

Solving surface structures from normal incidence x-ray standing wave data

Mark Basham and Roger A. Bennett

Department of Physics, University of Reading, Whiteknights, Reading, RG6 6AF, United Kingdom.

Abstract

A program is provided to determine structural parameters of atoms in or adsorbed on surfaces by refinement of atomistic models towards experimentally determined data generated by the normal incidence x-ray standing wave (NIXSW) technique. The method employs a combination of Differential Evolution Genetic Algorithms and Steepest Descent Line Minimisations to provide a fast, reliable and user friendly tool for experimentalists to interpret complex multidimensional NIXSW data sets.

PACS : 68.49.Uv, 68.43.Fg

Keywords: Normal incidence x-ray standing wave, Experimental Data Fitting, Genetic Algorithms, Adsorbate, Surface Structure.

Program Summary

Title of program: NIXSW Planewave Solver

Catalogue identifier:

Program obtainable from: CPC Program Library, Queen's University of Belfast, N. Ireland

Computers: Any Windows Compatible

Operating systems: Windows 2000 and XP

Programming language used: Borland C++ Builder 5

Memory required to execute with typical data: <10MB

Distribution format: Zipped Executable and source code

Libraries used: Borland Standard, included with distribution

Nature of physical problem: Using NIXSW experimental data to calculate atomic positions of adsorbates.

Restrictions on the complexity of the problem: Data from substrates must have cubic, tetragonal or orthorhombic crystal structures i.e. with 90° between conventional cell axes.

Typical running time: Seconds-Minutes dependant on the number of plane waves and the number of atomic sites.

Introduction

The determination of the geometric arrangement of atoms and molecules adsorbed at or within the surface of a material has been a central theme of scientific endeavour due to the importance that surfaces play in modern technologies. Understanding atomic arrangements has led to the development of the semiconductor industry through ultra-high quality growth of functional layers on surfaces; it has guided and shaped the development of catalysts which now enable harmful automobile exhaust emissions to be ameliorated and will be at the forefront of new technologies such as fuel cells, spin polarised electronic devices and functionalised nanomaterials. Methods of surface structure determination fall into two broad camps, the first employs diffractive techniques to sample relatively long range ordered structures such as clean surfaces (e.g. low energy electron diffraction and surface x-ray diffraction) and the second employs a chemically specific modulation of a localised excitation to discern local geometry (such as photoelectron diffraction and surface extended x-ray absorption fine structure). Recently the development of high brightness tuneable synchrotron x-ray sources has enabled a new technique to be developed in which diffraction from the crystalline bulk is used to locate and identify the atoms in the surface region. This class of technique is known as x-ray standing wavefield (XSW)

absorption and a variant at normal incidence (NIXSW) has proven particularly effective in determining the structure of a large range of adsorbates and overlayers on surfaces. In general the determination of surface structure requires the fitting of model candidate structures to the experimental data. In this paper we present a novel implementation of genetic algorithms and steepest descent minimisation to solve for complete data sets of pre-processed NIXSW data in the presence of complicating factors such as multiple adsorption sites, non-high symmetry adsorption or for experimental scatter in results. The program takes as an input series of measured coherent fractions and positions for each NIXSW experiment and outputs an optimised structural solution.

The NIXSW technique has recently been comprehensively reviewed by Woodruff [1] so here we will restrict ourselves to a basic description of the method and the data reduction commonly applied to produce the parameters which are the input to our program and from which conclusions are experimentally drawn. The origin of the technique stems from the formation of a standing x-ray wavefield by overlapping forward propagating x-rays into a crystal with back diffracted x-rays moving out from the surface when the x-ray wavelength satisfies a Bragg scattering condition. In practice the crystal is oriented in space such that the direction of forward propagation is along the normal to an atomic plane within the crystal and the Bragg condition is met by tuning synchrotron x-rays to the wavelength of twice the lattice spacing. For a single scattering (kinematic) analysis one would expect the Bragg condition to be satisfied at just one x-ray wavelength which gives rise to a perfect reflectivity. However, multiple scattering and crystal absorption lead to a broadening of the energy at which reflectivity occurs and within this energy range the phase of the

standing wave moves by π - that is the standing wavefield moves with respect to the scattering planes of atoms within the bulk of the crystal. At the surface this moving wavefield can locate nodes or antinodes on adsorbed and surface atoms. At nodal points the electric field strength is a minimum and x-ray absorption is also minimal whereas at antinodes the electric field strength is enhanced and so is the x-ray excitation of the atom in the field. By measuring the strength of absorption (by photoelectron emission, Auger electron emission or fluorescence) as a function of the x-ray energy (and hence relative position of the antinodes to the crystal planes) an elemental distribution can be produced for the atoms perpendicular to the scattering plane. There are of course many fine details to take account of but the most common method of analysis [2] yields two structural parameters: the coherent fraction f_{co} , and the coherent position D , which are the amplitude and phase of the 1-D Fourier representations of the elemental profile in the particular Bragg scattering direction, Eqn. 1.

$$f_{co} \exp\left(\frac{2\pi i D}{d_{h,k,l}}\right) = \int_0^{d_{h,k,l}} f(z) \exp\left(\frac{2\pi i z}{d_{h,k,l}}\right) dz \approx \sum_n f(z_n) \exp\left(\frac{2\pi i z_n}{d_{h,k,l}}\right) \quad \text{Eqn 1.}$$

In experiments we determine f_{co} and D from fitting experimental data [3] and use those to determine the real space positions (z) of the n absorbing atoms in the Bragg scattering direction. Normally the coherent position is written as a fraction of the interatomic spacing ($d_{h,k,l}$) of the scattering plane where h,k,l are the Miller indices of the plane, $P = D/d_{h,k,l}$. In our program we take f_{co} and P for each scattering plane as inputs. For a single atomic adsorption site a NIXSW experiment would return a

coherent fraction of 1 and the position would reflect the adsorbed atoms height z above the scattering plane.

For more than one absorbing atom within the profile, at differing heights z from the scattering plane, the coherent fraction and position become a coherent average over the spatial distribution. Further complications arise when multiple adsorption sites are present in differing concentrations. For simplicity these are taken here as the normalised fractional occupancies $f(z)$, however these may in reality be reduced slightly by thermal vibrations [1]. A single NIXSW experiment in any given direction generates two structural parameters describing the arrangements of the atoms with respect to the plane. However, a geometrical model for n adsorbed atoms generally requires a z position above the scattering plane and an occupancy (normalised to one) for each atom generating $2n-1$ free parameters. Thus to obtain unique solutions in this direction one must either measure successively higher Fourier components [4] or solve with other non-orthogonal scattering planes. Given a sufficient experimental sample size the real space positions and fractional coverages (or equivalently occupancies) for the full three dimensional problem can be recovered by triangulation.

In what follows is a description of our method for solving the full 3D geometry of a surface structure by reconstruction from the experimentally determined coherent fractions and positions in multiple Bragg scattering geometries. In some cases it is possible to constrain the fitting through a knowledge of adsorption site symmetry (eg different domains at a surface), through independent determination of coverage or structural parameters and we have maintained maximum flexibility in this regard.

Program Specification

The program presented here is designed to obtain a set of atomic coordinates (x,y,z) for each atom, which most closely resemble the collected experimental data. The atomic coordinates which we want are related to those experimental values by equation 1. For each experiment we have the amplitude and phase of the 1-D Fourier representations of the elemental profile in the particular Bragg scattering direction and so for multiple experiments we rapidly form a landscape of possible real space distributions that could fit the data. We aim to minimise the discrepancy between experimental results and the projected distribution. The following approach is therefore applied to calculate the error on a given array of atomic coordinates.

1. Use the atomic positions and proportions to calculate a theoretical value for the real and imaginary components in the summation of Eqn.1 for each atom in every Bragg direction that has been experimentally taken.
2. Take the summation over all the atoms of these contributions to determine the resultant complex number which represents the predicted NIXSW response of the model structure in each Bragg direction. This fit can then be plotted on an Argand Diagram for each Bragg direction.
3. Plot the experimental data on the Argand Diagram in each Bragg direction.
4. Use root mean square of the distances between the model point and the experimental point on the Argand Diagrams for all Bragg directions as an evaluation of the error in the fit.

This error function (Objective Function) needs to be minimised, as when the error is a minimum the atomic positions and proportions most accurately reflect the experimental data. This minimisation problem can be solved in a series of ways, but

first it is important to keep the number of free variables i.e. different atomic coordinates, as small as possible. This can be done simply by using surface symmetry inherent to the single crystalline substrates, for example on the $\text{TiO}_2(110)$ surface an adsorbate may bind to either side of the bridging row oxygen atoms that run in the (001) direction. Thus there are two sites with identical height above the surface and with the same displacement along the c-axis. Transposing the x and y coordinates gives the two sites (see Figure 1). This means that instead of treating this two atom problem as one with 6 independent variables, one for each direction of each atom, it can be treated as a problem of 3 independent variables. Including such symmetries allows for the inclusion of domains in overlayers in a simple manner, however, for chiral adsorption systems such symmetry constraints may not be appropriate and the user is free to implement as appropriate. The program is currently limited to crystal substrates which have cubic, tetragonal or orthorhombic crystal structures i.e. with 90° between conventional cell axes of the bulk. There are no assumptions made about the surface geometry or orientation.

Minimisation Algorithms

Now that an error function has been constructed, and the number of variable parameters in the problem reduced as much as possible by symmetry, a solution can be sought. This program uses two different methods to find the best atomic coordinates to match any set of experimental data. The two techniques complement one another and allow efficient searches of parameter space and tight refinement of final solutions.

Steepest Descent Line Minimiser

This simple algorithm is capable of finding local minima in configuration space with relative ease and speed [5]. It uses the following process:

1. Numerically calculate the derivative of the error function with respect to each of the variables.
2. Using this direction through N dimensional configuration space, perform an efficient line minimisation using a parabolic interpolation.
3. Repeat this process until the required level of accuracy is achieved.

This simple technique always finds local minima, which can be useful for checking configurations which the user thinks could be correct after some refinement. The error found at the minima, and displayed bottom left of the User Interface, can be used for determining which structure of many may be a better solution.

Differential Evolution Genetic Algorithms

This second and more complex technique is for solving the problem globally, in that it tries to find the best solution of all available ones [6,7]. This method uses the following process; a glossary of the terms used can be seen in Table 1

1. Create a population of genomes, the first epoch.
2. Perform two binary tournaments each between two random genomes
3. The two winners are 'bred' and create a child
4. With a small probability, mutate the new child genome
5. Add this new genome to the next epoch
6. Go to 2 and continue until the next epoch is filled
7. Once the next epoch has a full population, the old population can be removed and the process repeated from 2 but using this new epoch

8. Once a certain number of Epochs have been reached or all the genomes in the population are identical stop the process
9. Go through all the last populations' genomes and find the fittest individual.
The best fit for the system is described by this genome

All the structural parameters in the GA are periodic over the unit cell spacing, and so they are always rescaled to be between the boundaries of +/- 1 lattice constant of the appropriate direction. The part of the genome which is responsible for dealing with the proportion of the configuration is kept bound between 0 and 1. When the error function takes these proportions it scales them to be the correct proportions given the number of atoms in the system, but the values are left as they are in the GA. This gives a much smoother function for the GA to fit to, and hence improves the fit dramatically.

Using the program

The program uses data from pre-processed fitting of experimental NIXSW responses for multiple Bragg reflections to form a complete dataset from which the structure may be solved. The solution and the predicted coherent fractions and positions for each Bragg reflection are shown alongside the input experimental data, Figure 2. Once entered the experimental data and the model structures can be saved out to a comma separated variable file for ease of use.

The first step is therefore to input the information through the User Interface of the application, which on first running opens in a single window with four regions into which entries must be made. The unit cell parameters of the substrate (in Angstroms)

must be entered first. Simply click into the boxes and type the required values. The directions $\hat{a}, \hat{b}, \hat{c}$ define the co-ordinate system employed such that x is in the \hat{a} direction, y is in the \hat{b} direction and z is in the \hat{c} direction. The origin of this co-ordinate system is material dependant and defined by the structure factors employed in pre-processing the NIXSW data to give the coherent fractions and positions.

The experimentally determined coherent fractions and positions for each Bragg reflection are entered in the Wave parameters area. Right click within the wave parameters panel and select “Add Wave”. You are then able to enter values for the h, k and l Miller indices relating to the Bragg scattering direction and d_{hkl} (the interplanar separation in that direction) is calculated for you. The experimentally determined values for the Coherent Position (P) and Coherent Fraction (Fco) should then be entered. A single row of data entered in this way therefore comprises the results from a single NIXSW experiment (a Wave) and an option to include or exclude the Wave from the fitting is given in the last column (enter “Y” or “N”). This process should be repeated for as many waves as you have data for.

If there are multiple experiments for the same Bragg reflection it is best to add every experimentally determined result separately rather than to average the coherent fractions and positions and enter a single result. The fitting algorithm used in the structure search weights each Wave equally so multiple experiments enable a more refined fit in the presence of experimental scatter and allows outliers to be more readily identified. A visualisation of the quality of fit for each Bragg reflection is available as an Argand diagram and is described below.

Once this is done you should add the Adsorbate Parameters. This is done in a similar way, right click on the area and select “Add Atom”, and a new row is created in which entries may be made in the leftmost cells. The required values for the x, y, z and proportion boxes are not the actual coordinates but numbered links to parameters held in the Variables field to the left. A typical (x,y,z) position would be entered as (1,2,3) while symmetrical variants can be obtained by transposing or inverting such as (2,1,3) (2,-1,3) etc. The real space positions of these atoms are then looked up from the Variables field. In this way symmetric atomistic positions, such as domains, can be accounted for without increasing the number of variable parameters in the solution. This is important as the fewer variable parameters the faster and more feasible the fit is.

Starting values for input in the Variables field should be entered by right clicking and selecting “Add Variable”. The variable is numbered and that number corresponds to the numbers entered in the Adsorbate Field. The values entered here are the atomic positions in Angstroms and should close to the expected positions of the atoms (the GA minimiser searches much of the parameter space effectively so these initial guesses can be very crude). Values of the variables which are mapped to the Proportions entry in the Adsorbate Parameters are the fractional occupancies and are automatically scaled to sum to 1 over all the atoms in the developing model.

The option is given to include or exclude (Y/N) experimental Wave Parameters or Adsorbate Parameters in the fitting procedure. It is also possible to fix some of the Variables such that they become Parameters (V/P) in developing the model. This is of

use if some atomic positions are already known. Values labelled as Parameters in the Variable field are included in the analysis but are not changed in the fitting procedure.

Application

As an example of the use of the programme we demonstrate below the triangulation of the atomic positions of Rh atoms (in a metal-organic molecule) adsorbed on a $\text{TiO}_2(110)$ surface [8]. The first task is to enter the lattice constants of the substrate which has Tetragonal symmetry ($a=b=4.584$, $c=2.953$). Experiments were performed which gave the coherent fractions and positions for four different Bragg conditions, these are shown in Table 2

We now need to build a model structure, the simplest of which would be a single atom per surface unit cell. However, in this example the surface unit cell is rectangular (see Figure 1) and therefore we may expect four equivalent adsorption sites to be occupied at lowest symmetry (or equivalently four domains if the adatoms form long range structures). If our first atom is at position (1,2,3) (remembering 1,2,3 correspond to Variables 1,2,3) then the others are at (1,2,-3), (2,1,3) and (2,1,-3). They are symmetrically equivalent and therefore would be expected to be equally populated, which is our fourth Variable 4. Selecting solve and either minimiser results in an optimised fit for this model with an error of ~ 1.219 , Figure 2. Inspection by eye in comparison to the experimental data shows some significant discrepancies. Selecting Argand and view opens new windows showing the experimental data as blue dots and the model as red dots for each atom plotted on Argand diagrams for each Bragg reflection [2], Figure 4. With a perfect fit the sum of the atomic vectors (a red line and dot representing each atom's coherent fraction and position represented

as an Argand diagram) should end in the centroid of the blue dots in each direction. With one atom per unit cell there is clearly a poor fit so it would seem appropriate to add a second atom (5,6,7) with symmetrical equivalents and a new fractional coverage 8. Implementing a GA fit from here and refining through the SD minimiser should lead to an atomic arrangement with an error ~ 0.354 (the actual atomic arrangement determined is periodic in multiples of the crystal lattice constant and therefore several equivalent combinations may result, one example is shown in Figure 4). Noticeably the optimal fit is obtained if 78% of the atoms are in one position while 22% are in another and the c-axis values for these two sites are at the origin and half way along the unit cell. Chemical intuition must be used to determine whether the coordinates returned are reasonable or if the points lay a long way off the surface due to the multiples of the unit cell dimensions. Such multiples can be added or subtracted by hand and the SD minimiser re-run to refine. Producing the Argand diagrams now gives a much improved fit, Figure 5.

There are two options under the solve menu which alter the stopping behaviour of the steepest descent minimiser. The preferred method is to run the algorithm for a given number of steps which is set under the “Set number of SD steps” option (typically 1000). The alternative is to run the steepest descent algorithm until the error has fallen below a threshold per parabolic interpolation set by the “Set SD max error” option. It should be noted that this error is the difference between consecutive objective function calculations, and not the absolute value of the final error in the fit with respect to the experimental data.

Conclusions

The program provides a robust and intuitive exploration of the parameter space associated with atomic arrangements at surfaces determined from experimentally

derived reduced data NIXSW sets. Reliable structural models simulating the experimental data can be built rapidly and tested through manipulation of atomic positions, symmetries and surface occupancies.

Acknowledgements

We would like to acknowledge EPSRC (GR/T18738/01) and the Royal Society/Wolfson Foundation for funding and the CCLRC for experimental synchrotron beamtime which generated the test data.

Figure 1 – TiO₂(110) plan view of the surface. The white rectangle is a surface unit cell with the corners placed on 5-fold co-ordinated Ti atoms (small dark spheres). The larger red spheres are oxygen ions with bridging oxygen ions aligned vertically (c-axis) in the centre of the cell. The four crosses mark symmetrically equivalent surface sites $(x, y, z), (y, x, z), (x, y, \bar{z}), (y, x, \bar{z})$.

Figure 2 - All the Data Included and solved for 4 symmetric atomic positions.

Figure 3 - The Argand Diagrams for the 4 symmetric atom solution.

Figure 4 - All the Data Included and solved for 8 atomic positions.

Figure 5 - The Argand Diagrams for the 8 atom solution.

Table 1 - Glossary of genetic Algorithm terms

Table 2 - Experimental test data

References

- 1 D.P. Woodruff, Rep. Prog. Phys. 68 (2005) 743.
- 2 D.P. Woodruff, B.C.C. Cowie and A.R.H.F. Ettema, J. Phys.: Condens. Matter 6 (1994) 10633.
- 3 For example the NIXSW99#1 programme by R.G. Jones, Dept. of Chemistry, University of Nottingham, Nottingham, UK.
- 4 L. Cheng, P. Fenter, M.J. Bedzyk and N.C. Sturchio, Phys. Rev. Lett. 90 (2003) 255503.
- 5 "Numerical Recipes in FORTRAN - The Art of Scientific Computing" by William H. Press, Brian P. Flannery, Saul A. Teukolsky and William T. Vetterling (Cambridge University Press, ISBN-13: 9780521430647)
- 6 "Differential Evolution - A Practical Approach to Global Optimization" by Ken Price, Rainer Storn, and Jouni Lampinen (Springer, ISBN: 3-540-20950-6)
- 7 R. Storn and K. Price, Journal of Global Optimization, Vol. 11, (1997) 341.
- 8 R.A. Bennett, N.D. McCavish, M. Basham, V.R. Dhanak and M.A. Newton, Phys. Rev. Letts. 98 (2007) 056102.