

# Characterizing the Nanoscale Layers of Tomorrow's Electronics

## An Application of Fourier Analysis

Christopher Bishop Payne  
Princeton University

United States Department of Energy  
Office of Science, Science Undergraduate Laboratory Internship Program  
SLAC National Accelerator Laboratory  
Menlo Park, California

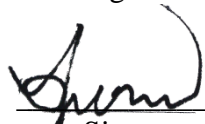
August 20<sup>th</sup>, 2011

Prepared in partial fulfillment of the requirement of the Office of Science, Department of Energy's Science Undergraduate Laboratory Internship under the direction of Apurva Mehta of the SSRL division of SLAC National Accelerator Laboratory.

Participant:

  
\_\_\_\_\_  
Signature

Research Advisors:

  
\_\_\_\_\_  
Signature

## **Abstract**

Characterizing the Nanoscale Layers of Tomorrow's Electronics: An Application of Fourier Analysis. CHRISTOPHER PAYNE (Princeton University, Princeton, NJ 08544) APURVA MEHTA (Stanford Synchrotron Radiation Lightsource (SSRL) Division of SLAC National Accelerator Laboratory, Menlo Park, CA 94025) MATTHEW BIBEE (Stanford University, Stanford, California, 94305)

Thin film applications are of great interest to the semiconductor industry due to the important role they play in cutting edge technology such as thin film solar cells. X-Ray Reflectivity (XRR) characterizes thin films in a non-destructive and efficient manner yet complications exist in extracting these characteristics from raw XRR data. This study developed and tested two different algorithms to extract quantity of layers and thickness information on the nanometer scale from XRR data. It was concluded that an algorithm involving a local averaging technique revealed this information clearly in Fourier space.

## I. Introduction

The next generation of electronic devices will power an increasingly technology dependent world, requiring industry to be able to manufacture semiconductor chips that are both smaller and composed of new materials. The semiconductor industry has rapidly grown over the past three decades as they have pushed the limits of fabrication technology into the nanometer resolution range. This exponential decrease in device size was hypothesized by Gordon Moore in 1965<sup>i</sup> and has led to increasingly cheaper devices with more processing power per area than their more expensive predecessors. Additionally, nearly all semiconductors are now made of silicon, however the use of new semiconductor materials could yield electronics that are more efficient and can operate under greater extremes.

To unlock these favorable attributes of these next generation semiconductors, we need to be able to characterize and analyze materials with a resolution of a few nanometers. Many different techniques have been developed to perform this analysis, yet the one we will focus on is called X-Ray Reflectivity (XRR). This technique quantifies parameters of materials by revealing the number of layers of which the material is composed of and each layer's corresponding thickness, density, and roughness value.<sup>ii</sup> These parameters are discussed in more detail in the theory section, but for now, it is sufficient to understand that XRR allows the user to know the physical structure of a material on the nanometer scale. It should also be noted that unlike other techniques, XRR does not destroy the sample nor does it require specially prepared samples such as Transmission Electron Microscopy<sup>iii</sup> demands.

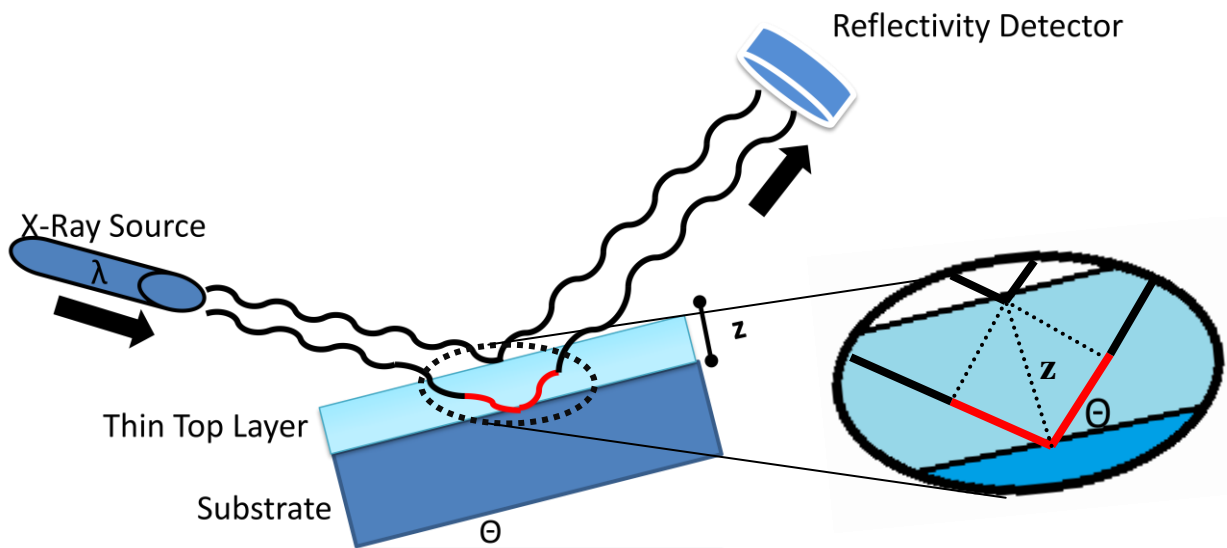
The specific focus of this project was to characterize the layers present on silicon carbide (SiC) wafers that GE is in the process of developing for future use in a wide range of applications that include wind turbines and hybrid-electric vehicles.<sup>iv</sup> It is important to keep in

mind though that the XRR analysis techniques discussed in this paper can be applied to a vast range of cutting edge semiconductor applications that involve nanometer scale layers such as thin film solar cells and power transistors. Having the capability to characterize these nanometer layers in the non-destructive manner XRR allows, is a crucial step on the path to developing the electronics of the future that the world urgently seeks.

## II. Materials and Methods

### *i. Reflectivity Theory*

When a multilayer system is illuminated by an x-ray beam of wavelength ( $\lambda$ ), the reflected beam profile measured as a function of the incidence angle, theta ( $\Theta$ ), contains information about the layer stratigraphy, density and roughness. Figure 1 shows a simple case of one thin layer on an infinitely thick substrate. The reflected beam in this case is a coherent summation of the beam reflected from the top surface and the beam reflected from the buried interface. The intensity of the reflected beam is then a function of the path length difference between the beams reflected from these two interfaces. This path length difference is shown in red in Figure 1.



**Figure 1** XRR Experimental Setup

Note the path length difference highlighted in red. This path length difference is a function of  $\Theta$  and  $Z$  as shown in Equation 1.

The path length difference that modifies the phase of the reflected wave can be approximated as

$$\textit{Extra Distance} = 2z\sin(\theta) \quad (1)$$

This extra distance corresponds to a phase difference at the point that the reflected beams interfere at the detector, dependent on the  $\lambda$  of the incident beam:

$$\textit{Phase Difference At Detector} = \frac{2z\sin(\theta)}{\lambda} \quad (2)$$

The beams reflected from the two interfaces interfere constructively if this phase difference is a multiple of  $2\pi$  and destructively if the phase difference is an odd multiple of  $\pi$ . This simple illustration, therefore, suggests that the intensity of the reflected beam is a function of the incidence angle ( $\Theta$ ) and is modulated to contain information about the layer thickness.

Under a more rigorous derivation, but still under the kinematic approximation, the reflected beam intensity is:

$$\textit{Intensity}(S) \approx \frac{1}{S^4} \left| FT\left(\frac{d\rho(z)}{dz}\right) \right|^2 \quad (3)^v$$

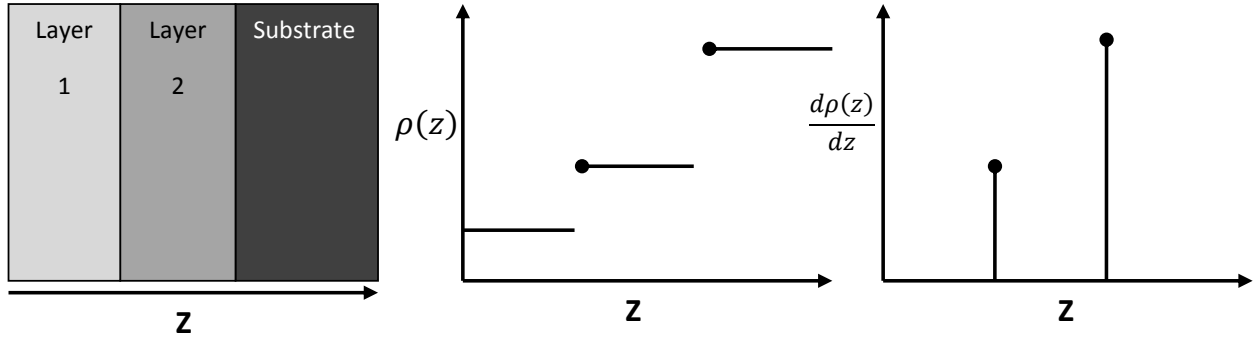
with  $\rho(z)$  being the electronic density a distance  $z$  (see Figure 1) into the sample and  $S$  defined as:

$$S = \frac{2\sin(\theta)}{\lambda} \quad (4)^{vi}$$

Equation 3 can be rewritten in the following manner:

$$\textit{Intensity}(S) \approx \frac{1}{S^4} FT\left(\frac{d\rho(z)}{dz} * \frac{d\rho(z)}{dz}\right) \quad (5)$$

As the reflected beam intensity is a function of the electron density gradient into the sample, it contains information about not only the thickness of the individual layers, but also the interfacial roughness and density difference between adjacent layers<sup>vii</sup>. Figure 2 represents an ideal situation in which  $\rho(z)$  of a particular layer is uniform and  $\rho(z)$  instantaneously changes at a layer interface.



**Figure 2** Relating density gradients to the sample's layers.

As Figure 2 depicts, the non-zero values of  $d\rho(z)/dz$  correspond to layer boundaries expressed in terms of  $z$ , thus we must arrange Equation 5 to get  $d\rho(z)/dz$  in terms of intensity. The  $1/S^4$  term in Equation 5 is a fall-off term which attenuates the intensity value over even a small  $S$  domain and is valid if the interfaces have no roughness. Experimentally we found this fall-off term ranges from  $1/S^4$  to approximately  $1/S^6$  for real multilayer samples, thus we developed an algorithm to calculate this fall-off term on a case by case basis so that it could be removed from the data. The algorithms that were tested are referred to as dN and 2N in the Results and Discussion section. With this fall-off term removed by using this algorithm, Equation 5 now has the form:

$$Intensity(S) - \text{Falloff Term} \approx FT\left(\frac{d\rho(z)}{dz} * \frac{d\rho(z)}{dz}\right) \quad (6)$$

The final step is to take the inverse Fourier Transform which will yield the self-convolution of  $d\rho(z)/dz$ :

$$FT^{-1}(Intensity(S) - \text{Falloff Term}) \approx \frac{d\rho(z)}{dz} * \frac{d\rho(z)}{dz} \quad (7)$$

Shortly, using some simulations, we will explore how the layer characteristics are encoded in Equation 7.

The main challenge in converting the original equation, Equation 3, into Equation 7, is to accurately calculate the fall-off term so that it can be removed. The following sections describe two algorithms for calculating the fall-off term. Subsequently, we will test the range of validity

of Equation 7, and explore how the various characteristics of the layers and layer stratigraphy are encoded in the formulation suggested by Equation 7.

*ii. Fall-Off and Conversion Algorithm Development Outline*

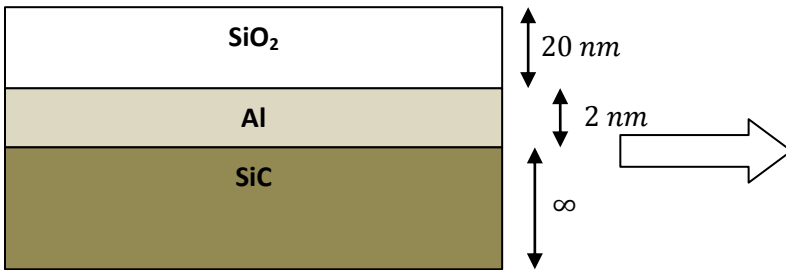
In order to find a solution to the problem outlined previously, I used MATLAB to create two algorithms that applied two different methods to remove the fall-off term from the intensity data. Independent of which method was used, the result was then transformed into Fourier space where sample characteristics such as number of layers present and layer thicknesses were compared with the information available in Fourier space. As will be discussed later, additional characteristics such as roughness or density of the layers are thought to be contained within the Fourier space, but were not fully explored in my study.

To optimize the fall-off subtraction algorithm and to test the range of validity of Equation 7, I used a MATLAB program called Multig, developed by Anneli Munkholm and Sean M. Brennan. I inputted sample characteristics such as the number of layers present and their corresponding thicknesses into the program and it outputted a simulated intensity versus  $\Theta$  curve. I then applied my algorithm to the simulated data to test whether the algorithm could return the sample characteristics I had simulated. I used this simple process to develop and tune my algorithm until the algorithm yielded the correct characteristics.

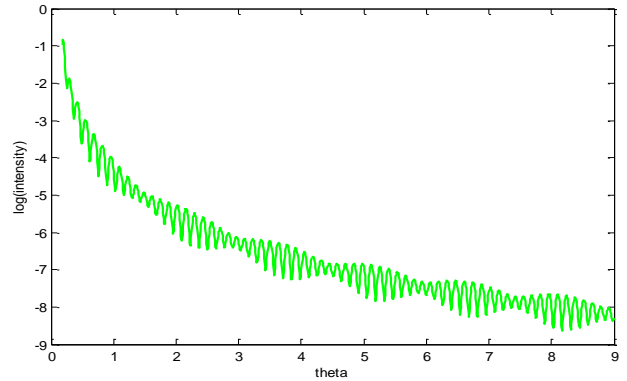
Below, the entire process of simulating raw intensity versus  $\Theta$  data and then transforming it into the accessible form of Equation 7 is outlined in five steps. Note that the two different fall-off removal methods,  $dN$  and  $2N$ , are introduced in Step 3, Fall-off Removal.

### Step 1, Simulate intensity versus $\Theta$ data using Multig

In this example, a 2 nm layer of aluminum (Al) between a 20 nm layer of silicon oxide (SiO<sub>2</sub>) is simulated on a SiC substrate (Figure 4). The Multig program then outputs a simulated intensity versus  $\Theta$  profile (Figure 3).



**Figure 4** Simulated sample inputted into Multig, no roughness in either layer.



**Figure 3** Resulting log(Intensity) versus  $\Theta$  simulated curve at 12keV

### Step 2, Restrict the $\Theta$ Domain

X-rays do not penetrate the sample until  $\Theta$  exceeds a critical angle. The kinematic approximation is known to be inaccurate near the critical angle, therefore, data points below a certain  $\Theta$  threshold are removed from the intensity versus  $\Theta$  data set to be operated on. Additionally, the reflectivity curve for higher  $\Theta$  values becomes progressively noisy, accordingly, data points above a certain threshold are removed as well. Techniques for selecting favorable threshold values, referred to as *theta\_low\_clip* and *theta\_high\_clip*, respectively, are discussed later. Additionally, the  $\log_{10}$  of the intensity data is taken to make it easier to operate on the data over approximately eight orders of magnitude.

### Step 3, Fall-off Removal

*Apply the dN Method*

The main idea behind this method is to take an averaged derivative of the intensity versus  $\Theta$  signal as described by Poust and Goorsky in *Enhanced Fourier Transforms for X-Ray*



*Scattering Applications*, which will minimize the original fall-off component present in the resulting dataset<sup>viii</sup>. The discrete version of this, as we are dealing with a finite dataset is described as:

$$I_j^{dN}(\theta) = \frac{1}{N} \sum_{i=1}^N \frac{I(\theta_{j+i}) - I(\theta_{j-i})}{\theta_{j+i} - \theta_{j-i}} \quad (8) \text{ ix}$$

Expressed in words, the algorithm takes the  $j^{\text{th}}$  intensity &  $\Theta$  data point and replaces it with the ‘local’ average derivative of the intensity. The ‘local’ area is defined by the dataset ranging from  $\theta_{j-N}$  through  $\theta_{j+N}$  and is centered on the  $j^{\text{th}}$  data point. By applying this process, the fall-off term is removed leaving a dataset in the form of Equation 6. As recommended by Poust and Goorsky,  $N$  is selected so that it is high enough to average out noise fluctuations but kept substantially less than the period of any possible thickness signals. As a result of taking the derivative, the dataset now trails the original dataset by a quarter of a period of the high frequency oscillations present in the original dataset.

*Or, alternatively apply the 2N Method*

The central idea behind this method is to construct a smooth and monotonic curve that removes the fall-off term from the original dataset while keeping all of its original frequency content. This is done by taking the local average of the intensity versus  $\Theta$  data and recording it as the fall-off value at that point. As alluded to in *Enhanced Fourier Transforms for X-Ray*

*Scattering Application*<sup>x</sup>:

$$I_j^{2N\_Falloff}(\theta) = \frac{1}{2N+1} \sum_{i=j-N}^{j+N} I(\theta_i) \quad (9)$$

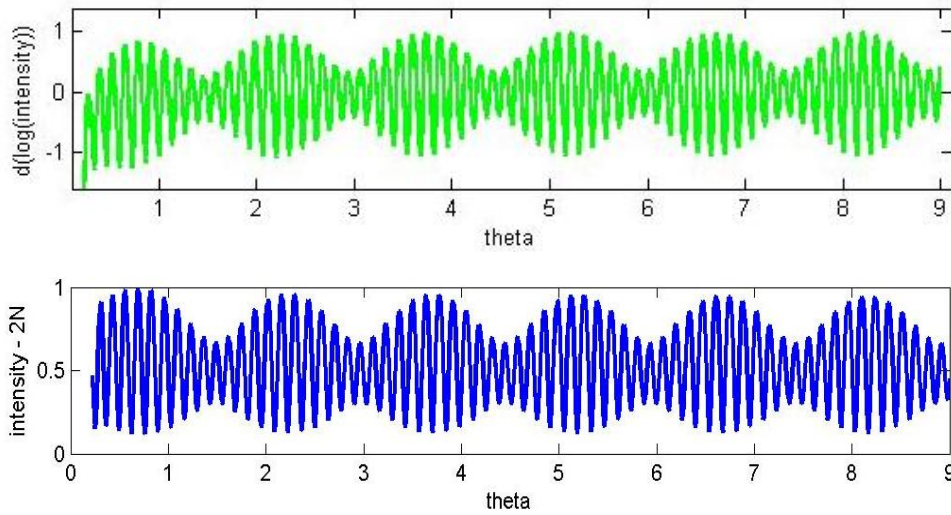
The method is called ‘2N’ as the local average includes the  $j^{\text{th}}$  point plus the  $N$  data points between  $\theta_{j-N}$  and  $\theta_j$  plus the  $N$  points between  $\theta_j$  and  $\theta_{j+N}$ . Effective selection of  $N$  will be discussed in detail later.

After quantifying the fall-off value at each point, we now subtract it from the corresponding original intensity values to remove the fall-off.

$$I_j^{2N}(\theta) = I_j^{Original}(\theta) - I_j^{2N\_Falloff}(\theta) \quad (10)$$

The last step in the 2N method is to take the anti-log of all the intensity values that have had the fall-off removed from them,  $I_j^{2N}(\theta)$ .

Regardless of whether the dN or 2N method is used, the fall-off has now been subtracted from the data, resulting in a dataset in the form of Equation 6, shown graphically in Figure 5.



**Figure 5** Graphical view of dataset after fall-off has been removed using dN (Top green graph) and 2N(Bottom blue graph)

Note, the dN result includes negative values, as it is the derivative of the log of the intensity. On the other hand, the 2N result is entirely positive because it is the antilog of the log of intensity minus the fall-off term:

$$2N \text{ Result (Blue Graph)} = 10^{I^{2N}} \text{ with } I^{2N} \text{ defined in Equation 10}$$

#### Step 4, Convert $\Theta$ into S

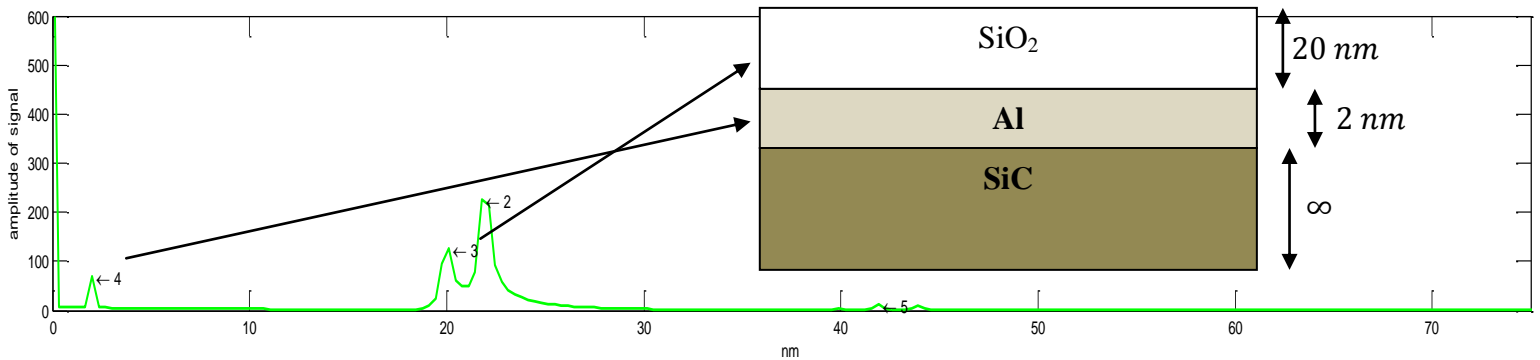
The  $\Theta$  values are converted to S according to Equation 11:

$$S = \frac{4\pi \sin(\theta - \theta_{critical})}{\lambda} \quad (11)$$

Note that this has an extra coefficient of  $2\pi$  compared to Equation 4 and that  $\theta_{critical}$  is subtracted off because there is no interference for  $\Theta$  below  $\theta_{critical}$  as there is no x-ray penetration of the sample. This equation favorably converts the  $\Theta$  axis [degrees] into  $S [m^{-1}]$ . For the more advanced reader, this transforms  $\Theta$  into a value related to momentum space.

### Step 5, Fourier Space Analysis

Lastly, we take the inverse Fourier transform of the extracted oscillation shown in Figure 5 in order to quantitatively know the number of layers and their corresponding thicknesses. In MATLAB, this is done using the built in `fft()` command and then discarding the results that correspond to negative frequencies – a mathematical byproduct of any application of `fft()`. The final result, expressed graphically, is the approximation of the convolution of the density gradient of the sample with itself as outlined in Equation 7. Figure 6 is the result from the simulation we started with:



**Figure 6** Fourier transform result

Note peak 4 corresponding to the 2nm layer and peak 3 corresponding to the 20 nm layer. Peak 2 is simply the sum of these two layers, equaling about 22nm. Also, we see the usefulness of transforming theta[degrees] into  $S[m^{-1}]$ , which results in an axis expressed in units of meters after the Fourier transform.

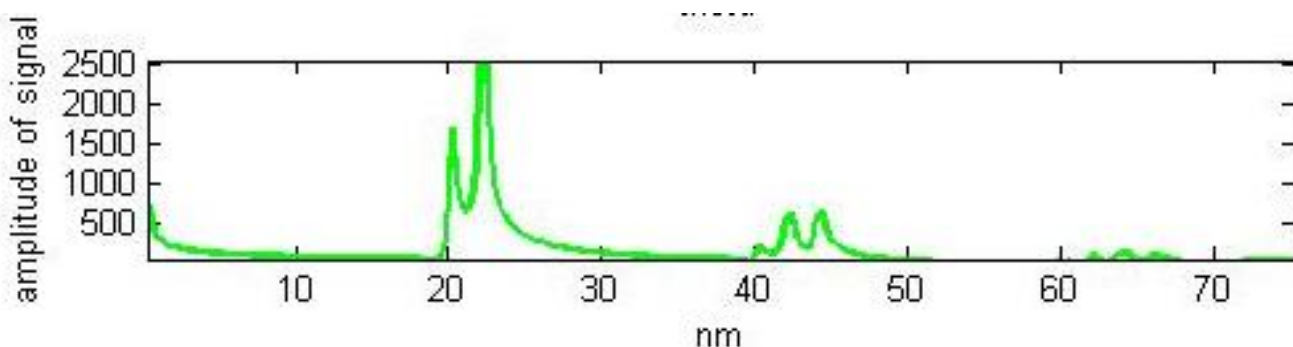
### III. Results

I will now discuss the results from testing the above algorithm against Multig simulations that included no roughness and roughness. This testing was designed to determine which of the dN or 2N method for Step 3, Fall-off Removal yielded structures closer to ones used for the simulation. Also, the parameters  $\theta_{low\_clip}$ ,  $\theta_{high\_clip}$ , and  $N$  were varied to see how to select them to achieve a consistently accurate algorithm.

#### *i. dN Results With No Roughness Introduced*

Using a simulation with no roughness present, such as the simulation of the 2nm of Al and 20 nm of SiO<sub>2</sub> discussed earlier, dN successfully extracted the number of layers present in the model and their corresponding thickness. It should be noted that although the oscillations resulting from applying the dN method in Step 3, Fall-off Removal are  $\frac{\pi}{4}$  out of phase with the original high frequency oscillations, it does not alter the frequency of the oscillations and thus does not alter the result of Step 5, Fourier Space Analysis.

One evident shortcoming of the dN method is the existence of a low frequency residue evident in Figure 7:



**Figure 7** Fourier transform result using the dN method on simulated data equivalent to that in Figure 4. Test run at 12keV, with a  $\theta_{low\_clip}$  of .2 and an  $N$  value of 3. The original data spans 9 degrees of  $\Theta$ .

This gradual low frequency curve spans from approximately 0 to 10 nm and ‘drowns out’ a peak that should be visible at 2nm in Figure 7. The existence of this 2nm layer is then only revealed through the difference in peaks at about 22 and 20 nm.

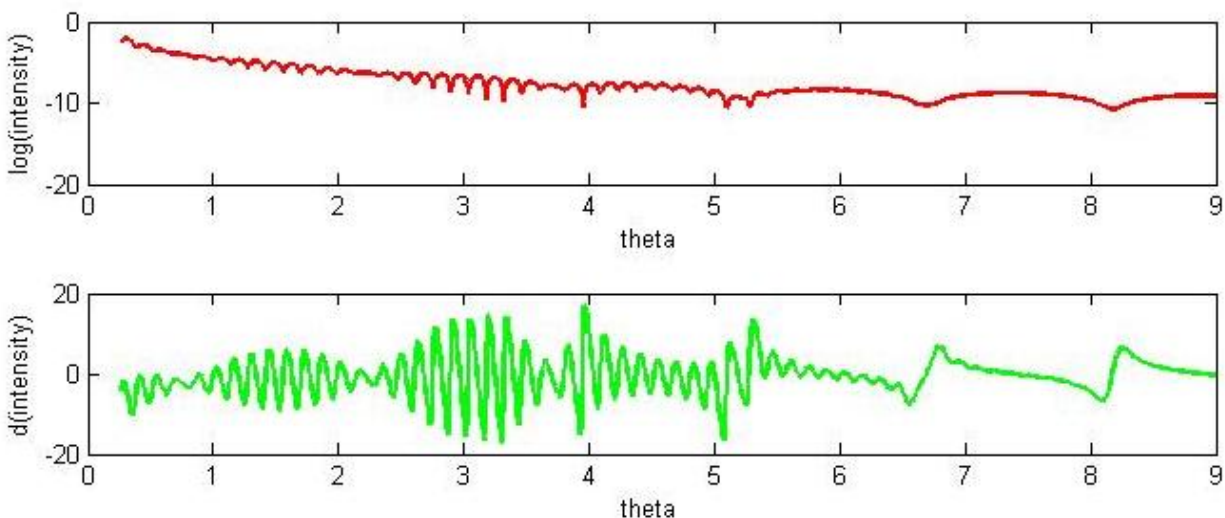
Another observation of the dN results in Fourier space is the relatively strong second harmonics signals present around approximately 45 nm in Figure 7. The presence of pronounced harmonics such as these indicates the dataset that is being Fourier transformed is not very sinusoidal in nature.

Lastly, the characteristics of a multilayer stack that are most readily evident from this method are the number and thickness of layers present. This capability has only been tested on one and two layer samples that are situated on a substrate.

#### *ii. dN Results on simulated roughness*

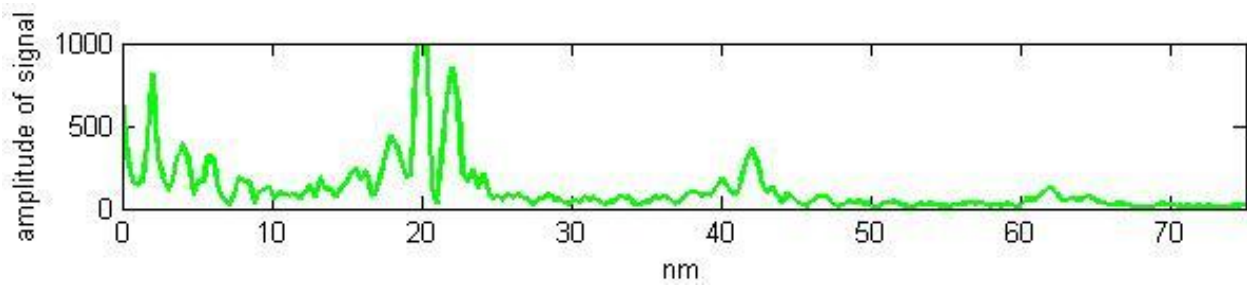
Under actual experimental conditions, roughness is present on the layers of any real sample being tested since no semiconductor sample can be perfectly smooth. To simulate this, the same simulation above is analyzed, yet with 1.5 Å of roughness added to the SiO<sub>2</sub> layer.

Figure 8 show how the roughness affects the intensity versus  $\Theta$  signal and the corresponding application of the dN method, while Figure 9 specifically shows the Fourier transform result.



**Figure 8** The top graph is the intensity versus theta dataset for a simulated sample defined in Figure 4 – with 1.5 Å of roughness added to the SiO<sub>2</sub> layer. The second graph is the result of applying the dN method in Step 3, Fall-off Removal. 13

Note how the high frequency component is greatly attenuated, especially after 7 degrees of theta. In turn, this affects the derivative in this region too, introducing a pulse like signal.



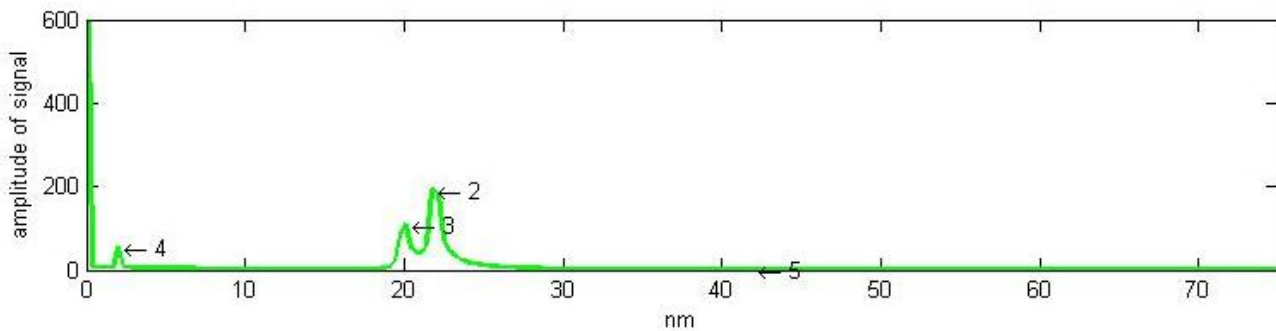
**Figure 9** The **Step 5, Fourier Space Analysis** result of the data shown in Figure 8.

Note the many different peaks that have been introduced by the addition of roughness, as compared to Figure 7 when no roughness was present

As can be seen in Figure 8, the roughness greatly attenuates the high frequency oscillation for larger  $\Theta$ . Although the Fourier transform result in Figure 9 is very noisy, a sizable peak is now centered over approximately 2nm. Additionally, the peak with the largest amplitude is now over 20nm, rather than the largest amplitude peak representing the summation of both layers (22nm). The Fourier transform also contains many large “ringing” oscillations, which do not correspond to any real layer spacing.

### *iii. 2N Results With No Roughness Introduced*

Using a simulation with no roughness, such as the 2nm of Al and 20 nm of SiO<sub>2</sub> discussed earlier, the 2N method extracted the number of layers present in the model and their corresponding thickness. This success is illustrated in the Figure 10 Fourier transform:



**Figure 10** The Fourier transform result from **Step 5, Fourier Space Analysis** using the 2N method on the dataset defined in Figure 4 with  $N$  set to 18 and  $\theta_{low\_clip}$  equal to  $0.18^\circ$ .

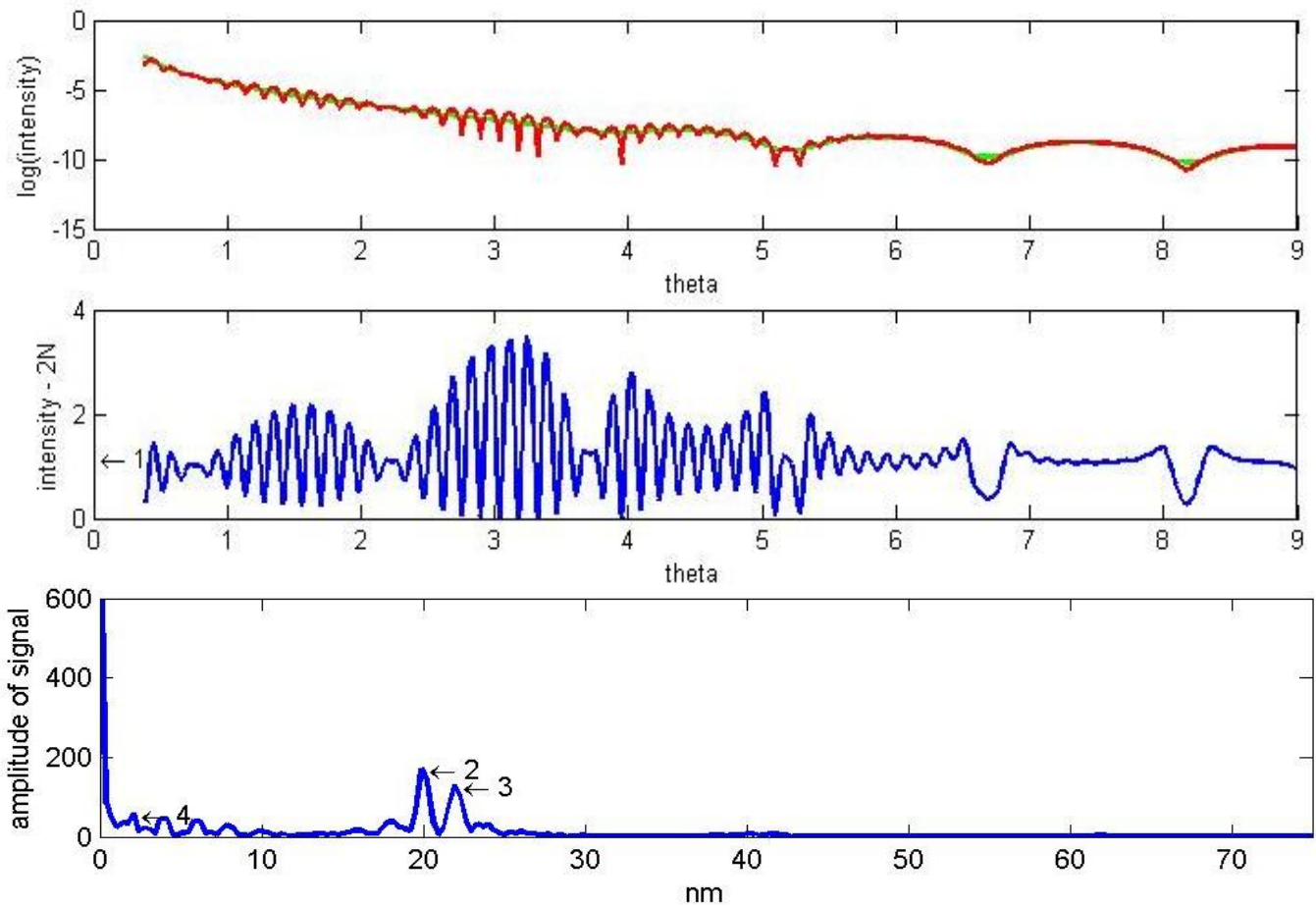
Note the existence of peak 4, indicating the 2nm layer. This peak is not visible in Figure 7, when the dN method is used.

Observing Figure 10, we see three distinct peaks. The first, peak four, corresponds to a thickness of 2.01 nm while the third corresponds to 20.1 nm. Peak two is approximately the sum of these

two layers. We notice that the second harmonics are extremely small as indicated by peak 5, suggesting that dN method enhances the higher harmonic contents. One anomaly is noted for a peak that appears to correspond to 0 nm. The reason for its existence and its physical meaning is not understood. We have only used the Fourier method so far to extract information about the number and thicknesses of layers. It is likely that the height, width and the shape of the peaks in the Fourier space contain information about roughness and density of the layers, but we haven't explored those characteristics yet.

*iv. 2N Results on simulated roughness*

Figure 11 shows the result of running the 2N algorithm on the same roughness simulation described previously in Figure 8 and 9 in the dN section



**Figure 11** The top graph is the intensity versus theta dataset for a simulated sample defined in Figure 4 – with 1.5 Å of roughness added to the SiO<sub>2</sub> layer. The second graph is the result of applying the 2N method in Step 3, Fall-off Removal. The bottom graph is the Fourier transform result from Step 5, Fourier Space Analysis using the 2N method. *N* is set to 18 and *theta\_low\_clip* equal to 0.3°.

Note the introduction of noise from the roughness, especially in the low frequency regions.

Just as with the dN simulation, there is the introduction of additional oscillations that do not correspond to any real layer thickness, especially in the low frequency region, though they are not as pronounced as in the case of dN method. These roughness-driven additional oscillations are rounded peaks, while the thickness of the 2 nm layer is denoted by a pointy peak of greater amplitude, as shown in peak 4 in Figure 11. Also, the peak with the largest amplitude (peak 2) now corresponds to 20nm, rather than the largest amplitude peak representing the summation of both layers (peak 4). The same effect occurred using the dN method.

v. Results of altering the  $\theta_{low\_clip}$  and  $\theta_{high\_clip}$  bounds on both algorithms

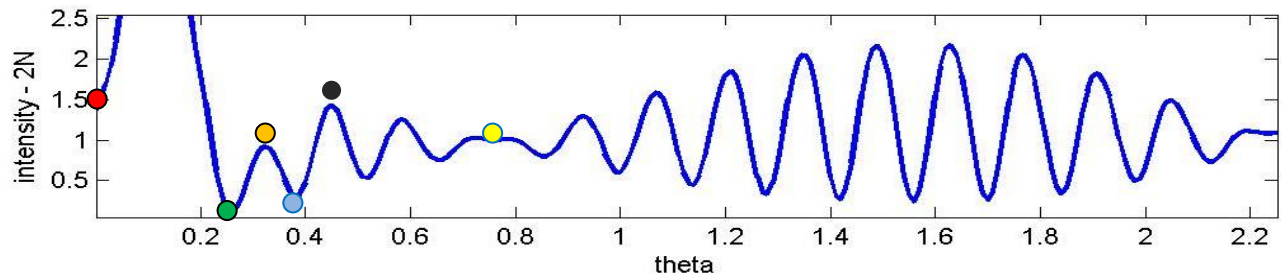
The most sensitive input parameter for both algorithms appears to be the lower  $\Theta$  bound used in Step 2, Restrict the  $\Theta$  Domain, a parameter defined as  $\theta_{low\_clip}$ . To highlight this fact, Table 1 illustrates the effect of altering the  $\theta_{low\_clip}$  used in applying the 2N method. For additional context, the location of the  $\theta_{low\_clip}$  is plotted in Figure 12.

$\theta_{low\_clip}$ Used [°]	2 nm Layer Result [nm]	20 nm Layer Result [nm]	Total Deviation [%]
● 0	1.97	20.03	1.78
● 0.22	2.02	19.86	1.78
● 0.33	2.38	20.46	23.58
● 0.38	2.05	19.89	3.72
● 0.45	2.42	20.40	25.13
● 0.75	2.50	20.42	27.50

**Table 1** Results of changing the lower bounds of theta,  $\theta_{low\_clip}$ , when the 2N method is applied to the simulation detailed in Figure 4 but with 1.5 Å of roughness added to the SiO<sub>2</sub> layer.  $N$  is held constant at 20 throughout. The colored dots correspond to the points on the Step 3, Fall-off Removal graph in Figure 12 below.

Total deviation is the percent that the 2nm result differs from the actual value plus the percent that the 20nm result differs from the actual value.

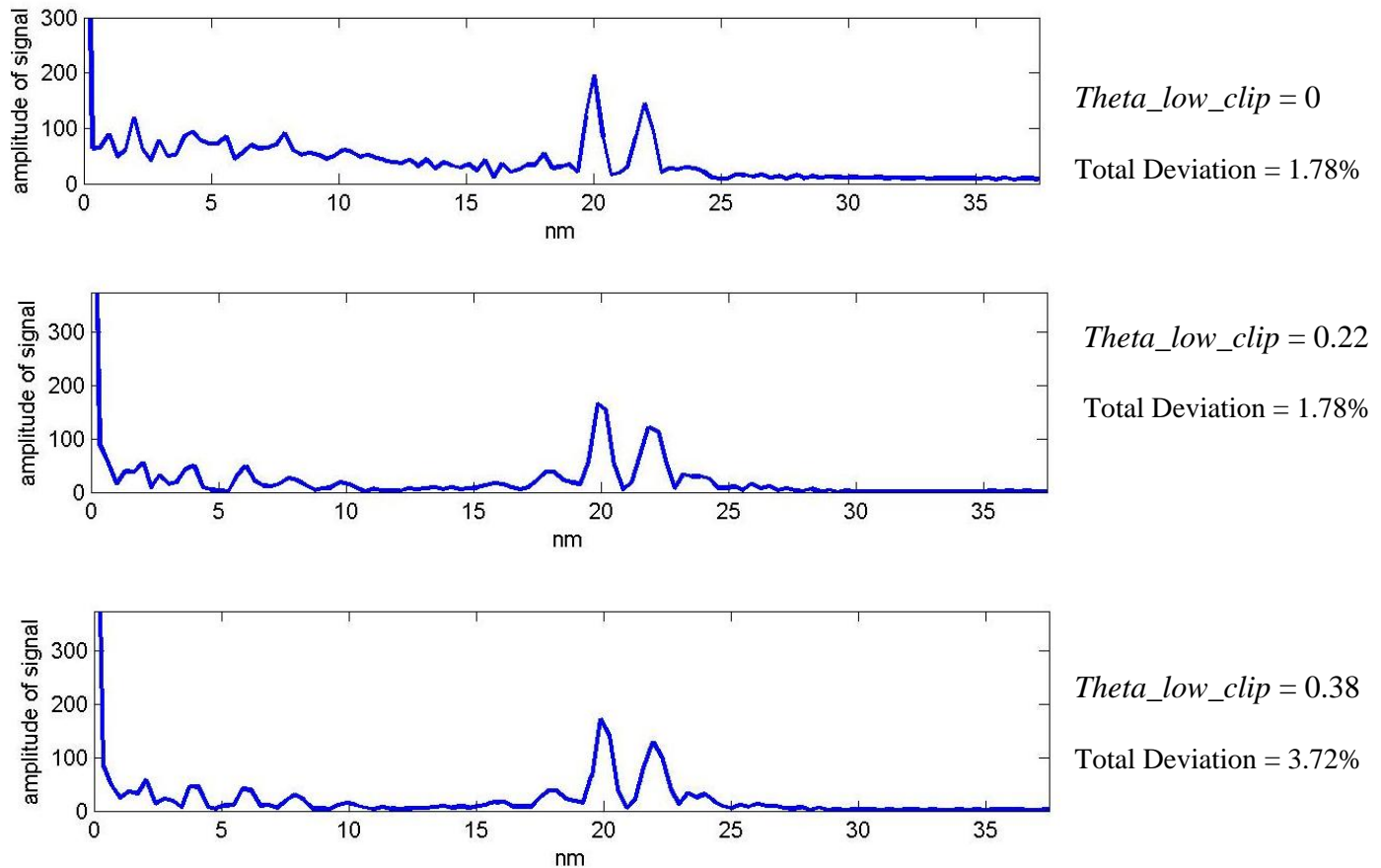




**Figure 12** A graph of the first part of the extracted oscillations resulting from the **Step 3, Fall-off Removal** with colored dots corresponding to Table 1.

Note the large spike below theta equal to 0.2 and how the oscillations then settle into a regular pattern after a theta equal to about 0.8.

To further analyze these findings, Figure 13 displays the Fourier transform of the three *theta\_low\_clips* that have the smallest deviations.



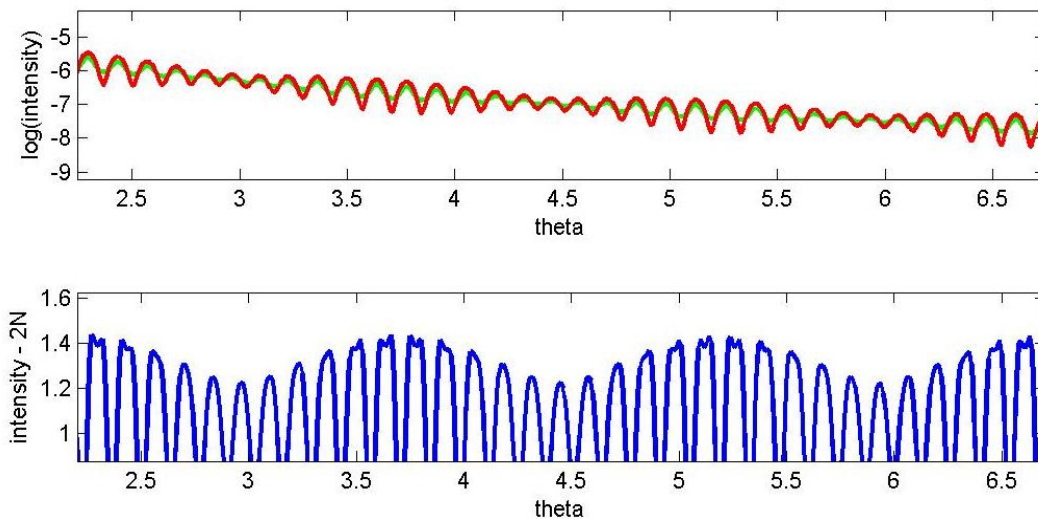
**Figure 13** Fourier transform graphs, the results of **Step 5, Fourier Space Analysis** on the dataset described in Table 1. All elements are held constant between the graphs, except for *theta\_low\_clip* values chosen because they deviated the least in Table 1. Note that although the *theta\_low\_clip* value of 0.38 thicknesses deviates the most, the false peaks in the transform are rounded, leaving the true peak at approximately 2nm pointed. The peaks in the other transforms are prominently pointed in nature, which could lead to multiple false readings of small layers

Looking at the results of Table 1, Figure 12, and Figure 13 it is observed that choosing the  $\theta_{low\_clip}$  to be the second minima in the high frequency oscillation,  $0.38^\circ$  in the example case, yields the best combination of layer thickness and true layer detection accuracy. Choosing alternative  $\theta_{low\_clips}$  that are more accurate in terms of layer thickness appear to falsely detect other thin layers.

Additional tests found that using a  $\theta_{high\_clip}$  always lowered the accuracy of the algorithm when used against various simulations. Tests were not tried on actual data.

*v. Results of altering N value used for the 2N method*

It was determined that the most sensible way to choose an appropriate N value was to observe the visual output of Step 3, Fall-off Removal, both the fall-off line created by the 2N method and the corresponding graphic with this fall-off line removed. As noted in Figure 14 and

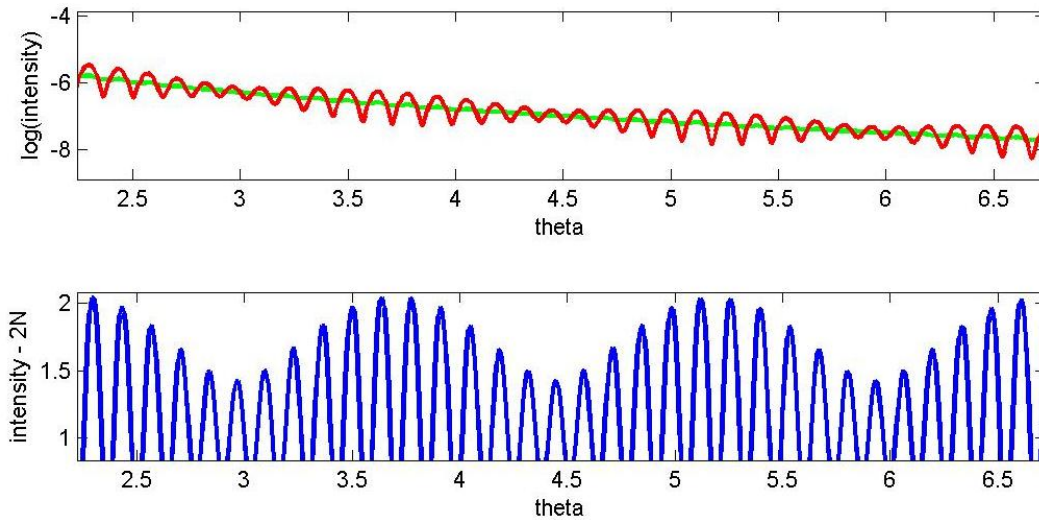


**Figure 14** An  $N$  value of 4 is used on the dataset defined in Figure 4. The green line in the top picture indicates the fall-off the 2N method is calculating.

In this case,  $N$  is too small, thus the green fall-off calculation too closely matches the original function resulting in the ‘clipping’ seen in the lower blue oscillation graph.

Figure 15, an appropriate  $N$  value allows the full oscillation to be extracted, rather than an oscillation with its peak clipped – such as Figure 14.

It is concluded that an  $N$  value equivalent to half the points in one period of the high frequency oscillation accurately maps the fall-off function present in the original data. Mathematically, this is intuitive because the  $2N$  method will span a full period of an oscillation when it takes a local average.



**Figure 15** An  $N$  value of 6 is used on the dataset defined in Figure 4. The green line in the top picture indicates the fall-off the  $2N$  method is calculating.

In this case, the green line accurately captures the fall-off with  $N$  chosen to be half the number of points in a period of the high frequency oscillation.

#### IV. Discussion & Conclusion

Separating the reflected beam intensity into a smooth, monotonic falloff term and an oscillatory terms, and then Fourier transforming the oscillatory terms is a quick and powerful way of extracting key characteristics of the layers from a sample containing multiple layers.

Based on the simulation test results described in the Results section, it appears that  $2N$  fall-off subtraction distorts the Fourier space data less than the  $dN$  method. Therefore, it is recommended that the  $2N$  algorithm be employed to assist in extracting information from the

XRR data, particularly the number of layers present and their thicknesses. The key distinction that sets the 2N algorithm apart from the dN algorithm is its low frequency sensitivity, a characteristic that is critical to the mission of detecting thin layers on samples. The low frequency artifact that exists in the dN method, as manifested in Figure 7, is simply unacceptable.

As noted in the Results section, it is recommended a *theta\_low\_clip* corresponding to the second minima of the high frequency oscillation be used with the 2N method. This recommendation is done with caution, due to the sensitivity of this parameter and it's recommended that this parameter's affect on the output of the algorithm be studied additionally. No *theta\_high\_clip* should be used, however, it may be useful to remove a high theta region of experimental data that traditionally contains a high noise to signal ratio – this needs to be further explored. Experience has determined that using an *N* value equal to at least half the number of points in the high frequency period yields the cleanest extraction of the fall-off function.

It is concluded that using the 2N fall-off removal method combined with the Fourier analysis is a powerful technique that can reveal the presence of nanoscale layers and their corresponding thicknesses. This study did not focus on the intensity and the shape of the peaks in the Fourier space, but there is some preliminary evidence that they may hold information about layer densities and order. It is recommended that this be explored in more detail.

Even if no connection is established, the current capabilities of this technique are very useful in analyzing new semiconductor structures that often contain multiple layers of varying thicknesses. Just knowing the number of layers present and their thicknesses eliminates many important unknown variables, making modeling of additional characteristics such as layer order and density less labor intensive. This partial information is often sufficient enough to guide the

generation of appropriate stratigraphic models which can then be refined against the raw XRR data using conventional goodness-of-fit metrics, ultimately resulting in a reliable characterization of complex multilayer samples.

## V. Acknowledgments

This work was supported by the SULI Program, U.S. Department of Energy, Office of Science. I would like to thank my mentor Apurva Mehta and Matthew Bibee, a graduate student at Stanford and a former SULI intern, for their assistance in addition to all the administrators of the SULI program for this wonderful and enriching experience. Finally, I would like to thank Yan Gao, of GE, for providing interesting samples for testing the effectiveness of the analysis approach taken here.

## VI. References

- 
- <sup>i</sup> Moore, Gordon E. (1965). "Cramming more components onto integrated circuits" . Electronics Magazine. pp. 4.
- <sup>ii</sup> Toney, Michael and Brennan, Sean. "Measurements of carbon thin films using x-ray reflectivity". Journal of Applied Physics 66 (4). 15 August 1989.
- <sup>iii</sup> Discussion with Apurva Mehta. August 4<sup>th</sup>, 2011.
- <sup>iv</sup> Fronheiser, Jody and Tilak, Vinayak. "Motivation for SiC Device Development and Relevance to NIST/GE Contract on SiC/SiO<sub>2</sub> Interface Measurement Technologies". March 27<sup>th</sup>, 2011.
- <sup>v</sup> O. Durand. "Characterization of multilayered materials for optoelectronic components by high-resolution X-ray diffractometry and reflectometry: contribution of numerical treatments". Thin Solid Films Volume 450, Issue 1, Proceedings of Symposium M on Optical and X-Ray Metrology for Advanced Device Materials Characterization, of the E-MRS 2003 Spring Conference. 22 February 2004. Pg. 51-59
- <sup>vi</sup> O. Durand.
- <sup>vii</sup> Toney, Michael and Brennan, Sean
- <sup>viii</sup> Poust, Benjamin and Goorsky, Mark. "Enhanced Fourier Transforms for X-Ray Scattering Applications". Fourier Transforms – Approach to Scientific Principles. InTech. 2011
- <sup>ix</sup> Poust, Goorsky
- <sup>x</sup> Poust, Goorsky