

# Direct and quantitative comparison of pixelated density profiles with high-resolution X-ray reflectivity data

P. Fenter,<sup>a\*</sup> S. S. Lee,<sup>a</sup> A. A. Skelton<sup>b</sup> and P. T. Cummings<sup>b,c</sup>

<sup>a</sup>Chemical Sciences and Engineering, Argonne National Laboratory, Argonne, IL, USA,

<sup>b</sup>Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN, USA, and <sup>c</sup>Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA. E-mail: fenter@anl.gov

A method for comparing pixelated density profiles (*e.g.* obtained from molecular dynamics or other computational techniques) with experimental X-ray reflectivity data both directly and quantitatively is described. The conditions under which such a comparison can be made quantitatively (*e.g.* with errors <1%) are determined theoretically by comparing calculated structure factors for an intrinsic continuous density profile with those obtained from density profiles that have been binned into regular spatial increments. The accuracy of the X-ray reflectivity calculations for binned density profiles is defined in terms of the inter-relationships between resolution of the X-ray reflectivity data (*i.e.* its range in momentum transfer), the chosen bin size and the width of the intrinsic density profile. These factors play a similar role in the application of any structure-factor calculations that involve the use of pixelated density profiles, such as those obtained from iterative phasing algorithms for inverting structures from X-ray reflectivity and coherent diffraction imaging data. Finally, it is shown how simulations of a quartz–water interface can be embedded into an exact description of the ‘bulk’ phases (including the substrate crystal and the fluid water, below and above the actual interface) to quantitatively reproduce the experimental reflectivity data of a solid–liquid interface.

© 2011 International Union of Crystallography  
Printed in Singapore – all rights reserved

**Keywords:** X-ray reflectivity; diffraction; molecular-dynamics simulations; pixelization; quartz–water interface.

## 1. Introduction

A key step in validating concepts concerning the controls on interfacial structures and processes is the comparison of experimental and computational results. A system can be considered to be well understood if these distinct approaches agree at a quantitative level; that is, when the differences between results are comparable with or smaller than the systematic or statistical uncertainties of the respective techniques.

When using diffraction or scattering-based techniques (*e.g.* X-rays, neutrons, electrons), this comparison can be complicated by the ‘phase problem’ (Als-Nielsen & McMorrow, 2001; Giacobozzo, 1992). Stated simply, scattering techniques measure the scattering intensity as a function of momentum transfer,  $\mathbf{Q} = \mathbf{k}_r - \mathbf{k}_i$  (where  $\mathbf{k}_i$  and  $\mathbf{k}_r$  are the wavevectors of the incident and reflected beams, respectively, with magnitude  $|\mathbf{k}| = 2\pi/\lambda$ , where  $\lambda$  is the X-ray wavelength). The X-ray reflectivity (XR) signal,  $R(\mathbf{Q})$  (*i.e.* the fraction of the incident

beam that is reflected by the sample), in the kinematic limit is proportional to the magnitude squared of the interfacial ‘structure factor’,  $F(\mathbf{Q})$ , which in turn is the Fourier transform (FT) of the interfacial density profile,  $\rho(\mathbf{r})$ , where  $\mathbf{r}$  is the position in space. For specular reflectivity [*i.e.*  $Q = 4\pi/\lambda \sin(2\theta/2)$ , where  $2\theta$  is the scattering angle], this is expressed as (Fenter, 2002)

$$R(Q) = (4\pi r_e / QA_{UC})^2 |F(Q)|^2, \quad (1)$$

where

$$F(Q) \equiv \int \rho(z) \exp(iQz) dz = |F(Q)| \exp[i\varphi(Q)].$$

Here,  $r_e$  is the classical electron radius,  $A_{UC}$  is the surface unit mesh area, and  $z$  is the height to the interfacial plane. If the phase of the structure factor,  $\varphi$ , were known, the full density profile could be obtained directly by the inverse FT. However, the phase information is lost in scattering measurements.

Consequently, it is necessary to by-pass the phase problem in order to obtain structural information from diffraction data.

The most common approach to analyzing scattering data is ‘model-based fitting’, in which a structural model is defined using a set of parameters (*e.g.* positions, occupancies and vibrational amplitudes for each atom) (Robinson & Tweet, 1992). These parameters often include both intrinsic factors (*i.e.* structure) as well as extrinsic factors [*e.g.* intensity scale factors, interfacial roughness (Robinson, 1986)], and they are optimized to match the data through non-linear least-squares fitting algorithms. Such an approach, while simple, has some significant limitations. Most notably, it is difficult to create an appropriately general model before knowing the salient features of the structure that is being determined. Consequently, this approach is characterized by substantial trial and error, and it is often not possible to prove if a given optimized model is the actual structure of interest (*i.e.* if it represents the global minimum in terms of the quality of fit or a local minimum). Even when a structural optimization is carried out successfully, the comparison with a computational study can be less than direct. For example, it can be difficult to assess whether the experimental and computational results are in true quantitative agreement, or whether residual disagreements reflect an inherent bias in the experimental data, the optimized fitting models or deficiencies in the computational approach. This traditional comparison between experimental and computational results has been performed recently for the quartz-(10 $\bar{1}$ 1)–water system for various choices of interatomic potentials (Skelton *et al.*, 2010).

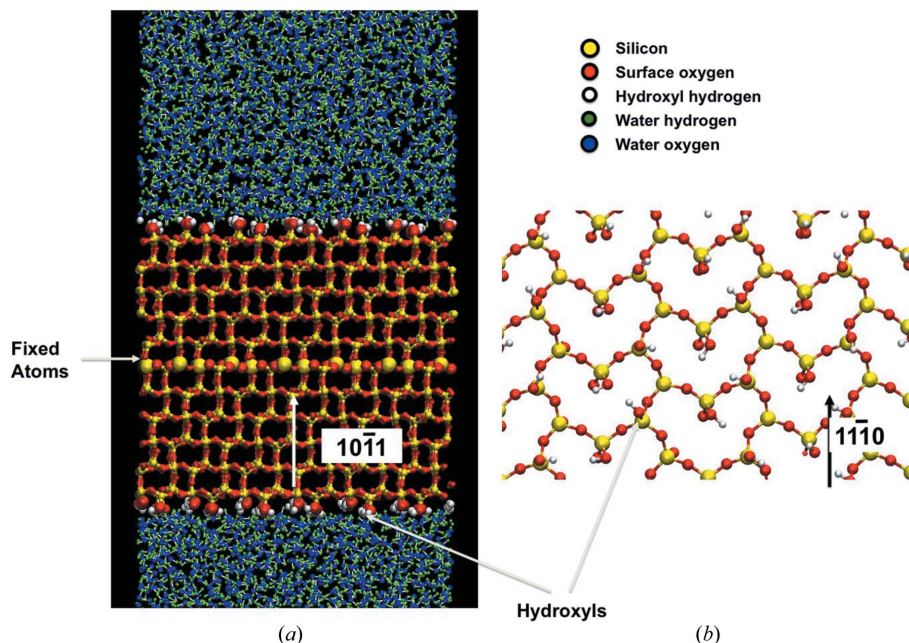
A recent approach is to use phase-sensitive direct methods that allow the phase problem to be by-passed (Miao *et al.*, 1998; Lyman *et al.*, 2005; Fenter & Zhang, 2005; Blasie *et al.*, 2003). In this approach, a ‘model-independent’ electron density profile is obtained from the data without any reference to parameterized structural models. The advantage of this approach is that no assumptions are made in the analysis process. Unfortunately, such approaches often do not provide quantitative agreement with the experimental data, because they do not readily incorporate interfacial structural relaxations (Fenter & Zhang, 2005) which are present at all interfaces, and they can be susceptible to systematic errors (especially when using only the one-dimensional information from specular reflectivity). Usually, such model-independent results are used to develop an appropriate model that can be optimized through least-squares results, and the model-dependent fitting results are considered to be more precise.

Here, we describe a new approach to directly and *quantitatively* evaluate the agreement between experimentally measured XR data and simulated structures (*i.e.* density profiles) obtained separately through computational techniques. This approach makes use of the trivial ability to calculate the scattering intensity for a time-averaged structure, as defined by a density profile,  $\rho(\mathbf{r})$ . The primary challenge is to generate a time-averaged density profile with sufficient spatial resolution that can reproduce accurately the XR data within available computational resources. The importance of spatial resolution derives from the use of discretized (or ‘binned’)

density profiles from the simulations. For example, increasing bin size reduces the number of simulation cycles needed to determine each bin’s density accurately. However, it also imposes a larger distortion of the density profile owing to pixelization, leading to systematic errors in the calculated XR signal. This becomes increasingly important when comparing non-specular reflectivity data with three-dimensional density profiles where the number of bins increases substantially. Specifically, we define the inter-related constraints between the bin size used in computational studies, the momentum transfer range of XR measurements and the intrinsic width of the structure of interest that allow for direct and accurate comparisons with known errors (*e.g.* <1%). To demonstrate the effectiveness of this approach, we also show how a simulated density profile of the quartz-(10 $\bar{1}$ 1)–water interface can be embedded into a bulk medium (*i.e.* calculated bulk structures of quartz and water) so that it can be directly compared with previously measured XR data with negligible artifacts. These results also define the conditions under which phase-retrieval algorithms and other model-independent fitting approaches (Miao *et al.*, 1999; Williams *et al.*, 2003; Blasie *et al.*, 2003; Lyman *et al.*, 2005; Fenter & Zhang, 2005) can be performed accurately when they utilize the FT of discretized structures (*i.e.* as obtained from a fast-Fourier transform).

## 2. Methods

We begin with an interfacial structure obtained from molecular dynamics (MD) simulation of the quartz-(10 $\bar{1}$ 1)–water interface (Skelton *et al.*, 2010). A snapshot from an MD simulation obtained using the Lopes *et al.* (2006) force field is shown in Fig. 1(*a*). The distribution of surface hydroxyls in the top surface above the silicon–oxygen plane is shown in the top view of the surface (Fig. 1*b*). This force field contains partial charges to model the electrostatic interactions and a Lennard–Jones potential to model short-range repulsion owing to electron cloud overlap and attractions owing to dispersion interactions. The atoms within the quartz surface are fully mobile, except the central layer of atoms which was held rigid throughout the simulation (as noted in Fig. 1*a*) to determine the intrinsic root-mean-square (r.m.s.) widths of interfacial species. Bond stretching, angle bending and dihedral potentials are used within the surface to maintain its structure. The Lopes *et al.* force field uses the TIP3P water model (Mark & Nilsson, 2001). The classical MD simulations were performed using the public-domain parallel-efficient classical MD code, *LAMMPS* (Plimpton, 1995). A 1 fs time-step was used with the velocity Verlet integration and Nosé/Hoover thermostat, in the *NVT* (constant number of molecules *N*, constant volume *V* and constant temperature *T*) ensemble. The simulation included 5040 atoms in the quartz crystal and 2320 water molecules. The dimensions of the orthorhombic box were 41.45298 × 39.231 × 80.2286 Å. The simulations were run for 2 ns and the particle–particle particle–mesh (PPPM) method was used to treat the long-ranged electrostatics (Hockney & Eastwood, 1988).


**Figure 1**

Side (a) and top (b) views of the quartz–water MD simulation cell. The fixed atoms at the center of the surface slab are indicated. Note that this simulation is periodically replicated in all three coordinate directions.

Specular XR measurements are sensitive to the one-dimensional electron density profile,  $\rho(z)$  (in units  $e^- \text{Å}^{-3}$ ), along the surface normal direction. This is derived from the time-averaged and laterally averaged number density profiles for each element,  $n_j(z)$ , derived from the simulation (Fig. 1)

$$\rho(z) = \sum_j Z_j n_j(z), \quad (2)$$

where  $Z_j$  is the atomic number of the  $j$ th element. The electron density profile of the quartz-(10 $\bar{1}1$ )–water interface (Fig. 2a) was derived from the MD simulation. This was obtained by making a time-averaged histogram of atom locations using a bin size of 0.02 Å. These results illustrate features that are typical of simulated density profiles including the spatial extent of the simulation cell ( $\sim 80$  Å) with 11 quartz unit-cell layers.

The XR signal for the quartz-(10 $\bar{1}1$ )–water interface (red line, Fig. 2b) is calculated from the MD-simulated structure. This is obtained using an extension of equation (1) that includes the necessary detail that the  $Q$ -dependent X-ray scattering for each element differs owing to the atomic form factor,  $f(Q)$ , of each atom. Here, the total structure factor is given by

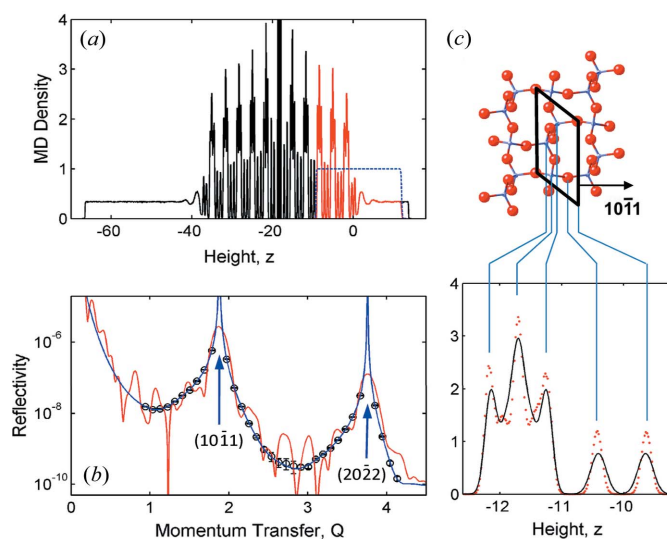
$$F(Q) = \sum_j f_j(Q) \text{FT}[n_j(z)], \quad (3)$$

where the sum is over the  $j$  elements in the structure, and the FT of the discretely binned density profile,  $n(z)$ , for each element,  $j$ , is obtained from the MD simulation.

This calculated XR signal is compared with experimental data (black circles, Fig. 2b) [see Schlegel *et al.* (2002) for details of the XR measurement and model-dependent analysis]. Also included in this comparison is a model-

dependent fit to the XR data (blue line, Fig. 2b) derived from Schlegel *et al.* (2002). The XR signal in Fig. 2(b) exhibits Bragg peaks [at  $Q = 1.88 \text{Å}^{-1}$  and  $3.76 \text{Å}^{-1}$  corresponding to the first- and second-order reflections associated with the quartz-(10 $\bar{1}1$ ) layer spacing,  $d_{10\bar{1}1} = 3.341 \text{Å}$ ]. The bulk electron density profile derived from the MD simulation reproduces the main features of the bulk crystallographic structure (e.g. the density peak locations), but the detailed shape of the two profiles differs significantly (Fig. 2c).

There are two significant artifacts that appear in the calculation of the XR from this simulation. First, the substrate Bragg peaks in XR data are weaker and broader than that observed in the model-dependent calculation. The Bragg peaks have a width in  $Q$  corresponding to  $2\pi/L$ , where  $L$  is the spatial extent of the material probed by the beam (in this case, controlled by the X-ray penetration depth). In the XR


**Figure 2**

(a) Electron density profile (in units of  $e^- \text{Å}^{-3}$ ) plotted as a function of height (in units of Å) obtained from a MD simulation of the quartz-(10 $\bar{1}1$ )–water interface (black line). The profile is shown from a side view as a function of height above the interface of interest near  $z = 0$  Å. The blue dashed line indicates the slicing function discussed in the text, and the portion of the electron density shown in red is that extracted from the total profile using the slicing function. (b) The X-ray reflectivity signal,  $R$ , as a function of momentum transfer,  $Q$  (in  $\text{Å}^{-1}$ ), calculated from the full MD-simulated profile in (a) (red line), a best-fit model-dependent fit to the XR data (blue line), and the measured XR signal for the quartz-(10 $\bar{1}1$ )–water interface (black circles with experimental error bars). (c) Side view of the bulk quartz lattice with a surface unit cell indicated in the black parallelogram. Also shown is a detailed comparison of the bulk unit cell structure obtained from crystallographic parameters (black line) and an MD-simulated quartz structure whose density profile has been discretized with 0.02 Å bins (red points).

study of the quartz–water interface structure, this corresponds to a thickness of many micrometers, leading to sharp Bragg peaks with an XR magnitude that approaches  $\sim 1$ . The MD-derived reflectivity calculation also shows additional artifacts in the form of intensity oscillations corresponding to the sizes of the simulated density region, with periods of  $\sim 2\pi/35 \text{ \AA}^{-1} \simeq 0.18 \text{ \AA}^{-1}$  corresponding to the width of the simulated quartz crystal, and  $2\pi/80 \text{ \AA} \simeq 0.08 \text{ \AA}^{-1}$  corresponding to size of the simulation box (this is mostly apparent at small  $Q$ ). It is therefore not possible to make a quantitative (or even a qualitative) assessment of the level of agreement between the experimental data and the MD simulation based on this comparison. These artifacts remain even if one incorporates a semi-infinite substrate density profile in the XR calculation (Sakuma & Kawamura, 2009). In this case, intensity oscillations would be present owing to the artificial truncation of the water profile near  $z = 12 \text{ \AA}$ . A more subtle issue is whether there are artifacts in the calculated XR data due to the binning process. Finally, we need to address whether the detailed differences between the bulk electron density profiles seen in Fig. 2(c) lead to significant errors.

The quality of agreement between the XR data and the calculated reflectivity (from both the model-dependent fit and the MD-simulated profile) is quantified in two ways. The  $\chi^2$  function measures the quality of fit with respect to the experimental uncertainties in the data,  $\chi^2 = \sum_i [(R_i - R_{ci})/\sigma_i]^2 / (n - n_p)$ , where  $R_i$ ,  $\sigma_i$  and  $R_{ci}$  are the measured reflectivity, its uncertainty and the calculated reflectivity, the sum is over all data points ( $i = 1 \dots n$ ), and  $n_p$  is the number of parameters used in the optimization.  $\chi^2 \simeq 1$  for a fit that is consistent with the data to within the experimental uncertainties. A second function is the ‘R-factor’, where  $r = \sum_i |R_i - R_{ci}| / R_{ci} / (n - n_p)$ , and is the average fractional deviation between the data and the calculation, independent of experimental uncertainties.

### 2.1. Structure factor of a binned density profile

We develop a formalism to define the relationships between the bin size in the binned density profiles and the calculated XR intensities to define the conditions under which a comparison between these quantities can be made accurately and quantitatively. In making such a comparison it is necessary first to consider the sensitivity of the scattering intensities to the binning of the density into individual pixels of size  $\Delta$ . Scattering techniques probe a structure through a coherent summation of the electron density within an associated real-space ‘resolution’,  $\delta z_{\text{res}} \simeq \pi/Q_{\text{max}}$ , characterized by the maximum momentum transfer,  $Q_{\text{max}}$ , along a particular direction (Fenter & Sturchio, 2004). This concept of X-ray scattering resolution is fully analogous with that in a microscope or other optical instruments (Halliday & Resnick, 1978). It therefore can be anticipated that low-resolution data (e.g.  $Q_{\text{max}} < 1.5 \text{ \AA}^{-1}$  or  $\delta z_{\text{res}} > 2 \text{ \AA}$ ) would be relatively insensitive to perturbations from discretizing a density profile into small (e.g.  $\Delta < 0.1 \text{ \AA}$ ) bins. However, specular XR data are often collected to  $Q_{\text{max}} = 6 \text{ \AA}^{-1}$  to  $8 \text{ \AA}^{-1}$ , corresponding to real-space resolutions of  $\leq 0.5 \text{ \AA}$ . At what point does the calculated

reflectivity profile become sensitive to the discretization of the density profile?

We first define the binned density,  $\rho_b$ , from the intrinsic continuous density profile,  $\rho(z)$ . Here, the density functions are assumed to be one-dimensional and have an integrated weight,  $NZ$ , where  $Z$  is the atomic number of each atom and  $N = \sum n(z)$  is the total number of atoms in the profile. (For simplicity in this derivation, we initially ignore differences in the atomic form factors for each atom.) From this we obtain

$$\rho_b(z_b) = \int_{-\infty}^{\infty} \rho(z)W(z, z_b) dz = \rho(z) \otimes W(z), \quad (4)$$

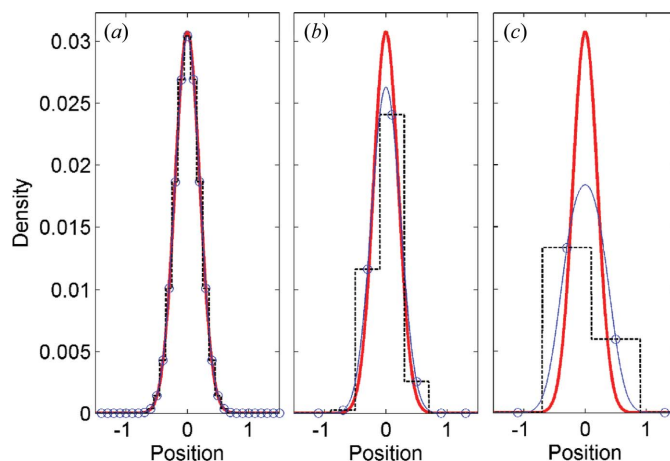
where  $\otimes$  denotes the convolution of two functions, and  $W$  is a window function defined as

$$W(z, z_b) = \begin{cases} 1 & \text{for } |z - z_b| < \Delta/2, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

where  $W$  is a continuous function of  $z$ , for discrete values of  $z_b$ . In the case of the MD-simulated profiles,  $\rho_b$  is a discrete profile sampled at the bin size  $\Delta$  (i.e.  $z_b = 0, \Delta, 2\Delta \dots$ ), as shown in Fig. 3 (circles) compared with the intrinsic density profile (red lines, Fig. 3). The integrated weight of the window function is  $\int_{-\infty}^{\infty} W dz = \Delta$ . Notably, the sum of the integrated weight of the profile is unaffected by the transformation [i.e.  $\sum \rho_b = \int_{-\infty}^{\infty} \rho(z) dz$ ]. However, in order to compare the absolute electron densities for the binned and intrinsic profiles, we scale the binned profile by a factor of  $\delta z/\delta z_b$ , corresponding to the change of coordinates from  $z$  to  $z_b$ . From this the structure factor of the binned profile,  $F_b$ , can be calculated,

$$F_b(Q) \equiv \sum_{z_b} \rho_b(z_b) \exp(iQz_b). \quad (6)$$

The binned density profile,  $\rho_b$ , in turn is the convolution of the continuous density profile with a window function [equation



**Figure 3** Intrinsic (red line) and discretely binned (blue circles) density profiles for a Si atom having an intrinsic width of  $\sigma = 0.2 \text{ \AA}$  for bin sizes of (a)  $0.1 \text{ \AA}$ , (b)  $0.4 \text{ \AA}$  and (c)  $0.8 \text{ \AA}$ , shown as a function of position (in  $\text{\AA}$ ). Also plotted are the discretely binned density profile shown as a step-wise continuous profile (black dashed lines) and the continuously binned density profiles,  $\rho_{b\text{-cont}}$  (blue line).

(4)]. When calculating the binned structure factor of multi-element systems, one needs to apply (6) for each element and then obtain the total structure factor that takes into account the atomic form factor of each element using (3).

Equation (6) has a notable short-coming when the binned density profile,  $\rho_b$ , is sampled at the bin size,  $\Delta$  (as is typical for creating a binned density profile). In this case the calculated structure factor depends on the details of the binning process, both in terms of the bin size,  $\Delta$ , and in terms of the offset between the bin location and the intrinsic profile. While the former factor can be corrected for (see below), the latter factor cannot. This can lead to significant numerical errors in the calculated reflected signal. For example, when the bin size is comparable with the intrinsic width of a feature (for example, a Gaussian profile with r.m.s. width  $\sigma$ ), the structure factor of the discretely binned density profile (*e.g.* a single bin) shows intensity minima when the momentum transfer,  $Q$ , ‘resolves’ the bin size (*i.e.* at  $Q \simeq \pi/\Delta$ ). Such minima are not present in the structure factor of the intrinsic Gaussian profile, and cannot be exactly corrected for without prior knowledge of the intrinsic profile (which is not available). As will be shown below, these factors do, however, depend in a systematic way on the relative sizes of  $\Delta$ ,  $\sigma$  and the momentum transfer,  $Q$ , for a desired accuracy (*e.g.* 1%).

A more accurate and precise result is obtained by considering a ‘continuously binned density profile’,  $\rho_{b\_cont}$ , with a point spacing,  $\delta z_b$ , obtained using (4) when  $z_b$  is considered to be an almost continuous variable (*i.e.* when  $\delta z_b \ll \Delta$ ; blue lines, Fig. 3). In particular, when a continuously binned profile is used, the structure factor of this quantity can be obtained exactly through the FT of the continuously binned profile of  $\rho_{b\_cont}(z_b)$ ,

$$F_{b\_cont}(Q) = \int_{-\infty}^{\infty} \rho_{b\_cont}(z_b) \exp(iQz_b) dz_b = \text{FT}[\rho(z) \otimes W(z)]. \quad (7a)$$

Using the convolution theorem (Giacovazzo, 1992), we obtain

$$F_{b\_cont}(Q) = \text{FT}[\rho(z)] \text{FT}[W(z)] = F(Q) [\sin(Q\Delta/2)/(Q\Delta/2)]. \quad (7b)$$

Here,  $F(Q)$  is the intrinsic structure factor of the original continuous density profile. Consequently we can define the ‘bin-corrected’ structure factor as

$$F_{b\_corr}(Q) = F_b(Q) / [\sin(Q\Delta/2)/(Q\Delta/2)]. \quad (8)$$

This quantity provides an estimate of the intrinsic structure factor,  $F(Q)$ , directly from the simulated density profiles. Its accuracy is evaluated below. While we do not explicitly convolute the simulated density profile with the window function, we note that the discretely binned density profile,  $\rho_b$  (circles, Fig. 3), simply samples the continuously binned density profile,  $\rho_{b\_cont}$  (blue line, Fig. 3). Consequently we anticipate that this formalism can be used for the discretely binned profiles and will become more accurate as the bin size is reduced.

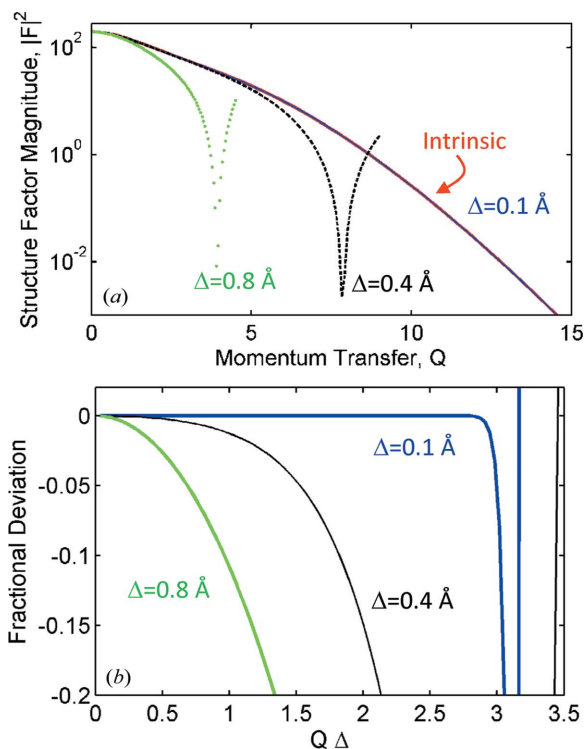
## 2.2. Comparison of binned versus intrinsic structure-factor magnitudes

The utility of this formalism is assessed by making a closed-form comparison of X-ray scattering intensities calculated for an intrinsic structure and the structure after it has been binned. The deviations between the two approaches reflect the limitations of calculating the scattering intensities from binned profiles. Here, we choose an atom with a Gaussian distribution located at the origin as the intrinsic density distribution. This choice is made for two reasons. First, the ability to accurately calculate the scattering intensity from an atom implies the same ability for any arbitrary structure. Second, the structure factor is easily calculated,  $F(Q) = f(Q) \exp[-0.5(Q\sigma)^2]$ , where  $\sigma$  is the r.m.s. width of the atom distribution. We choose a single Si atom located at  $z = 0 \text{ \AA}$ , with  $\sigma = 0.2 \text{ \AA}$  (with a full width at half-maximum,  $\text{FWHM} = 2.35\sigma = 0.47 \text{ \AA}$ ). The comparison of the intrinsic and binned structures is made for bin sizes  $\Delta = 0.1, 0.4$  and  $0.8 \text{ \AA}$ , and as a function of  $Q$ .

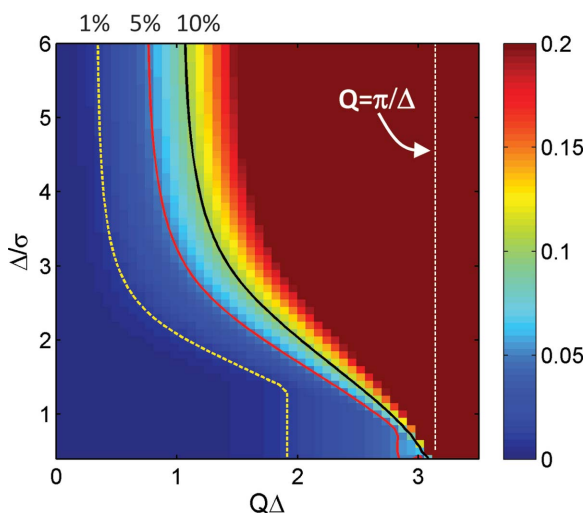
The intrinsic and binned electron density profiles are shown in Fig. 3. The choice of  $\Delta = 0.1 \text{ \AA}$  leads to a binned density profile that closely resembles the intrinsic profile, with only small perturbations. The use of larger bin sizes,  $\Delta = 0.4$  and  $0.8 \text{ \AA}$ , leads to substantial modifications in the density profile.

The effect of binning on the calculated XR signal is shown in Fig. 4(a). These calculations compare the modulus square of the intrinsic structure factor [*i.e.* related to the scattering intensity by equations (1) and (3)] with that from the bin-corrected structure factor [equation (8)]. In each case the modulus square of the bin-corrected structure factor appears to reproduce the exact calculation at small  $Q$ , but with substantial deviations that appear at  $Q\Delta = \pi$ , typically seen as a sharp minimum in the intensity, corresponding to the point where the individual bins are resolved. A finer-scale assessment of this formalism can be obtained by plotting the fractional errors in the calculated intensities [*i.e.*  $(|F_{b\_corr}|^2 - |F|^2)/|F|^2$ ] between the magnitudes of the bin-corrected structure factor,  $F_{b\_corr}$ , and the intrinsic structure factor,  $F$  (Fig. 4b). Here, we plot the data as a function of  $Q\Delta$ , which puts the node from the width of each pixel size at the same value. These plots show that each calculation is essentially exact at sufficiently small values of  $Q\Delta$ , but that the range of  $Q$  over which the calculation has a specified accuracy (*e.g.* 10%) depends on the bin size. For  $\Delta = 0.1 \text{ \AA}$ , the results are accurate to within 1% error for  $Q\Delta < 2.8$  (*i.e.*  $Q < 28 \text{ \AA}^{-1}$ ). Meanwhile, the calculation for  $\Delta = 0.8 \text{ \AA}$  shows that 1% accuracy is achieved only for  $Q\Delta < 0.3$  ( $Q < 0.38 \text{ \AA}^{-1}$ ). That is, the accuracy is controlled by two factors: the resolution of the data ( $Q < \pi/\Delta$ ) as well as the relative bin size,  $\Delta$ , with respect to the intrinsic r.m.s. width,  $\sigma$ .

The systematic variation of the calculation error is shown, in Fig. 5, as a function of the dimensionless variables  $Q\Delta$  and  $\Delta/\sigma$ . Since the calculated reflectivity varies with the exact details of the pixel locations, the results shown in Fig. 5 represent the worst case evaluated by comparing the results for different pixel locations (calculated for bins shifted by all increments of  $\Delta/8$ ). These results confirm the general obser-



**Figure 4**  
 (a) The variation of the structure-factor magnitude,  $|F|^2$ , as a function of momentum transfer,  $Q$  (in units of  $\text{\AA}^{-1}$ ) of the distributions shown in Fig. 3, including the intrinsic structure factor (red line), and the bin-corrected structure factors,  $|F_{b,corr}|^2$ , for each of the bin sizes,  $\Delta = 0.1, 0.4$  and  $0.8 \text{ \AA}$  (blue, black and green lines, as indicated). Note how the results for  $\Delta = 0.1 \text{ \AA}$  provide an almost exact reproduction of the intrinsic structure factor. (b) Fractional errors between the intrinsic and bin-corrected intensities, plotted as a function of the product of the momentum transfer and bin size,  $Q\Delta$ .



**Figure 5**  
 Systematic variation of errors between the bin-corrected structure-factor magnitude,  $|F_{b,corr}|^2$ , and the exact structure-factor magnitude (indicated as a color map), as a function of the dimensionless quantities  $Q\Delta$  and  $\Delta/\sigma$ . The yellow-dashed, red and black lines indicate the contours associated with 1%, 5% and 10% errors. The condition associated with ‘resolving’ a structure,  $Q = \pi/\Delta$ , is indicated by the dashed white line. For the sake of clarity, errors are shown with a maximum value that saturates at a value of 20%.

variations shown above and reveal the systematic trends in the accuracy of the bin-corrected structure-factor calculations. There are two distinct regimes. The accuracy is ultimately controlled by the resolution of the data,  $Q\Delta < \pi$ , when the pixel size is comparable with, or smaller than, the intrinsic width (i.e.  $\Delta < 1.5\sigma$ ). This is the expected result. When the pixel size becomes larger than the intrinsic width, however, the  $Q$ -range over which the calculations are accurate becomes substantially reduced, with a continuous transition between these two extreme limits. For calculations with 1% accuracy, the  $Q$ -range is modified by the sensitivity to factors such as the specific offset between the bins and the center of the intrinsic distribution. The range of parameters consistent with 1% accuracy can be summarized as  $Q\Delta < 1.8$  when  $\Delta/\sigma < 1.5$ , and  $Q\Delta < 0.5$  for  $\Delta/\sigma > 3$ .

This calculation provides a quantitative and practical approach to determine the optimal binning size when an MD simulation result is compared with the experimental XR. For example, for the quartz structure, the most limiting component to the structure factor calculation is due to the silicon atom, with an r.m.s. vibrational amplitude of  $0.09 \text{ \AA}$ . A bin size of  $0.02 \text{ \AA}$  leads to  $\Delta/\sigma \simeq 0.2$ . With  $Q_{max} \simeq 5 \text{ \AA}^{-1}$ , we obtain  $Q_{max}\Delta = 0.1$ . These parameters are therefore well within the region in which the reflectivity is accurate to  $< 1\%$ . Under these conditions an accurate calculation of the XR signal can be obtained.

### 2.3. Comparison of MD simulations with experimental XR data

We now describe the process of incorporating this calculation of an MD-simulated structure into a structure-factor calculation of a solid–liquid interface. The main challenge is to avoid artifacts associated with discontinuities in the density profile at the boundary between the simulated and known structures. The total structure factor of an interface,  $F_{tot}$ , is the sum of the structure factors of the individual components [equation (3)]. In order to calculate reflectivity data that can be compared with experimental values, it is necessary to ‘embed’ the simulated structure into the calculated density profiles of the two bulk phases whose structures are known (the unrelaxed bulk substrate crystal, and the featureless fluid phase far above the interface). When including the structure factor of the portion of the density profile simulated by MD,  $F_{MD}$ , the total structure factor can be written as

$$F_{tot} = F_{UC} F_{CTR} + F_{int} + F_{MD} + F_{water}, \quad (9)$$

where  $F_{UC}$ ,  $F_{CTR}$ ,  $F_{int}$  and  $F_{water}$  are the structure factors of the bulk unit cell, the ‘crystal truncation rod’, the interfacial region (including any relaxed layers, adsorbed species, modulation of the bulk water, etc. that are not included in  $F_{MD}$ ) and bulk fluid, respectively. For the case of specular reflectivity,  $F_{CTR} = 1/[1 - \exp(iQc)]$ , where  $c$  is the substrate layer spacing along the surface normal direction. Since the water structure factor is meant only to reproduce the bulk water properties, it can be written as an error function profile defined by a position,  $z_w$ , an r.m.s. width,  $u_w$ , and the bulk

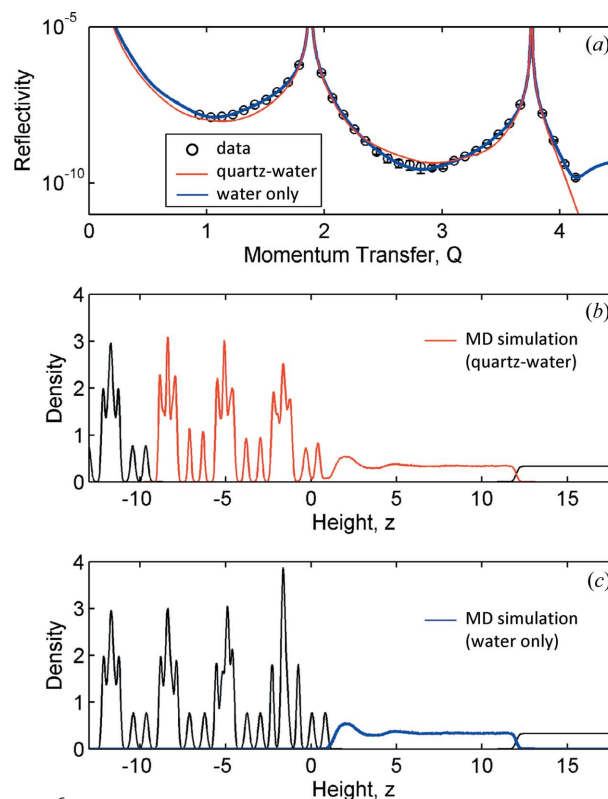
electron density of water,  $\rho_w = \sum Z_i/V_W = 0.33 \text{ e}^- \text{ \AA}^{-3}$ , where  $V_W = 30.3 \text{ \AA}^3$  is the volume of one water molecule in bulk water. In this case,  $F_{\text{water}} = (A_{\text{UC}}/V_W)f_{\text{water}}(Q)\exp(iQz_w)\exp[-0.5(Qu_w)^2]/iQ$ , where  $f_{\text{water}}$  describes the atomic form factor from a single water molecule.

Embedding the simulated profile into the bulk-like structures of the substrate and fluid profiles must be done without any artificial discontinuities at the boundaries of any two structure components. This is a two-step process. First, we select a portion of a MD profile that is of interest, by multiplying the simulated profile by a slicing function:  $[\text{erf}(z, z_1, \sigma_1) + 1][1 - \text{erf}(z, z_2, \sigma_2)]/4$ , where erf is the error function profile,  $z_i$  are the top and bottom limits of the simulation that will be incorporated, and  $\sigma_i$  are the widths of the interfacial slicing function, as shown in Fig. 2(a). Within the crystal, the inflection point of the error function profile can be set to a zero-density region between crystalline layers so that there is no discontinuity between the ‘bulk-like’ material and the density obtained from the MD simulation. On the fluid side of the crystal, the slicing position is chosen as any location where the MD-simulated fluid density profile matches the bulk water density. The same position and width ( $z_2$  and  $\sigma_2$ ) are used to define the termination of the MD-simulated profile and the start of the bulk water profile so that there is no discontinuity in the density profile at this boundary.

**2.3.1. Comparison with the MD-simulated quartz-(10 $\bar{1}$ 1)–water interface.** Using the process described above, we can calculate the reflectivity corresponding to the quartz-(10 $\bar{1}$ 1)–water interface using the MD-simulated density profile for the top three quartz layers as well as the first  $\sim 10 \text{ \AA}$  of the interfacial water profile (red line, Fig. 6b). This is embedded within the closed-form calculation for the respective bulk structures, including the ideal quartz crystal substrate density profile and the bulk water profile (black line, Fig. 6b).

The embedding process has a few parameters that need to be optimized. First, the reflectivity calculation is critically sensitive to surface structural relaxations. In the case of specular reflectivity, this is controlled by the vertical displacements of the MD-simulated density profile with respect to the crystallographic atom positions in the substrate crystal. This can be defined, initially, with reasonable precision by allowing the binned density profile to overlap with that of the ideally terminated layer, and adjusting their relative positions so that the two structures overlap. It is necessary, however, to allow this relative spacing to be optimized as a fitted parameter.

A few non-structural parameters influence the XR signal and need to be optimized. These include an intensity scale parameter (to compensate for errors in the estimated absolute reflectivity signal), as well as parameters that describe the surface roughness and the attenuation of the X-ray beam through the sample cell. The structure factor of an ideally rough surface (*i.e.* one in which the surface height changes with position, but exposes the same local structure at every point) can be written  $F = F_{\text{rough}}(Q)F_{\text{ideal}}(Q)$ . There are multiple functional forms that have been developed for different types of surface roughness. A widely used form



**Figure 6**

(a) X-ray reflectivity calculations (*i.e.* reflectivity versus  $Q$ , in units of  $\text{\AA}^{-1}$ ) for the quartz-(10 $\bar{1}$ 1)–water interface structure, compared with the data obtained experimentally by Schlegel *et al.* (2002). The electron density profile (with density and height plotted in units of  $\text{e}^- \text{ \AA}^{-3}$  and  $\text{\AA}$ , respectively) in (b) corresponds to the structure used to calculate the red line in (a), while the profile in (c) corresponds to the blue line in (a). In (b) and (c) the red and blue lines, respectively, indicate the portion of the profile obtained from the MD simulation. The black line in (b) indicates a bulk-like structure for both quartz and water. In (c), the black line indicates a bulk-like quartz layer along with three relaxed interfacial layers (optimized by least-squares fitting) and the bulk water structure. The heights in (b) and (c) are plotted in units of  $\text{\AA}$ .

assumes that the surface has a series of partial layers having occupation factors  $\beta, \beta^2, \beta^3 \dots$  (Robinson, 1986). In this case,  $F_{\text{rough}} = (1 - \beta)^2/[1 + \beta^2 - 2\beta\cos(Qc)]$ , where the r.m.s. interfacial roughness has a value of  $c\beta^{1/2}/(1 - \beta)$ . A second parameter is needed to describe the attenuation of the X-ray beam through the water and Kapton film above the sample surface. The lack of data for  $Q < 1 \text{ \AA}^{-1}$  make these data insensitive to this parameter (*i.e.* its uncertainty is very large when it is allowed to vary). It was set to a typical value of  $20 \mu\text{m}$ .

Using this approach, we can see that the qualitative features of the reflectivity data, *e.g.* the expected crystal truncation rod shape, are reproduced by the combined MD-simulated interfacial profile and the closed-form calculation of the bulk contributions (red line, Fig. 6a). At this level, there do not appear to be any obvious artifacts in the simulation. For instance, if there were errors in matching the MD-simulated and closed-form calculations, one would expect to observe oscillations in the reflectivity profile with a period of about  $2\pi/20 \text{ \AA} = 0.3 \text{ \AA}^{-1}$ , associated with the size of the MD-simulated region,  $\sim 20 \text{ \AA}$  (*e.g.* as seen in Fig. 2b).

An evaluation of the simulated data shows that the embedded MD-simulated profile does not, however, provide *quantitative* agreement with the experimental data. The reflectivity calculation in Fig. 6(a) has a quality of fit of  $\chi^2 = 28$  and  $r = 0.4$  (with  $n_p = 3$ ); that is, the average deviation is substantially larger than the uncertainty for each data point and is substantially worse than the model-dependent fit [the blue curve in Fig. 2(b) is the result of a re-analysis of the XR data having a quality of fit of  $\chi^2 = 0.63$  and  $r = 0.034$ , which is improved with respect to that reported previously (Schlegel *et al.*, 2002)]. The poorer quality of fit for the MD-simulated profile is not due to the binning process since the data are characterized by  $Q_{\max} = 4 \text{ \AA}^{-1}$  and a bin size  $\Delta = 0.02 \text{ \AA}$ . Comparing with the results in Fig. 5, these results correspond to conditions for which we expect to be able to obtain a quantitative calculation with <1% errors for  $Q_{\max}\Delta \simeq 0.1$  and  $\Delta/\sigma_{\text{Si}} \simeq 0.2$ , which are well within the region for accurate calculation of the reflectivity signal. Therefore, we tentatively conclude that the poorer quality of fit is because of some aspect of the MD-calculated profile.

**2.3.2. Comparison with the MD-simulated interfacial water profile.** It is apparent that the simple reflectivity calculation based on the interfacial profile from the MD simulation, combined with the extrinsic parameters, does not provide a quantitative reproduction of the measured data. It is reasonable to expect that the lack of full quantitative agreement may be due to deficiencies in the simulated structure of the quartz substrate, especially given the detailed differences in electron density between the simulated density profile and that obtained from the crystallographic structure for the bulk portion of the structure, seen in Fig. 2(c). This idea can be tested by using only the simulated water and by treating all other aspects of the interfacial structure (*i.e.* the substrate structural relaxations) through a model-dependent fitting approach. This comparison is a hybrid of the model-dependent fitting approach along with the MD-simulated water profile (*i.e.* without the substrate portion; Fig. 6c). This was performed by fitting the interfacial quartz structural parameters to describe the correlated relaxation profiles of the top three layers of quartz, as described previously (Schlegel *et al.*, 2002). We also optimized the relative location of the MD-simulated and calculated profiles, as described above. The total number of fitted parameters was then 14. This approach leads to a significantly better fit to the reflectivity data (Fig. 6a), with quality of fits of  $\chi^2 = 0.93$  and  $r = 0.042$ , indicating a quantitative reproduction of the experimental data. This is only marginally worse in quality to that obtained by model-dependent fitting (*i.e.*  $\chi^2 = 0.63$  and  $r = 0.034$  when all aspects of the interfacial structure were optimized), suggesting that any deviation in the calculation of the reflectivity owing to the binning of the MD-simulated profile are small compared with the experimental uncertainties. A comparison of the quartz portion (*i.e.* for  $z < 0$ ) of two interfacial profiles shows that the interfacial structure obtained by fitting to the XR data (Fig. 6c) differs significantly from that obtained from the full MD simulation (Fig. 6b). Given the high sensitivity of XR to interfacial structure, we infer that the inability of the quartz-

water MD simulation to quantitatively reproduce the XR data is associated with differences in the MD-simulated quartz structure, which ultimately would be due to the potentials that describe the quartz interfacial region in the MD simulation. This observation is also reinforced by the small but significant differences between the bulk quartz structure derived from MD simulation with respect to the known crystal structure, as seen in Fig. 2(c).

We also have evaluated the role of bin size by comparing the data with the same simulation while using a fivefold-larger bin size,  $\Delta = 0.1 \text{ \AA}$ , again using only the MD-simulated water profile. All other details in the comparison were equivalent, including the optimization of the interfacial roughness and the intensity scale factor. This comparison leads to qualities of fit,  $\chi^2 = 0.95$  and  $r = 0.042$ , which are almost identical to those obtained with  $0.02 \text{ \AA}$ -sized bins. This is expected since the values of  $Q_{\max}\Delta = 0.1$  and  $\Delta/\sigma = 1$  for this comparison are well within the parameter region defined by intrinsic systematic errors of <1% (Fig. 5).

### 3. Discussion and conclusions

The present results demonstrate a procedure by which a pixelated structure (*e.g.* from a MD simulation, or other computational approach) can be compared directly with XR data to test the consistency between the computational and experimental results. This process is attractive for a number of reasons. This provides a distinct route to understanding the experimental XR data because it makes no detailed assumptions about the nature of the interfacial structure (*i.e.* as would be implicit with the creation of any parameterized model to fit the XR data), except those that are implicit to the computational studies (*e.g.* parameterized interaction potentials). Such an approach could be used, for instance, to identify likely interfacial structures so that an appropriate choice of structural model can be made for fitting. On the other hand, this approach opens up a new window through which the quality of interaction potentials used in the MD simulations can be evaluated. Normally, interaction potential parameters are obtained by fitting to experimental bulk properties (densities *etc.*) or quantum mechanical calculations. This ability to directly and quantitatively compare the MD profiles with measured XR data opens up the possibility of evaluating the relative accuracy of various potentials to reproduce the measured interfacial structures of a given system [*e.g.* the Lopes force field *versus* ClayFF (Cygan *et al.*, 2004)]. More generally, it might be possible to optimize the interaction potentials based on measurements of interfacial structures.

One result of this comparison was the demonstration that the systematic errors in comparing the measured data with pixelated density profiles is controlled by two factors: the inter-relationships between the pixel size,  $\Delta$ , and the data resolution ( $\pi/Q_{\max}$ ), and the intrinsic width of the structural features,  $\sigma$ . These relationships are defined by the dimensionless quantities  $Q_{\max}\Delta$  and  $\Delta/\sigma$ . One implication of this observation concerns model-independent data analysis approaches that involve the FT of pixelated densities derived



from data [e.g. for XR data (Blasie *et al.*, 2003; Lyman *et al.*, 2005; Fenter & Zhang, 2005) or more generally approaches such as coherent diffraction imaging (Miao *et al.*, 1999; Williams *et al.*, 2003)]. These approaches will be subject to systematic errors whenever a component of the intrinsic structure is sharp compared with the resolution of the data.

We obtained quantitative agreement between the XR data and MD simulation only when the structure of the MD-simulated quartz surface was excluded from the calculation. This suggests that the Si and O potentials of the Lopes force field did not reproduce the actual quartz interfacial structure that is observed experimentally. This highlights a benefit of this approach in that it can help to isolate the features in the simulation that do not agree with the data. (A complete analysis would also consider whether there exists any interfacial water profiles that, when combined with the MD-simulated quartz structure, would reproduce the experimental XR data.)

The present results made a comparison using only specular reflectivity data which are sensitive to the vertical structure at an interface (*i.e.* along the surface normal direction). Inclusion of non-specular reflectivity data (*i.e.* ‘crystal truncation rod’ data) provides information about the lateral interfacial structures. One challenge in making a similar direct comparison between MD simulated structures and the non-specular data will be the need to define the density in each voxel of linear dimension,  $\Delta$ , in a simulation cell of volume  $V \simeq 80 \text{ \AA} \times A_{\text{UC}} = 2600 \text{ \AA}^3$ . For  $\Delta = 0.02 \text{ \AA}$ , this would require that the density in each of  $\sim 10^8$  voxels be independently and accurately defined (compared with the current one-dimensional result requiring only  $\sim 10^3$  pixels), therefore requiring substantially greater computational resources (both in the amount of time that would be needed to simulate to obtain adequate statistics and the computational power required for the binning itself). However, the generalized assessment of the systematic errors as a function of the dimensionless parameters  $Q_{\text{max}}\Delta$  and  $\Delta/\sigma$  (Fig. 5) provides a way of making an intelligent choice of voxel dimensions. This would be defined by the details of the structure of interest,  $\sigma$ ; the quality and extent of the experimental data (*e.g.* defined by the resolution of the data,  $Q_{\text{max}}$ , which may differ in the vertical and lateral directions); and ultimately the experimental uncertainties which provide an estimate for the level of accuracy that is needed.

This work was supported by the Geosciences Research Program, Office of Basic Energy Sciences, US Department of

Energy. PF and SSL were supported under Contract DE-AC02-06CH11357 to UChicago Argonne, LLC, as operator of Argonne National Laboratory (‘Mineral-Fluid Interactions’). AAS was supported through the project ‘Nanoscale Complexity at the Oxide–Water Interface’ (ERKCC41) at Oak Ridge National Laboratory. A portion of this research was conducted at the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Division of Scientific User Facilities, US Department of Energy. The X-ray reflectivity data were collected at beamline 12-ID-D, Advanced Photon Source. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357 to UChicago Argonne, LLC, as operator of Argonne National Laboratory.

## References

- Als-Nielsen, J. & McMorrow, D. (2001). *Elements of Modern X-ray Physics*. Chichester: John Wiley and Sons.
- Blasie, J. K., Zheng, S. & Strzalka, J. (2003). *Phys. Rev. B*, **67**, 224201.
- Cygan, R. T., Liang, J. J. & Kalinichev, A. G. (2004). *J. Phys. Chem. B*, **108**, 1255–1266.
- Fenter, P. A. (2002). *Rev. Mineral. Geochem.* **49**, 149–220.
- Fenter, P. & Sturchio, N. C. (2004). *Prog. Surf. Sci.* **77**, 171–258.
- Fenter, P. & Zhang, Z. (2005). *Phys. Rev. B*, **72**, 081401(R).
- Giacovazzo, C. (1992). *Fundamentals of Crystallography*. Oxford University Press.
- Halliday, D. & Resnick, R. (1978). *Physics*, Part 2. New York: John Wiley and Sons.
- Hockney, R. W. & Eastwood, J. W. (1988). *Computer Simulation Using Particles*, Special Student ed. Philadelphia: Hilger.
- Lopes, P. E. M., Murashov, V., Tazi, M., Demchuk, E. & MacKerell, A. D. (2006). *J. Phys. Chem. B*, **110**, 2782–2792.
- Lyman, P. F., Shneerson, V. L., Fung, R., Harder, R. J., Lu, E. D., Parihar, S. S. & Saldin, D. K. (2005). *Phys. Rev. B*, **71**, 081402(R).
- Mark, P. & Nilsson, L. (2001). *J. Phys. Chem. A*, **105**, 9954–9960.
- Miao, J., Sayre, D. & Chapman, H. N. (1998). *J. Opt. Soc. Am. A*, **15**, 1662–1669.
- Miao, J. W., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
- Plimpton, S. (1995). *J. Comput. Phys.* **117**, 1–19.
- Robinson, I. K. (1986). *Phys. Rev. B*, **33**, 3830–3836.
- Robinson, I. K. & Tweet, D. J. (1992). *Rep. Prog. Phys.* **55**, 599–651.
- Sakuma, H. & Kawamura, K. (2009). *Geochim. Cosmochim. Acta*, **73**, 4100–4110.
- Schlegel, M. L., Nagy, K. L., Fenter, P. & Sturchio, N. C. (2002). *Geochim. Cosmochim. Acta*, **66**, 3037–3054.
- Skelton, A. A., Kubicki, J. D., Fenter, P., Wesolowski, D. J. & Cummings, P. T. (2010). *J. Phys. Chem. C*. Submitted.
- Williams, G. J., Pfeifer, M. A., Vartanyants, I. A. & Robinson, I. K. (2003). *Phys. Rev. Lett.* **90**, 175501.