanoparticles.

## 2. Data simulation and fitting procedures

### 2.1. Nanoparticle models

**2.1.1. Core/shell particles.** We have considered spherical bimetallic nanoparticles that consist of an Au core and a Pd shell, as synthesized, for example, by Ding *et al.* (2010). Both metals exhibit similar face-centered cubic structures with the lattice parameters  $a = 4.08 \text{ \AA}$  (Au) and  $a = 3.89 \text{ \AA}$  (Pd). Each species was assigned its own value of atomic displacement parameter (*i.e.*  $\sigma_{\text{Au}}$  and  $\sigma_{\text{Pd}}$ ). We assumed an epitaxial cube-on-cube relationship between the core and shell lattices with an incoherent core/shell interface. Three distinct models of such composite Au/Pd particles were considered: monodisperse, polydisperse with a lognormal size distribution and monodisperse with a specified fraction of dumbbell-type agglomerates.

**2.1.2. CdSe particles with stacking faults.** Nanoparticles of covalent semiconductor compounds, such as CdSe, are of interest because of their quantum-dot optical properties.

These compounds crystallize with either hexagonal wurtzite or cubic sphalerite structures that differ by a stacking sequence of the close-packed atomic layers. In practice, CdSe nanoparticles commonly exhibit high number densities of stacking faults, which result in the intergrowth of the hexagonal (stacking  $ABAB\dots$ ) and cubic ( $ABCABC\dots$ ) regions (Korsunskiy *et al.*, 2007; Masadeh *et al.*, 2007; Yang *et al.*, 2013). We considered a model of monodisperse CdSe nanoparticles having a wurtzite structure with a given probability of stacking faults. As in practical cases, which will be analyzed later, we introduced distinct scale factors for the SAS and WAS data. The assumption of constant X-ray scattering factors (commonly equated with atomic numbers) used to calculate  $G(r)$  in real space may introduce significant errors. Fortunately, in the case of CdSe, the errors caused by this approximation are negligible, as was proved by comparing  $G(r)$  calculated *via* the Fourier transform of  $S(Q)$  with  $G(r)$  calculated directly in real space.

### 2.2. Simulating experimental data

The ‘experimental’ data for each model were simulated by calculating  $S(Q)$  ( $0 < Q < 30 \text{ \AA}^{-1}$ ) according to equation (1) for a population of 100 nanoparticles. The center of each particle was selected randomly within the unit-cell volume to mimic real situations and to avoid artifacts associated with discontinuous changes in the total number of atoms as the particle size varies during the fit. The ‘experimental’  $G(r)$  was obtained as the Fourier transform of  $S(Q)$  according to equation (2).

### 2.3. Calculating fitted data

The scattering function,  $S(Q)$ , was again calculated using equation (1) for both SAS and WAS ranges.  $G(r)$  was calculated directly in real space according to equations (2) (second part) and (3). The baseline  $\gamma_0(r)$  was calculated either exactly, as the Fourier transform of the SAS part of  $S(Q)$ , or using a closed-form analytical approximation. For the simulated experimental data, we assumed  $Q_{\text{max}}(\text{SAS}) = Q_{\text{min}}(\text{WAS}) \simeq 0.8 \text{ \AA}^{-1}$ . For the real experimental data, we adopted the actual values of  $Q_{\text{max}}(\text{SAS}) = 0.53 \text{ \AA}^{-1}$  and  $Q_{\text{min}}(\text{WAS}) = 0.82 \text{ \AA}^{-1}$ , as determined by the corresponding instrumental setups. The effects of the omitted data for  $Q_{\text{max}}(\text{SAS}) < Q < Q_{\text{min}}(\text{WAS})$  are relatively small, as has been discussed previously (Mullen & Levin, 2011).

Computing  $S(Q)$  is notoriously slow and becomes a limiting factor for particle sizes of only a few nanometres. Here, we adopted the relatively fast computational approach described by Cervellino *et al.* (2006), which was further modified to allow for correlated atomic displacements and the resulting  $r$  dependence of the Debye–Waller factors. According to this approach, the calculations are divided into two tasks: (i) computing the unique interatomic distances,  $d_i$ , and their multiplicities,  $m_i$ , and (ii) computing the Debye sum over the sampled distance range. In cases of ideal uniform crystals having cubic or parallelepiped shapes, the  $m_i$  values can be calculated analytically. However, for crystals either having

more complex shapes/substructures (e.g. core/shell) or containing defects, numerical techniques become necessary.

After  $d_i$  and  $m_i$  have been determined, the Debye sum is computed using a recurrent relation (Cervellino *et al.*, 2006). For polydisperse particles, the structural models were built and stored in the form of spherical layers, which could be added or removed to generate a particle of a given size.

### 3. Experimental data collection

Commercial samples of CdSe nanoparticles dispersed in toluene were obtained from NN-Labs.<sup>1</sup> The particles were covered with octadecylamine (ODA) to prevent agglomeration. The concentration of particles was 2 mg ml<sup>-1</sup>. X-ray total scattering data from these samples were collected at beamline 11-ID-B (Advanced Photon Source, Argonne National Laboratory, IL, USA) using an incident-beam energy of ~90.45 keV ( $\lambda = 0.13702 \text{ \AA}$ ) and a two-dimensional amorphous silicon detector. The samples were sealed in polyamide film capillaries. The intensities scattered from the sample were extremely weak because of the small particle concentrations, and long recording times (7 h)<sup>2</sup> were required. The scattering from toluene (with the appropriate concentration of ODA added) was measured separately (data collection time ~7 h) and subtracted from the signal acquired from the nanoparticle suspension. The intensity diffracted by the particles was corrected for Compton scattering and normalized to yield  $S(Q)$  and  $G(r)$  using the *PDFGetX2* software. The  $Q$  range included in the Fourier transform was limited to  $Q_{\max} = 17 \text{ \AA}^{-1}$  because of the very weak but nevertheless significant (relative to the weak signal from the sample) parasitic scattering that contaminated the data at larger values of  $Q$ .

Pinhole-geometry SAS measurements using a two-dimensional position-sensitive detector were carried out at beamline 15-ID-D (ChemMatCARS, Advanced Photon Source, Argonne National Laboratory, IL, USA). This instrument is described in detail elsewhere (Cookson *et al.*, 2006). The CdSe nanoparticle suspensions were exposed to the X-ray beam in a 1.4 mm-diameter quartz capillary positioned in the incident-beam path. Samples were irradiated with a  $0.5 \times 0.3 \text{ mm}$  monochromatic X-ray beam ( $E = 12 \text{ keV}$ ,  $\lambda = 1.03 \text{ \AA}$ ), and experiments were performed at ambient temperature (~298 K). The data collection times were of the order of a few minutes. The collected data were corrected for parasitic and solvent (toluene) scattering, circularly averaged about the incident-beam direction to yield one-dimensional SAS data, and intensity calibrated against a standard glassy carbon sample. All of the SAS data reduction was carried out within the *Irena* package and modeling macros for *Igor Pro* (Ilavsky & Jemian, 2009; Ilavsky, 2012).

<sup>1</sup> The use of brand or trade names does not imply endorsement of the product by NIST.

<sup>2</sup> The measurement times could be reduced significantly by using higher concentrations of particles in suspensions or letting the toluene evaporate completely. However, in this study, for reference purposes, we wanted to perform the WAS and SAS measurements on identical samples having the same degree of particle agglomeration.

### 4. Optimization approaches for combined analyses

Combined fitting of several data sets requires simultaneous minimization of multiple objective functions. The problem of such multiobjective minimization can be solved using either *a priori* (scalarization) or *a posteriori* (Pareto) methods (Cohon & Marks, 1975).

#### 4.1. Scalarization methods

The weighted-sum method (Zadeh, 1963) is the most common scalarization approach. In this method, a single aggregate objective function (AOF) is formed as a weighted sum of individual objective functions; the weights have to be assigned *a priori*. This AOF is then minimized using a global optimization algorithm. The principal issue with this approach is related to justifying any given selection of weights.

Minimization of the AOF was accomplished using a differential evolution algorithm (DEA; Price *et al.*, 2005) implemented in the *DEoptim* software (Ardia *et al.*, 2011). We used preliminary testing to determine a set of the DEA parameters that enables accurate determination of the structural variables within reasonable computation times. The adopted values of the DEA parameters include the number of population members,  $N_p = 150$ , the number of generations,  $N = 150$ , the crossover probability,  $CR = 0.85$ , and the differential weighting factor,  $F = 0.7$ . For each nanoparticle model and/or type of fit, we performed >50 independent runs to account for the stochastic nature of the DEA. This algorithm was also used for structural refinements using single data sets.

#### 4.2. Pareto methods

Pareto methods search for parameter vectors that simultaneously minimize the collection of individual objective functions without requiring assignment of weights. Evolutionary methods, such as the nondominated sorting genetic algorithm-II (NSGA-II; Deb *et al.*, 2002), the strength Pareto evolutionary algorithm-II (SPEA-II; Zitzler *et al.*, 2001) and the Pareto archived evolution strategy (PAES; Knowles & Corne, 1999), form the most important class of such algorithms. These algorithms use the principle of dominance to select optimal solutions. A solution (*i.e.* a parameter vector)  $x_1$  is said to dominate another candidate solution  $x_2$  if both conditions 1 and 2 are true:

(1) The solution  $x_1$  is no worse than  $x_2$  in all objectives.

(2) The solution  $x_1$  is strictly better than  $x_2$  in at least one objective.

A set of solutions  $P'$  is said to be nondominated by a set of solutions  $P$  if no member of  $P'$  is dominated by any member of  $P$ . NSGA-II and similar algorithms use natural-selection-inspired operations of mutation and crossover to produce a generation composed of candidate parameter vectors and then select the nondominated parameter vectors from this group to survive into the next generation. The set of optimal solutions forms the so-called Pareto front. The advantage of Pareto methods is that no weighting scheme is needed. We used the NSGA-II algorithm implemented in the package *mco*

**Table 1**

Estimates of parameters for the Au/Pd nanoparticles obtained by fitting different types of the simulated data.

The fits of  $S^{\text{SAS}}(Q)$  alone have been performed using the parametric functions. Subscripts C and S denote the core and shell, respectively. The column labeled Shell specifies the shell thickness ( $r - r_C$ ).

Model	Fitting scheme	$a_C$ (Å)	$a_S$ (Å)	$r_C$ (Å)	$r$ (Å)	Shell (Å)	$\ln \mu$	$p$	$\sigma_C$ (Å <sup>2</sup> )	$\sigma_S$ (Å <sup>2</sup> )
Monodisperse particles	True value	4.08	3.89	8	10	2	–	–	0.02	0.04
	$S^{\text{SAS}}(Q)$	Fixed	Fixed	7.8	10	2.2	–	–	–	–
	$\log[S^{\text{SAS}}(Q)]$	Fixed	Fixed	8.9	9.8	0.9	–	–	–	–
	$G^*(r)$	4.080 (1)	3.890 (2)	7.98 (5)	10.00 (2)	2.01 (5)	–	–	0.0199 (3)	0.041 (1)
	$S^{\text{WAS}}(Q)$	4.081 (2)	3.891 (3)	7.96(7)	9.98 (3)	2.02 (8)	–	–	0.0198 (4)	0.041 (2)
Polydisperse particles	True value	4.08	3.89	11.2	14	2.8	0.049	–	0.02	0.04
	$S^{\text{SAS}}(Q)$	Fixed	Fixed	10.1	14.0	3.9	0.054	–	–	–
	$\log[S^{\text{SAS}}(Q)]$	Fixed	Fixed	11.9	13.9	2.0	0.051	–	–	–
	$G^*(r)$	4.0798 (4)	3.891 (1)	11.13 (8)	13.92 (9)	2.79 (5)	0.036 (8)	–	0.0199 (2)	0.040 (1)
	$G^*(r) + [\text{parametric } \gamma_0(r)]$	4.078 (1)	3.883 (1)	11.12 (4)	13.64 (6)	2.52 (4)	0.005 (4)	–	0.0205 (2)	0.035 (1)
	$G^*(r) + \log[S^{\text{SAS}}(Q)]$	4.08 (2)	3.9 (2)	11.2 (4)	13.9 (4)	2.7 (1)	0.043 (1)	–	0.021 (5)	0.04 (1)
	$G^*(r) + S^{\text{SAS}}(Q)$	4.0795 (10)	3.890 (1)	11.15 (7)	13.9 (1)	2.79 (5)	0.037 (4)	–	0.0199 (2)	0.039 (1)
	$S^{\text{WAS}}(Q)$	4.0804 (8)	3.891 (1)	11.09 (8)	13.9 (1)	2.81 (4)	0.034 (5)	–	0.0198 (3)	0.040 (1)
	Agglomerated particles	True value	4.08	3.89	8	10	2	–	0.39	0.02
$S^{\text{SAS}}(Q)$		Fixed	Fixed	7.8	9.8	2	–	0.42	–	–
$\log[S^{\text{SAS}}(Q)]$		fixed	fixed	9.2	9.9	0.7	–	0.44	–	–
$G^*(r)$		4.075	3.887	8.35	10.25	1.9	–	0.14	0.0206	0.038
$G^*(r) + S^{\text{SAS}}(Q)$		4.073	3.887	8.5	10.4	1.95	–	0.36	0.0205	0.042

(Trautmann *et al.*, 2013), which was parallelized to achieve acceptable computing times.

The major difficulty with Pareto methods is selecting the best possible set of parameters from a multitude of the optimal Pareto solutions. A rigorous answer to this problem is currently unavailable (Coello Coello *et al.*, 2007). In fact, selection of a single point from a convex Pareto front is equivalent to solving a scalarized problem with specific weights assigned to individual data sets. Indeed, simultaneous minimization of the two functions  $f(x)$  and  $g(x)$  with respect to a parameter  $x$  is equivalent to minimization of a set of functions (*e.g.* Das & Dennis, 1997)

$$\{J_w(x) = (1 - w)f(x) + wg(x)\}_{w \in (0;1)}. \tag{5}$$

From this set, one can select a function that exhibits the least sensitivity to small changes in the input and, therefore, can be assumed to correspond to the optimal choice of weight,  $w$ . Concurrently, such a function exhibits the highest sensitivity to the fitted parameters at the point of its minimum. We suggest that the best estimate of the parameter  $x$ ,  $x^*$ , can be obtained by minimizing the following expression:

$$\left| \frac{\delta x^*}{\delta w} \right| w(1 - w), \tag{6}$$

where  $x^* = x^*(w)$  is the parameter estimate corresponding to the global minimum of  $J_w(x^*)$ . The optimal value of  $w$  corresponds to the sharpest minimum of  $J_w(x)$  at the point  $x = x^*(w)$  for which the experimental functions  $f(x^*)$  and  $g(x^*)$  are the closest to their respective minima. The scattering function,  $S(Q)$ , depends on several variables and, therefore, the above-made arguments are not directly applicable to the analyses of the diffraction data. Nevertheless, we found minimization of

expression (6) effective even for the types of data used in the present study.

We introduce the following change of variables,

$$s = \begin{cases} \ln(w) + \ln(1 - w), & 0 < w \leq 0.5, \\ -\ln(w) - \ln(1 - w) - 4 \ln 2, & 0.5 < w < 1, \end{cases} \tag{7}$$

so that  $|x^{*'}(s)| = |x^{*'}(w)|w(1 - w)$ . Each of the fitted parameters,  $x$ , is then plotted as a function of  $s$ . The value of  $x^*$  is assumed to correspond to the minimum of the standard deviation of  $x(s)$ ; we determined this standard deviation,  $\sigma_x$ , using the five nearest points in the plot of  $x(s)$ .

## 5. Results and discussion

### 5.1. Simulated data

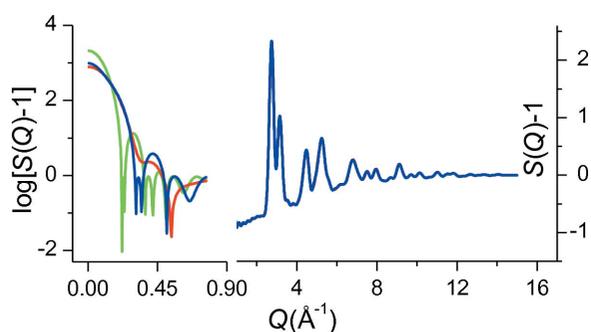
**5.1.1. Parametric SAS fits.** Parameterized SAS models (in  $Q$  space) are available for the classes of core/shell nanoparticle models studied here. We fitted these models to the  $S^{\text{SAS}}(Q)$  data for the Au/Pd nanoparticles using either linear or logarithmic intensity scales for the estimates of the objective function. In all cases, the lattice parameters for the core and shell components were kept fixed at their correct values. For the monodisperse non-agglomerated particles, using the linear scale yields accurate estimates of the particle size and core/shell dimensions (Table 1). Fitting the same models on the logarithmic scale returns an incorrect value for the core radius. For the polydisperse particles, the estimates of parameters are far less accurate compared to the monodisperse case, with better results obtained by fitting on the logarithmic scale; the standard deviation for the core radius,  $\mu$ , is reproduced relatively well. Conversely, for the monodisperse but agglomerated particles, all the parameters, including the probability of agglomerates, are reproduced satisfactorily for the linear-scale

fits, while fitting on the logarithmic scale returns a significantly incorrect value of the core radius.

The fitting scale (linear *versus* logarithmic) affects the accuracy of the refined parameters because different nanoparticle characteristics are manifested most prominently over different  $Q$  ranges, which may have disparate intensity scales. For example, the difference between the SAS data for the mono- and polydisperse particles is concentrated primarily at  $Q > 0.2 \text{ \AA}^{-1}$ , for which the intensities are relatively weak; thus, a logarithmic scale that emphasizes this part of the data over the lower- $Q$  part is preferred. The opposite is true for the agglomerate particles, for which most of the difference occurs at low  $Q$ .

**5.1.2. Nonparametric fits of SAS and WAS.** Both the  $S^{\text{WAS}}(Q)$  and the corresponding  $G(r)$  functions are only weakly affected by the particle-size distributions (Figs. 1 and 2). Studies of nanoparticle structures using  $G(r)$  commonly employ approximate analytical descriptions of  $\gamma_0(r)$ . Fig. 3 illustrates the discrepancies between the analytical expression and the actual  $\gamma_0(r)$ . We compared the structural parameters obtained using both the analytical and the exact (*i.e.* the Fourier transform of SAS) forms of  $\gamma_0(r)$ , while fitting  $G(r)$  for the polydisperse core/shell particles. The fit using the exact  $\gamma_0(r)$  returned accurate estimates of all the parameters, apart from  $\mu$  (a dispersion of the core size), which exhibited a small, albeit significant, bias from its true value. In contrast, all the parameters determined using the analytical  $\gamma_0(r)$  were grossly incorrect, falling outside of the corresponding histograms (Fig. 4); importantly, these fits failed to reveal the polydisperse nature of the particles. Thus, the analytical  $\gamma_0(r)$  can introduce significant errors into the parameters refined from  $G^*(r)$  and even lead to adoption of a qualitatively incorrect model (*e.g.* monodisperse instead of polydisperse). All of the examples described below incorporate the exact form of  $\gamma_0(r)$ .

**Combined fits using an aggregate objective function.** As mentioned above, fitting  $G^*(r)$  for the polydisperse Au/Pd particles returns a biased estimate of the parameter  $\mu$ . A similar result is obtained by simultaneous fitting of  $G^*(r)$  and  $S^{\text{SAS}}(Q)$ . In contrast, simultaneous fitting of  $G^*(r)$  and  $\log[S^{\text{SAS}}(Q)]$  (Table 1) significantly improves the accuracy of the refined value of  $\mu$ , which again highlights the importance

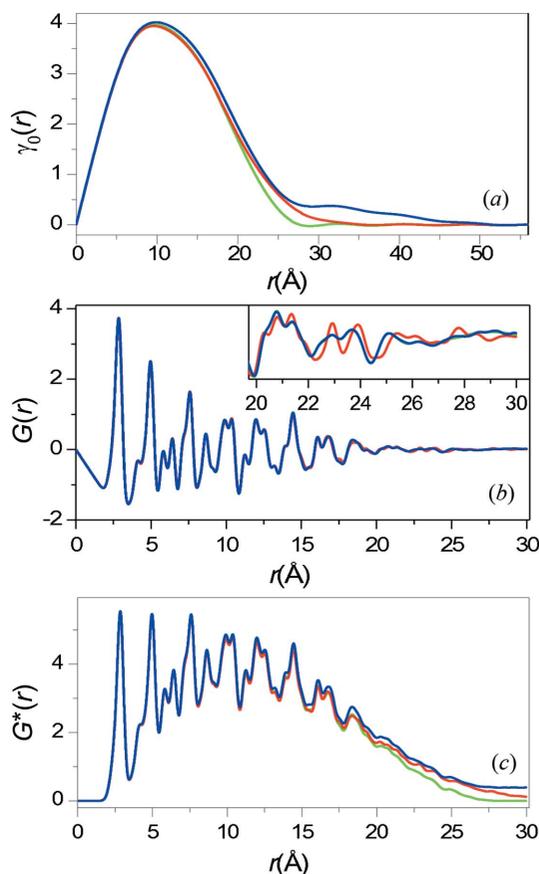


**Figure 1**

The total scattering function  $S(Q)$  generated for the monodisperse (green line), polydisperse (red line) and agglomerated (40% of dimers, blue line) Au/Pd particles over the SAS and WAS  $Q$  ranges. The differences are noticeable only in the SAS range.

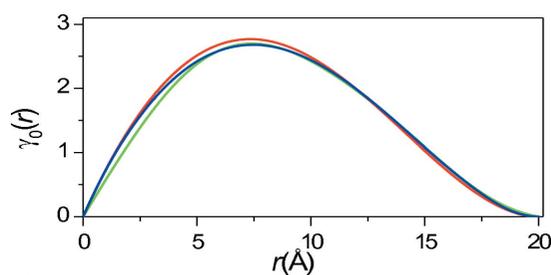
of selecting the appropriate scale (*i.e.* linear or logarithmic) for fitting  $S^{\text{SAS}}(Q)$ . Note that  $G^*(r)$  incorporates the Fourier transform of  $S^{\text{SAS}}(Q)$  on the linear scale, which explains why fitting  $G^*(r)$  alone yields inaccurate estimates of  $\mu$ .

Agglomerates that consist of randomly oriented particles have a relatively weak effect on  $S^{\text{WAS}}(Q)$  and  $G(r)$  (Figs. 1 and 2). Fitting  $G^*(r)$  alone returns a nonzero but significantly incorrect value of the probability of agglomerates,  $p$ , while other structural and particle-size parameters are reproduced



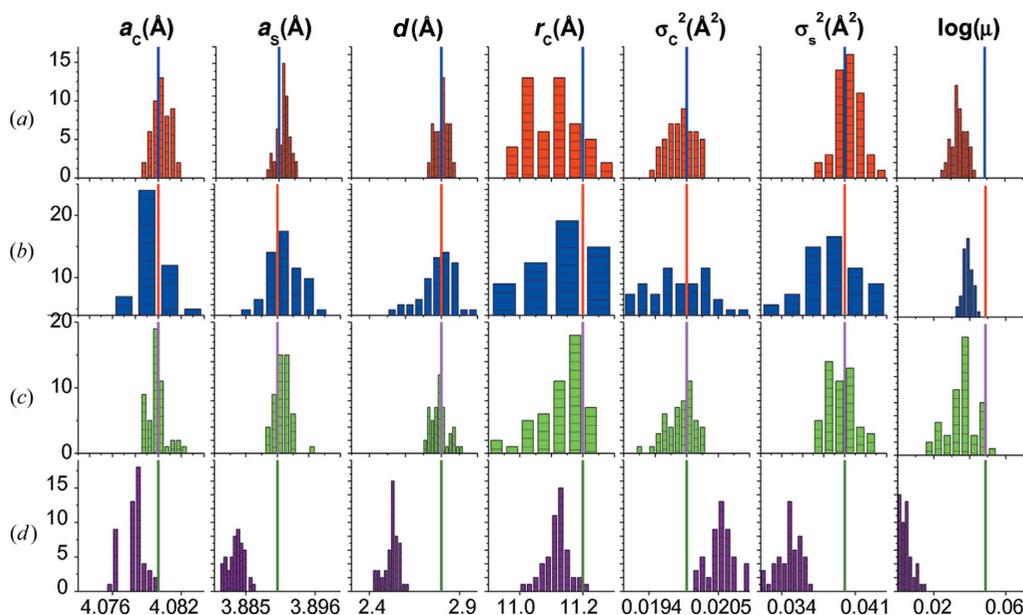
**Figure 2**

Real-space representations of the total scattering data generated for the monodisperse (green line), polydisperse (red line) and agglomerated (40% of dimers, blue line) Au/Pd particles: (a)  $\gamma_0(r)$ , (b)  $G(r)$  and (c)  $G^*(r)$ .



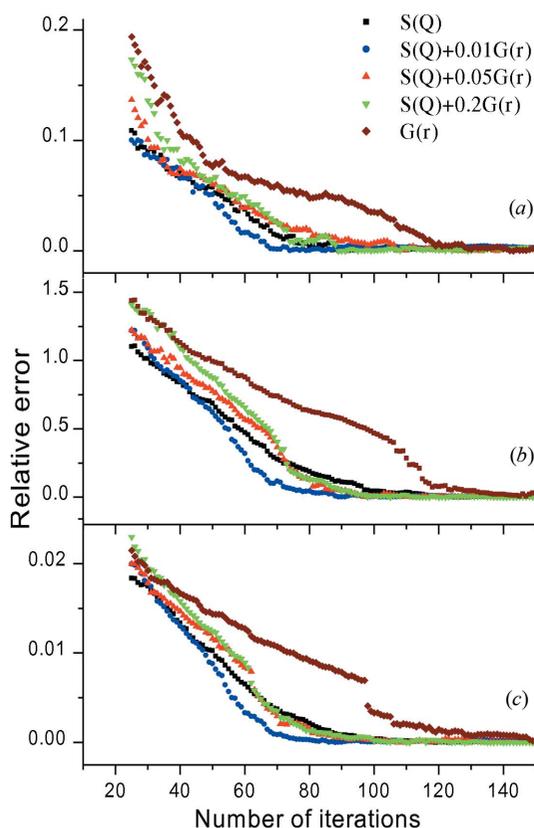
**Figure 3**

Comparison of the exact  $\gamma_0(r)$  with its different approximations for the polydisperse Au/Pd nanoparticles: (green) the Fourier transform of the SAS data, which corresponds to the exact  $\gamma_0(r)$ , (blue) the spherical core/shell model (Guinier & Fournet, 1955) and (red) the spherical uniform-particle model (Glatter, 1979) adapted for core/shell particles.



**Figure 4** Histograms of the parameter estimates for the polydisperse Au/Pd particles obtained by fitting (a)  $S^{\text{WAS}}(Q)$ , (b)  $G^*(r) + S^{\text{SAS}}(Q)$ , (c)  $G^*(r)$  with the  $\gamma_0(r)$  term described as the Fourier transform of  $S^{\text{SAS}}(Q)$  and (d)  $G^*(r)$  with the analytical  $\gamma_0(r)$ . The vertical stick marks indicate the true parameter values.

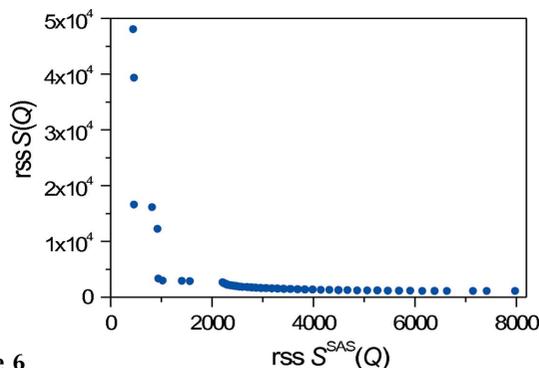
satisfactorily (Table 1). By contrast, a combined fitting of  $G^*(r)$  and  $S^{\text{SAS}}(Q)$  yields accurate values of all the parameters,



**Figure 5** Dependence of the relative error in the fitted parameters for the Au/Pd nanoparticles on the number of iterations for different fitting schemes: (a) core lattice parameter  $a_c$ , (b) core radius  $r_c$  and (c) core Debye–Waller factor  $\sigma_c^2$ . The weights assigned to  $G^*(r)$  in the combined fits of  $G^*(r) + S^{\text{WAS}}(Q)$  were 0.01, 0.05 and 0.2, as indicated in the legend.

including the probability  $p$ . In principle, this probability can be recovered by fitting  $S^{\text{SAS}}(Q)$  using an analytical function. Thus, in cases for which such analytical expressions are available, the most effective approach would be to determine the agglomeration parameter from a separate fitting of  $S^{\text{SAS}}(Q)$  and then assume non-agglomerated particles while fitting either  $G^*(r)$  or  $S^{\text{WAS}}(Q)$ .

Our results suggest that fitting in real space results in a slower convergence to the correct values of parameters than fitting in reciprocal space (Fig. 5). For the combined fits, the rate of convergence depends strongly on the weights used: that is, a non-optimal choice of weights can significantly slow the convergence. We also note that fitting the SAS data using up to 200 particles to calculate the signal still results in significant uncertainties, which then project into larger uncertainties in the estimates of the refined variables from the combined SAS/WAS analyses. Obviously, the size of particle populations used in the fitting is limited by the computing time.



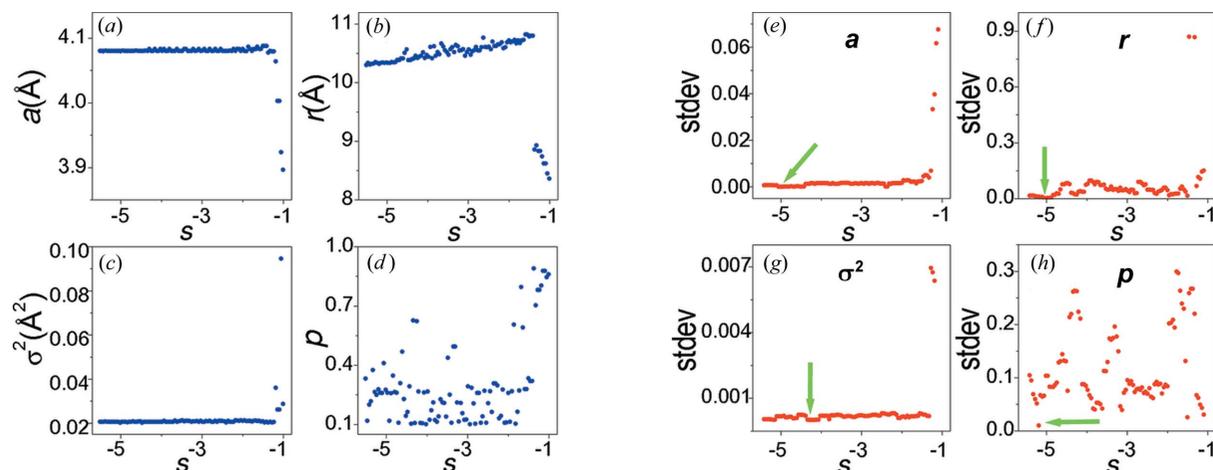
**Figure 6** The Pareto front of optimal solutions for a combined fit of  $S^{\text{SAS}}(Q) + S^{\text{WAS}}(Q)$  for the monodisperse agglomerated Au nanoparticles. The plot illustrates correlations between residual sums of squares (rss) for the WAS and SAS data.

**Table 2**

Estimates of structural and particle-size parameters obtained for the CdSe model using the simulated data and the Pareto algorithm.

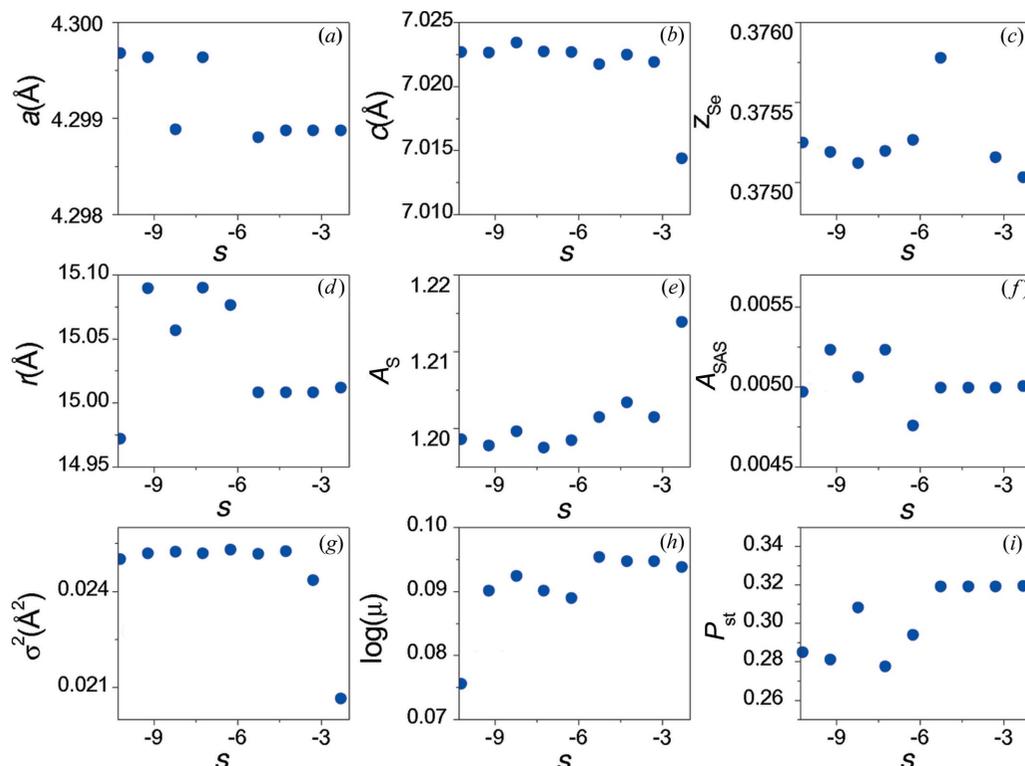
The labels Pareto min and Pareto max denote limiting values of the parameter estimates for the entire set of Pareto solutions. The solution with the minimum standard deviation is indicated as ‘Pareto best’. The label ‘ $S^{\text{WAS}}(Q)$  @  $S^{\text{SAS}}(Q)$ ’ denotes a fit of  $S^{\text{WAS}}(Q)$  where the parameters of the particle-size distribution were kept fixed at the values determined using the parametric fit of  $S^{\text{SAS}}(Q)$ .

	$a$ (Å)	$c$ (Å)	$Z_{\text{Se}}$	$r$ (Å)	$A_s$	$A_{\text{SAS}}$	$\delta$ (Å <sup>2</sup> )	$\sigma$ (Å <sup>2</sup> )	$\ln \mu$	$p_{\text{st}}$	bkg
True value	4.3	7.02	0.375	15	1.2	0.005	6.5	0.025	0.0953	0.3	0.004
Pareto min	4.2988	7.014	0.3729	14.97	1.197	0.0048	5.65	0.0206	0.0755	0.25	0.0036
Pareto max	4.3002	7.024	0.3758	15.09	1.214	0.0052	6.55	0.0253	0.0953	0.32	0.0045
Pareto best	4.3000 (5)	7.023 (1)	0.3752 (1)	15.01 (3)	1.199 (1)	0.0050 (1)	6.50 (2)	0.0253 (1)	0.0946 (25)	0.32 (1)	0.0040 (0)
$S^{\text{WAS}}(Q)$ @ $S^{\text{SAS}}(Q)$	4.3002 (5)	7.023 (1)	0.3750 (3)	15.05	1.201 (3)	0.0053	6.51 (3)	0.0249 (2)	0.0974	0.30 (1)	0.0044



**Figure 7**

(a)–(d) The Pareto set of optimal solutions for the parameters of the monodisperse agglomerated Au particles obtained from the combined fit of  $S^{\text{SAS}}(Q) + S^{\text{WAS}}(Q)$ ; (e)–(h) the corresponding standard deviations (stdev) of these parameters. The variable  $s$  is defined according to equation (7). The arrows indicate the ‘best’ Pareto solutions defined according to the criterion of the minimal standard deviation.



**Figure 8**

The Pareto set of optimal solutions for the parameters of the CdSe nanoparticles obtained using a combined fit of the simulated  $S^{\text{SAS}}(Q) + S^{\text{WAS}}(Q)$  data.

**Combined fits using the Pareto algorithm.** We applied the Pareto algorithm to simultaneous fitting of  $G^*(r)$  and  $S^{\text{SAS}}(Q)$  for the monodisperse agglomerated Au particles. The resulting Pareto front and the set of Pareto solutions are presented in Figs. 6 and 7, respectively. Analyses of the dependencies  $x(s)$  suggest the following best parameter estimates:  $a = 4.0805 \text{ \AA}$ ,  $r = 10.3 \text{ \AA}$ ,  $\sigma^2 = 0.0205 \text{ \AA}^2$ ,  $p = 0.28$ . These estimates are very close to the corresponding target values:  $a = 4.08 \text{ \AA}$ ,  $r = 10 \text{ \AA}$ ,  $\sigma^2 = 0.02 \text{ \AA}^2$ ,  $p = 0.3$ .

The results of Pareto analysis for the CdSe nanoparticles are summarized in Table 2 and Figs. 8 and 9. Our approach yields fairly accurate estimates of all the parameters despite having only nine points in the Pareto front. The structural parameters obtained by fitting  $S^{\text{WAS}}(Q)$  with the values of  $r$  and  $\mu$  fixed according to a parametric fit of  $S^{\text{SAS}}(Q)$  are also close to the target values. Thus, in the case of error-free data, a sequential SAS/WAS fit appears to be effective. However, in practice, given the presence of systematic errors, a combined fit using the Pareto algorithm is expected to produce more accurate and reliable results than the sequential fit.

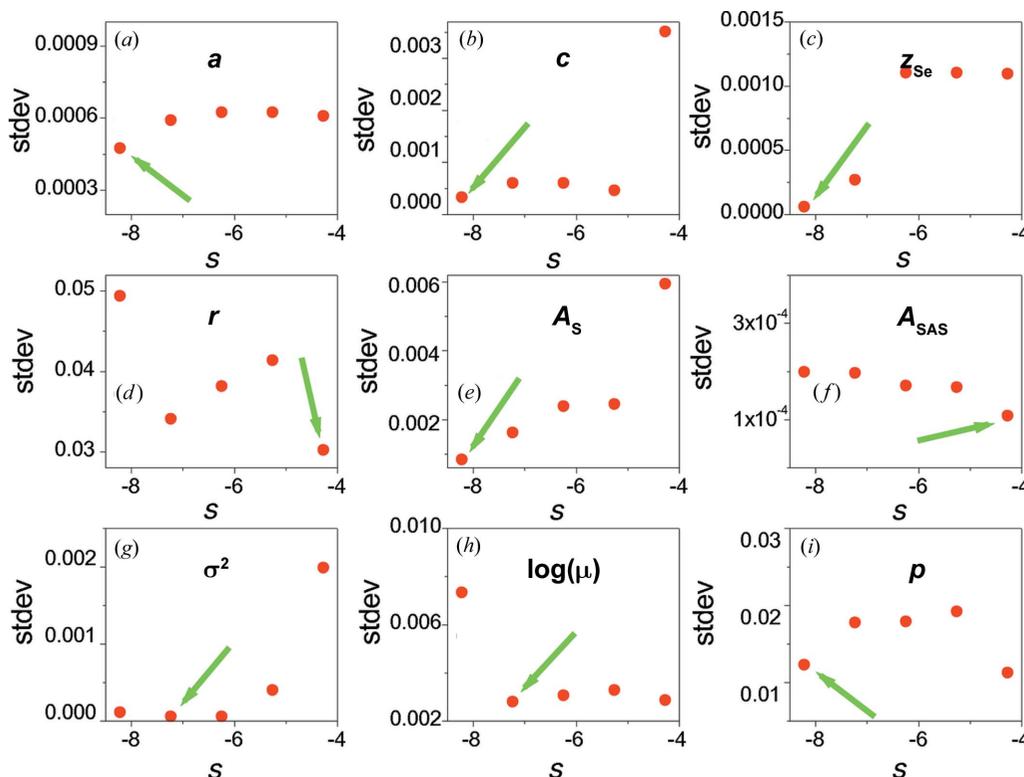
**5.1.3. Software and hardware.** The tools for generating nanoparticle models and calculating total scattering data, as well as the wrapper functions for optimization libraries, have been implemented in the R package *nanop* (Gagin & Mullen, 2013), available for noncommercial use under the terms of the GNU General Public License. Structural refinements have been performed using an SMP machine with 48 2500 MHz AMD Opteron CPUs. The computing time varied from several hours to one week, depending on the problem.

## 5.2. Experimental data

The experimental data for the CdSe nanoparticles were fitted using a model of wurtzite structure with stacking faults. The WAS data were corrected for the contributions of scattering from the toluene and ODA. However, for the SAS data, the contribution from ODA was not subtracted. Modeling the effects of ODA, while using the Debye equation to calculate the scattering function, is challenging. Fortunately, the X-ray scattering power of ODA is significantly less than the scattering from CdSe. Parametric fits of  $S^{\text{SAS}}(Q)$  that account for the presence of the ODA shell yield a CdSe particle size of  $r = 25.7 \text{ \AA}$ , as opposed to  $r = 26.2 \text{ \AA}$  determined by assuming polymer-free CdSe particles. This systematic error proved to be insignificant relative to the discrepancy between the particle sizes obtained with and without the SAS data (see below).

The structural variables included the hexagonal lattice parameters ( $a$ ,  $c$ ), the fractional  $z$  coordinate of the Se atoms, the atomic displacement parameters (constrained to be equal for Cd and Se), a correlation parameter for the Cd and Se displacements ( $\delta$ ), and the probability of stacking faults,  $p_{\text{sf}}$ . The nonstructural parameters included the mean particle radius ( $r$ ), its standard deviation ( $\mu$ ), scale factors for the SAS ( $A_{\text{SAS}}$ ) and WAS ( $A_{\text{S}}$ ) data, and a constant background value (bkg) for the SAS signal.

The DEA parameters were the same as described for the simulated data. The particle-size distribution and the occurrence of stacking faults were modeled by calculating the signal for over 50 particles and then comparing the average with the



**Figure 9**

Standard deviations (stdev) for the parameters of the CdSe nanoparticles for the Pareto set of optimal solutions shown in Fig. 9. The arrows indicate the 'best' Pareto solutions defined according to the criterion of the minimal standard deviation.

**Table 3**

Structural and particle-size parameters for the CdSe nanoparticles obtained by fitting the experimental data.

The label ‘ $S^{\text{WAS}}(Q) @ S^{\text{SAS}}(Q)$ ’ denotes a fit of  $S^{\text{WAS}}(Q)$  with the parameters  $r$  and  $\mu$  fixed at the values determined from the parametric  $S^{\text{SAS}}(Q)$  fit. In the combined scalarized fits the weight of  $S^{\text{SAS}}(Q)$  was set to 0.05, based on the analyses of the simulated data. The results quoted for the Pareto fit were extracted from the Pareto sets using the minimal standard deviation procedure.

Fitting scheme	$a_c$ (Å)	$a_s$ (Å)	$z$	$r$ (Å)	$A_s$	$\delta_2$	$\sigma^2$ (Å <sup>2</sup> )	$p_{\text{sf}}$	$\mu$	$A_{\text{SAS}}$	bkg
$S^{\text{WAS}}(Q)$	4.2964 (10)	7.014 (7)	0.3785 (9)	22.1 (5)	1.06 (1)	6.49 (6)	0.0265 (10)	0.33 (2)	0.080 (15)	–	–
$S^{\text{WAS}}(Q) @ S^{\text{SAS}}(Q)$	4.2964 (8)	7.007 (6)	0.3770 (8)	26.5 (1)	1.02 (1)	6.58 (12)	0.0274 (11)	0.33 (1)	0.088 (1)	–	–
$S^{\text{WAS}}(Q) + S^{\text{SAS}}(Q)$	4.2963 (28)	7.014 (6)	0.3769 (11)	26.3 (2)	1.03 (1)	6.49 (13)	0.0287 (9)	0.31 (1)	0.093 (3)	0.00656 (40)	0.0046 (3)
$G^*(r)$ [parametric $\gamma_0(r)$ ]	4.2948 (2)	6.999 (1)	0.3779 (4)	20.2 (2)	1.19 (1)	6.39 (1)	0.0253 (1)	0.33 (2)	0.075 (17)	–	–
$G^*(r)$	4.2951 (19)	7.001 (4)	0.3772 (10)	26.0 (2)	1.12 (1)	6.49 (2)	0.0260 (5)	0.33 (1)	0.091 (5)	0.00656 (2)	–
$G^*(r) + S^{\text{SAS}}(Q)$	4.2951 (20)	7.005 (20)	0.3774 (12)	26.2 (2)	1.12 (2)	6.46 (10)	0.0264 (10)	0.34 (2)	0.094 (7)	0.00649 (2)	0.0046 (1)
Pareto fit	4.2950 (2)	7.015 (2)	0.3790 (2)	26.0–26.5	1.051 (4)	6.48 (9)	0.0256–0.0263	0.325 (5)	0.0108 (3)	0.0068 (1)	0.0048 (1)

experimental data. For each type of fit, we performed five independent DEA runs. For the Pareto fits, the number of population members  $N_p$  was set at 348 and the number of generations  $N$  at 200 to obtain a sufficient number of points in the Pareto front.

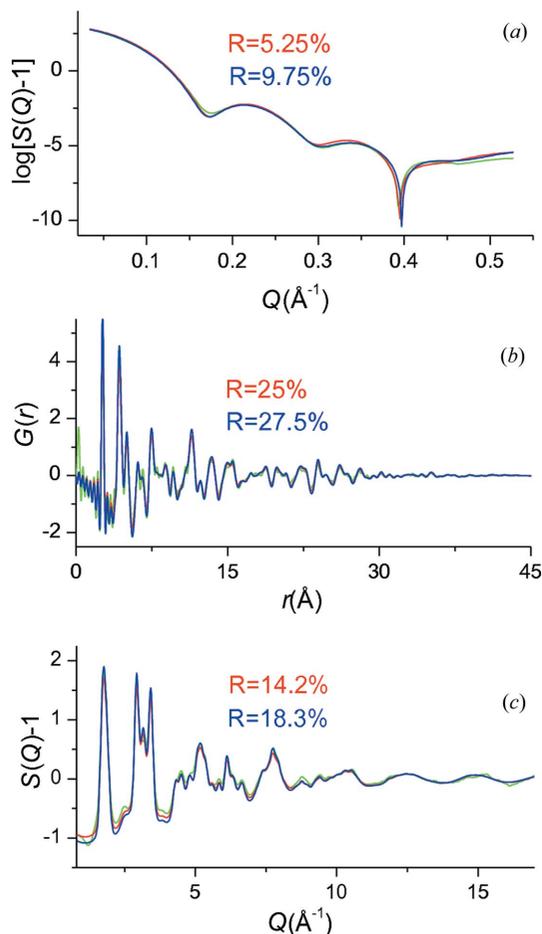
The following types of fits were used: (1)  $S^{\text{WAS}}(Q)$ , (2)  $S^{\text{WAS}}(Q)$  with particle-size parameters  $r$  and  $\mu$  fixed at values

determined from the parametric SAS fits, (3)  $G(r)$  with the  $\gamma_0(r)$  term described as the Fourier transform of  $S^{\text{SAS}}(Q)$ , (4)  $G(r)$  with the  $\gamma_0(r)$  term described using an analytical function, (5) combined  $S^{\text{SAS}}(Q) + S^{\text{WAS}}(Q)$ , and (6) combined  $G(r) + S^{\text{SAS}}(Q)$ . The Pareto algorithm was used for fit (6). The results are summarized in Table 3.

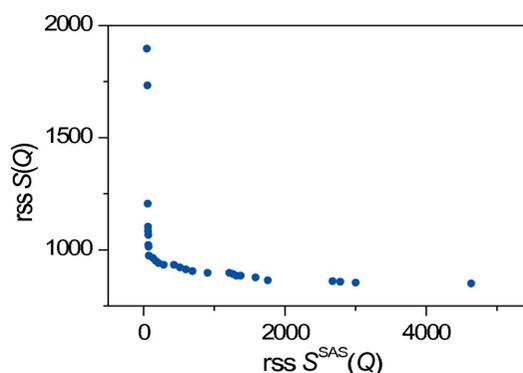
In all cases, the agreement between the experimental and calculated signals was satisfactory (Fig. 10). The principal difference between the fits that include the SAS data and those that rely on the WAS data alone is in the estimate of the particle size (Table 3 and Figs. 11–13). All the combined fits converge on  $r = 26$  (1) Å, whereas fits (1) and (4) yield  $r = 22$  (1) Å; the differences in the other parameters appear to be insignificant relative to the uncertainties estimated from the five fits. Overall, these results suggest that including the SAS data in the fit (either in  $Q$  or  $r$  space) is advantageous and yields significantly more accurate particle sizes than fits of the WAS data alone. The combined fitting using an aggregate objective function leaves the question of the optimal data weighting open, but the Pareto algorithm combined with the analysis developed here addresses this problem. Comparison of residual values for the different types of fits again confirms a better convergence of the reciprocal-space fits.

**6. Conclusions**

According to our results, SAS data are necessary for accurate determination of the nanoparticle size, agglomeration parameters *etc.* Reciprocal-space fits converge faster to the correct



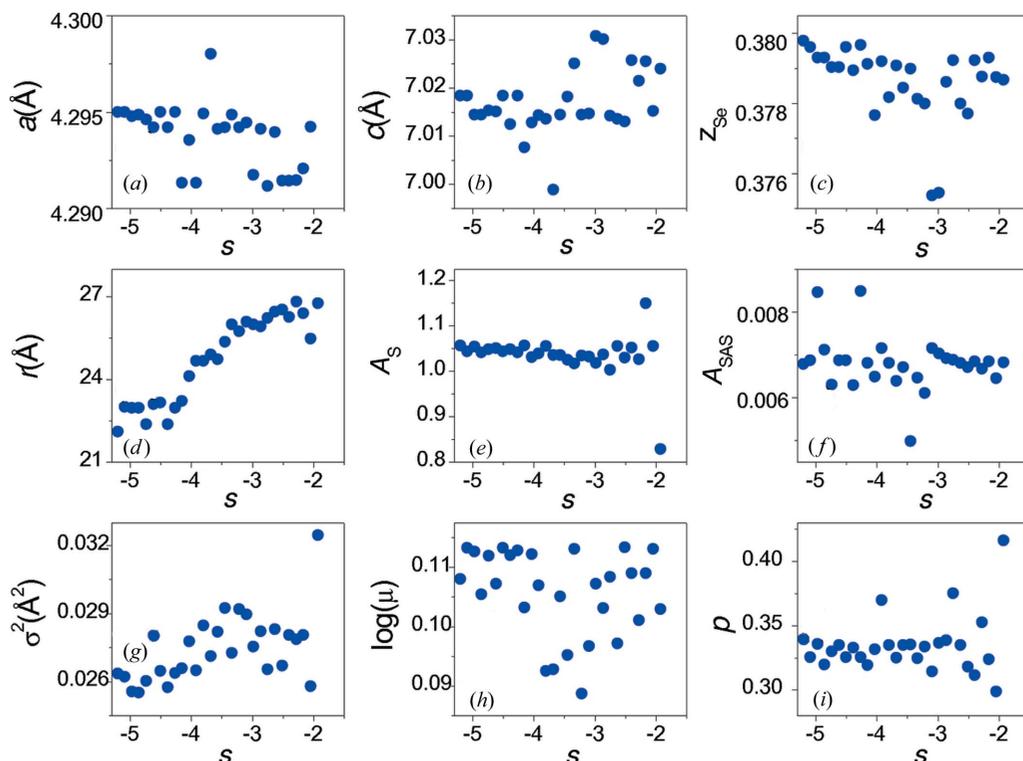
**Figure 10** Experimental (green) and calculated (red, blue) signals for the CdSe nanoparticles: (a)  $S^{\text{SAS}}(Q)$ , (b)  $G^*(r)$  and (c)  $S^{\text{WAS}}(Q)$ . The calculated signals correspond to the most optimal nanoparticle parameters obtained from the Pareto fits of the  $S^{\text{WAS}}(Q) + S^{\text{SAS}}(Q)$  (red) and  $G(r) + S^{\text{SAS}}(Q)$  (blue) data. The  $R$  factors that characterize the quality of each fit are indicated in the plots.



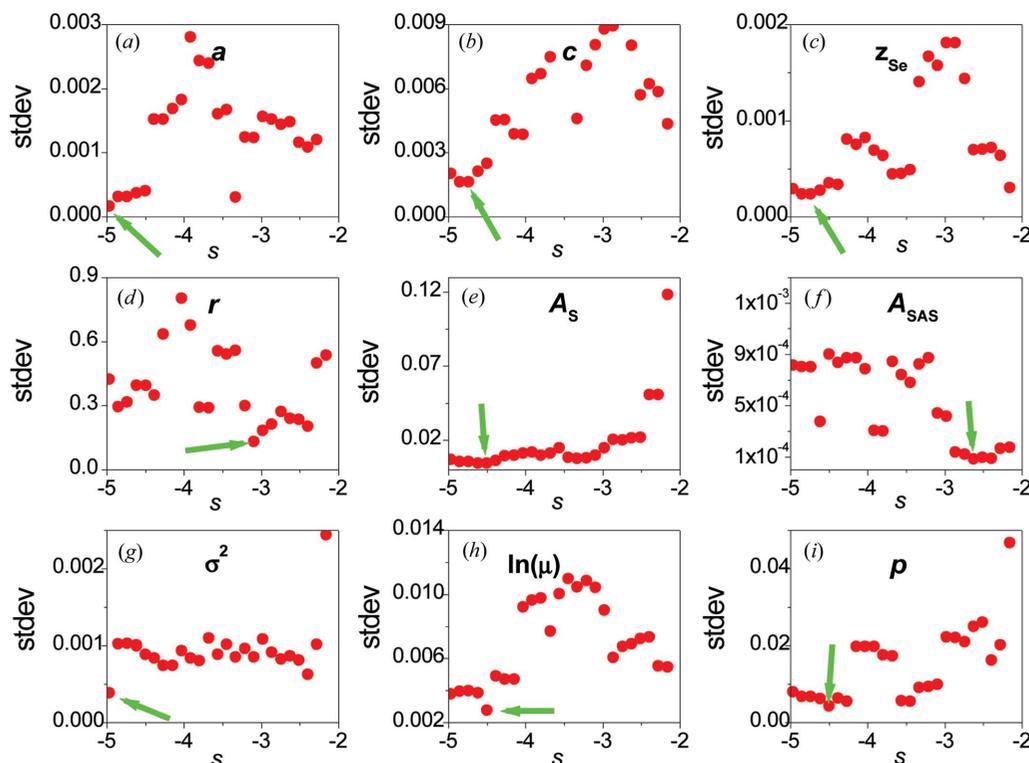
**Figure 11** The Pareto front of optimal solutions obtained from the combined fit of the experimental  $S^{\text{SAS}}(Q) + S^{\text{WAS}}(Q)$  for the CdSe nanoparticles.

values of parameters than the real-space fits and, therefore, appear to be more robust. Real-space fits of the atomic PDF using the Fourier transform of the SAS data to describe the

baseline consistently produce more accurate results than those obtained using approximate analytical expressions. For the nanoparticle types considered in this work, which included



**Figure 12**  
The Pareto set of optimal solutions for the parameters of the CdSe nanoparticles obtained from the combined fit of the experimental  $S^{SAS}(Q) + S^{WAS}(Q)$ .



**Figure 13**  
Standard deviations (stdev) of the parameters of the CdSe nanoparticles for the Pareto set of optimal solutions shown in Fig. 12. The arrows indicate the ‘best’ Pareto solutions defined according to the criterion of the minimal standard deviation.

composite metal and quantum-dot CdSe nanoparticles in dilute suspension, the structural parameters appeared to be only weakly correlated with the size/agglomeration features and therefore could be accurately determined from fitting the WAS data alone. However, in real situations where the nanoparticle concentration and associated SAS structure factor term may be more significant, the degree of correlation between these parameters is unknown. Therefore, simultaneous fitting of WAS and SAS data will provide more reliable values of the parameter estimates than fitting these data sets separately. Recording both SAS and WAS in a single data set would facilitate the combined analyses. The Pareto methods proved to be effective solutions for the combined-technique structural refinements and appear advantageous over more traditional approaches that minimize the aggregate objective functions with weights assigned to individual data sets. We proposed a promising scheme for selecting a set of optimal parameters from the results of the Pareto analyses. This scheme has been first validated using the simulated data and then applied to the analysis of the experimental results.

This study has been initiated in collaboration with K. Mullen (UCLA). Use of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE), Office of Science, by Argonne National Laboratory, was supported by the US DOE under contract No. DE-AC02-06CH11357. ChemMatCARS APS sector 15 is principally supported by the National Science Foundation/Department of Energy under grant No. NSF/CHE-0822838. The authors are grateful to K. Chapman (APS) and P. Chupas (APS) for their technical assistance with the X-ray total scattering measurements, and to J. Ilavsky (APS) and Y.-S. Chen (ChemMatCARS) for their assistance with the SAS measurements.

## References

- Ardia, D., Mullen, K., Peterson, B. & Ulrich, J. (2011). *DEoptim*. R Package Version 2.2-2. <http://CRAN.R-project.org/package=DEoptim>.
- Billinge, S. J. L. & Kanatzidis, M. G. (2004). *Chem. Commun.* pp. 749–760.
- Billinge, S. J. & Levin, I. (2007). *Science*, **316**, 561–565.
- Bowron, D. T., Soper, A. K., Jones, K., Ansell, S., Birch, S., Norris, J., Perrott, L., Riedel, D., Rhodes, N. J., Wakefield, S. R., Botti, A., Ricci, M. A., Grazi, F. & Zoppi, M. (2010). *Rev. Sci. Instrum.* **81**, 033905.
- Cannas, C., Musinu, A., Navarra, G. & Piccaluga, G. (2004). *Phys. Chem. Chem. Phys.* **6**, 3530–3534.
- Cargill, G. S. (1971). *J. Appl. Cryst.* **4**, 277–283.
- Cervellino, A., Giannini, C. & Guagliardi, A. (2006). *J. Comput. Chem.* **27**, 995–1008.
- Chaudhuri, S., Chupas, P., Morgan, B. J., Madden, P. A. & Grey, C. P. (2006). *Phys. Chem. Chem. Phys.* **8**, 5045–5055.
- Coello Coello, C., Lamont, G. B. & Veldhuizen, D. A. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed. New York: Springer.
- Cohon, J. L. & Marks, D. H. (1975). *Water Resour. Res.* **11**, 208–220.
- Cookson, D., Kirby, N., Knott, R., Lee, M. & Schultz, D. (2006). *J. Synchrotron Rad.* **13**, 440–444.
- Das, I. & Dennis, J. E. (1997). *Struct. Optim.* **14**, 63–69.
- Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. (2002). *IEEE Trans. Evol. Comput.* **6**, 182–197.
- Ding, Y., Fan, F., Tian, Z. & Wang, Z. L. (2010). *J. Am. Chem. Soc.* **132**, 12480–12486.
- Dmowski, W., Egami, T., Swider-Lyons, K. E., Love, C. T. & Rolison, D. R. (2002). *J. Phys. Chem. B*, **106**, 12677–12683.
- Egami, T. & Billinge, S. J. L. (2003). *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*. Oxford: Pergamon.
- Farrow, C. L. & Billinge, S. J. L. (2009). *Acta Cryst.* **A65**, 232–239.
- Gagin, A. V. & Mullen, K. M. (2013). *nanop*. Version 2.0–2. <http://cran.r-project.org/web/packages/nanop>.
- Gilbert, B. (2008). *J. Appl. Cryst.* **41**, 554–562.
- Gilbert, B., Huang, F., Lin, Z., Goodell, C., Zhang, H. & Banfield, J. F. (2006). *Nano Lett.* **5**, 605–610.
- Glatter, O. (1979). *J. Appl. Cryst.* **12**, 166–175.
- Glatter, O. & Kratky, O. (1982). Editors. *Small Angle X-ray Scattering*. New York: Academic Press.
- Guinier, A. & Fournet, G. (1955). *Small-Angle Scattering of X-rays*. New York: Wiley.
- Hall, B. D. (2000). *J. Appl. Phys.* **87**, 1666–1675.
- Howell, R., Proffen, T. & Conradson, S. (2006). *Phys. Rev. B*, **73**, 94107.
- Hwang, S., Petkov, V., Rangan, K. K., Shastri, S. & Kanatzidis, M. G. (2002). *J. Phys. Chem. B*, **106**, 12453–12458.
- Ilavsky, J. (2012). *J. Appl. Cryst.* **45**, 324–328.
- Ilavsky, J. & Jemian, P. R. (2009). *J. Appl. Cryst.* **42**, 347–353.
- Knowles, J. & Corne, D. (1999). *Proceedings of the 1999 Congress on Evolutionary Computation*, Vol. 1, pp. 98–105. Piscataway: IEEE.
- Korsounski, V. I., Neder, R. B., Hradil, K., Barglik-Chory, C., Müller, G. & Neufeind, J. (2003). *J. Appl. Cryst.* **36**, 1389–1396.
- Korsunskiy, V. I. & Neder, R. B. (2005). *J. Appl. Cryst.* **38**, 1020–1027.
- Korsunskiy, V. I., Neder, R. B., Hofmann, A., Dembski, S., Graf, C. & Rühl, E. (2007). *J. Appl. Cryst.* **40**, 975–985.
- Kumpf, C. (2006). *Appl. Phys. A*, **85**, 337–343.
- Kumpf, C., Neder, R. B., Niederdraenk, F., Luczak, P., Stahl, A., Scheuermann, M., Joshi, S., Kulkarni, S. K., Barglik-Chory, C., Heske, C. & Umbach, E. (2005). *J. Chem. Phys.* **123**, 224707.
- Masadeh, A., Božin, E., Farrow, C., Paglia, G., Juhas, P., Billinge, S., Karkamkar, A. & Kanatzidis, M. (2007). *Phys. Rev. B*, **76**, 115413.
- Michel, F. M., Antao, S. M., Chupas, P. J., Lee, P. L., Parise, J. B. & Schoonen, M. A. A. (2005). *Chem. Mater.* **17**, 6246–6255.
- Mullen, K. & Levin, I. (2011). *J. Appl. Cryst.* **44**, 788–797.
- Neder, R. B. & Korsunskiy, V. I. (2005). *J. Phys. Condens. Matter*, **17**, S125–S134.
- Petkov, V. (2005). *Mol. Simul.* **31**, 101–105.
- Petkov, V., Gateshki, M., Choi, J., Gillan, E. G. & Ren, Y. (2005). *J. Mater. Chem.* **15**, 4654–4659.
- Price, K. V., Storn, R. M. & Lampinen, J. A. (2005). *Differential Evolution – A Practical Approach to Global Optimization*. Berlin, Heidelberg: Springer-Verlag.
- Proffen, T., Billinge, S. J. L., Egami, T. & Louca, D. (2003). *Z. Kristallogr.* **218**, 132–143.
- Rayleigh (1914). *Proc. R. Soc. London Ser. A*, **90**, 219–225.
- Trautmann, H., Steuer, D. & Mersmann, O. (2013). *mco*. R Package Version 1.0.12. <http://CRAN.R-project.org/package=mco>.
- Vogel, W., Borse, P. H., Deshmukh, N. & Kulkarni, S. K. (2000). *Langmuir*, **16**, 2032–2037.
- Vogel, W., Urban, J., Kundu, M. & Kulkarni, S. K. (1997). *Langmuir*, **13**, 827–832.
- Warren, B. E. (1969). *X-ray Diffraction*. Reading: Addison Wesley.
- Yang, X., Masadeh, A. S., McBride, J. R., Božin, E. S., Rosenthal, S. J. & Billinge, S. J. (2013). *Phys. Chem. Chem. Phys.* **15**, 8480–8486.
- Zadeh, L. A. (1963). *IEEE Trans. Autom. Control*, **AC-8**, 59–60.
- Zanchet, D., Hall, B. D. & Ugarte, D. (2000). *Characterization of Nanophase Materials*, editor Z. L. Wang, pp. 13–36. Weinheim: Wiley-VCH.
- Zitzler, E., Laumanns, M. & Thiele, L. (2001). *TIK-103*. Department of Electrical Engineering, Swiss Federal Institute of Technology, Zurich, Switzerland.