



JOURNAL OF  
APPLIED  
CRYSTALLOGRAPHY

**Volume 48 (2015)**

**Supporting information for article:**

**Determining the Shape and Periodicity of Nanostructures using  
Small Angle X-ray Scattering**

**Daniel F. Sunday, Scott List, Jasmeet S. Chawla and R. Joseph Kline**

# Determining the Shape and Periodicity of Nanostructures using Small Angle X-ray Scattering

Daniel F. Sunday,<sup>1</sup> Scott List,<sup>2</sup> Jasmeet S. Chawla,<sup>2</sup> R. Joseph Kline<sup>1\*</sup>

1. National Institute of Standards and Technology, Materials Science and Engineering Division,  
Gaithersburg, MD 20899
2. Intel Corporation, Components Research Division, Hillsboro OR 97124 USA

\* [rjkline@nist.gov](mailto:rjkline@nist.gov)

Supplemental Information

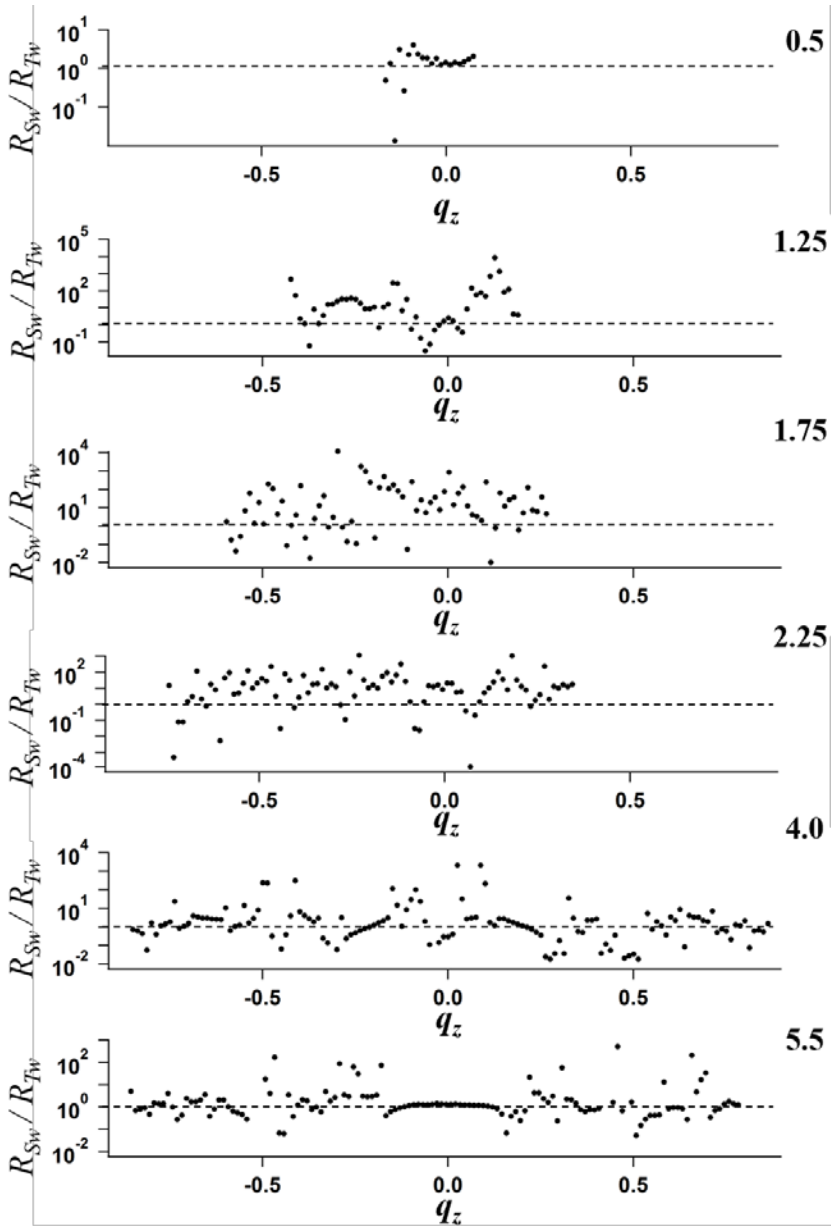


Figure S1: Plots of  $R_{Sw}/R_{Tw}$  ( $R_{Sw}$  is residuals for single width model,  $R_{Tw}$  is the residual for the two width model) for selected peak orders of PQ1. The peak order is indicated by the number in the top right. Values larger than 1 indicate an improvement in the goodness of fit for the two width model compared to the single width. Dashed line denotes a  $R_{Sw}/R_{Tw}$  of 1.

### Description of the Shape Model

The simulated scattering was calculated from a model where the base unit was a trapezoid, which has the advantage that the Fourier transform can be calculated analytically, avoiding the need for a more computationally intensive numerical approach. Figure S2 defines how the coordinates are assigned. The trapezoid is asymmetric to allow the stack to reproduce the shark fin

shape of the lines. The Fourier transform is calculated over the area of the trapezoid (with coordinates and slopes defined in Figure S2) which produces Eq S2. Figure S3 shows in detail how the model parameters are assigned.  $A(q)$  is calculated for each individual trapezoid and summed to provide the scattering amplitude for the entire structure.

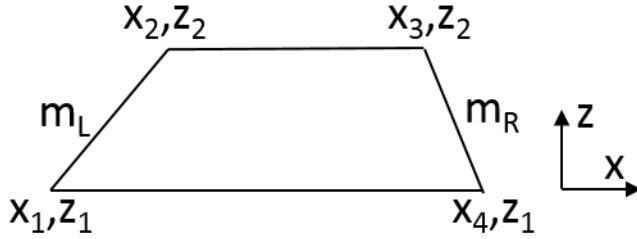


Figure S2: Definition of coordinate system for derivation of analytical Fourier transform of an asymmetric trapezoid,  $m_L$  is the slope of the left edge and  $m_R$  is the slope of the right edge.

$$S1 \quad A(q) = \int_{A(r)} e^{-iqr} dr$$

$$S2 \quad A(q) = \frac{e^{-iq_x \left(\frac{-z_1+x_1}{m_l}+x_1\right)}}{q_x \left(\frac{q_x}{m_l} + q_z\right)} \left( e^{-i\left(\frac{q_x}{m_l}+q_z\right)z_2} - e^{-i\left(\frac{q_x}{m_l}+q_z\right)z_1} \right) - \frac{e^{-iq_x \left(\frac{-z_1+x_4}{m_r}+x_4\right)}}{q_x \left(\frac{q_x}{m_r} + q_z\right)} \left( e^{-i\left(\frac{q_x}{m_r}+q_z\right)z_2} - e^{-i\left(\frac{q_x}{m_r}+q_z\right)z_1} \right)$$

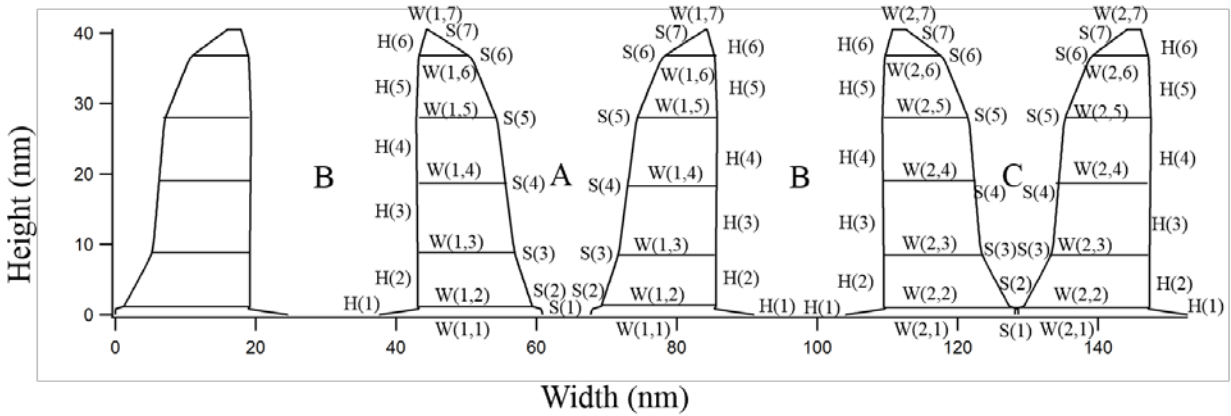


Figure S3: Model image with each free parameter labeled. For the Single Width model  $W(1,:) = W(2,:)$ .  $W(1,2)$  is the width of the base of the 2<sup>nd</sup> trapezoid from the bottom in the first pair of

lines. H corresponds to the trapezoid height and S corresponds to the shift in the edge of the trapezoid relative to the bottom left point on the lowest trapezoid in the stack. The ratio of W and S determines if the trapezoid is symmetric or skewed.

### Description of the Monte Carlo Markov Chain (MCMC) model

The MCMC algorithm searches the local space of the goodness of fit ( $GF$ ) and generates a model population. The variation of the model parameters in the population is reflective of the sensitivity of the parameters to the experimental data. The probability of a given model being in the population is proportional to its  $GF$  ( $GF$  is calculated according to Eq S4, where  $I_{Sim}$  represents the simulated intensity and  $I_{Exp}$  represents the experimental intensity). The algorithm is initialized by using the best known fit as the initial condition. Candidate models are generated by making small, random changes in each of the model parameters simultaneously. If the  $GF_{i+1} < GF_i$  then the candidate model is accepted, otherwise an acceptance probability ( $\alpha$ ) is generated based on the relative fit quality of the candidate model compared to the best known model. Eq S3 is used to calculate  $\alpha$ , where  $GF_B$  is the goodness of fit of the best known model. (Mosegaard & Sambridge, 2002) A random number  $y \in [0,1]$  is then generated and if  $y < \alpha$  the candidate model is accepted. The step size of the chains were tuned so the probability of accepting a new step was between 0.3 and 0.4, which has been shown to result in the fastest convergence of the chains. (Mosegaard & Tarantola, 1995) Each parameter (H,W,S) was allowed to vary by at least an order of magnitude in each direction from the initial position, ensuring that none of the parameters were trapped near a limit or in a local minima. Each simulation consisted of 10 chains with 3 million steps. The model population developed from the MCMC is then resampled every 100 steps to remove correlations between points, for a total population of about 300,000. Figure S4 shows a correlation plot for the key parameters, the line spacings, height and widths. Using this plot it can be shown that most of the parameters are completely uncorrelated. A diagonal distribution of points is indicative of correlations between two parameters. Mild correlations between the A-B and B-C spacings are observed, and also between Width1 and Width2.

$$S3 \quad \alpha = e^{-0.5(GF_i - GF_B)}$$

$$S4 \quad GF = \frac{|I_{Sim} - I_{Exp}|^2}{I_{Exp}^2}$$

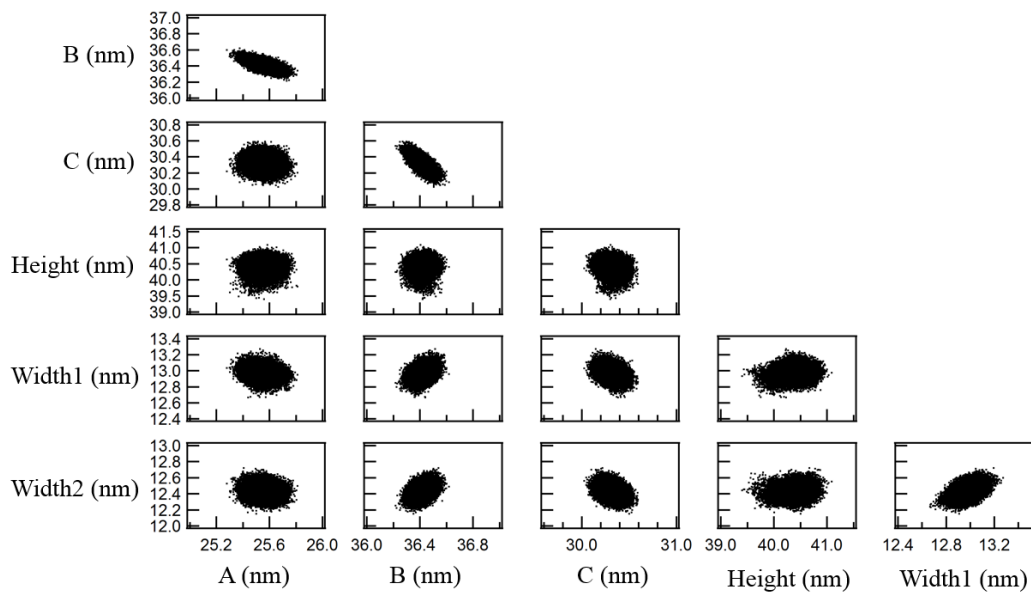


Figure S4: Correlation plots for A, B and C spacings, Height, Width1 and Width2. Diagonal distributions indicate correlations between parameters.

### References

Mosegaard, K. & Sambridge, M. (2002). *Inverse Probl.* **18**, R29.

Mosegaard, K. & Tarantola, A. (1995). *J. Geophys. Res.* **100**, 12431–12.