Solving for the phase of STEM probes in real space.

K M Atkinson 1, F Sweeney and J M Rodenburg

University of Sheffield, Department of Electronic and Electrical Engineering, Mappin Street, Sheffield, S1 3JD, UK

1 E-mail: elp06kma@sheffield.ac.uk

Abstract. The combination of a Gerchberg-Saxton phase retrieval algorithm with a genetic algorithm is used to retrieve the phase of a TEM or STEM probe and pinpoint its defocus. Results from modelled data and initial experimental work are included.

1. Introduction
The intensity distribution of a probe in TEM or STEM can easily be recorded but the corresponding phase is unknown. Knowledge of this information is essential in using phase retrieval methods such as ptychography. Although in STEM it is possible to calculate aberration constants from the Ronchigram, errors in these values can change when moving from a defocused probe to a focused probe. We have attempted to by-pass these problems by retrieving the phase from a series of several intensity measurements at varying levels of defocus.

2. Phase Retrieval
As the intensity images of the probe are recorded without any known background support this rules out using many phase retrieval methods, however we can use Gerchberg-Saxton method [1]. This technique uses two intensity images recorded at different distances (defoci) from an object. An estimate of the phase is then combined with the known amplitude of the first image to produce a complex wave. This wave is propagated to the second image plane. The resulting amplitude is then compared with the actual amplitude recorded in the second plane and the difference used to form an error metric. The estimated phase is retained whilst the estimated amplitude is discarded and replaced by the correct amplitude for the second plane. The wave is then propagated back to the first plane and the process repeated. By continuing to iterate between the two images the correct phase is reached.

Here we have extended this technique to include more than two images, propagating from one to the next until the final image in the sequence is reached at which point we propagate the wave directly back to the first image. Tests on simulated probe images showed that the use to four of five images rather than just two significantly increased the speed at which the algorithm converged.

3. Genetic Algorithm
The Gerchberg-Saxton algorithm relies on the propagation distance between images been known accurately. If incorrect propagation distances are used the phase estimate may still improve over successive iterations however the error metric will never reach zero. When recording images the microscope can provide an estimated defocus however this is unlikely to be accurate enough. We must
therefore search through a range of potential defocus values close to this estimate for each probe image so there will be many possible combinations of values for a sequence of several images.

We have chosen to employ a genetic algorithm to search through this range of combinations. The algorithm performs the Gerchberg-Saxton phase retrieval on a selection of combinations for a set number of iterations and compares the resulting error metrics to rank the selected combinations from best to worst. The next generation of combinations is then created by the techniques of crossover and mutation [2]. It should contain many combinations similar to the best combinations from the previous generation in order to examine these sections of the search space more thoroughly, however it should also contain some significantly different combinations to avoid being caught in local minima. Over successive generations this algorithm should identify the most suitable combination of defocus values to use in phase retrieval.

4. Results
Initially the combined Gerchberg-Saxton genetic algorithm (GS-GA) was tested with modelled data to establish if it was working correctly and how robust it was under different circumstances. Afterwards some initial experimental data was evaluated using the algorithm.

4.1. Testing with modelled data
The model consisted of four probe images with differing defocus values, -600nm, -330nm, +30nm (focus) and +345nm respectively (where a positive value indicates overfocus and a negative value indicates underfocus). These probe images are all generated from an aperture with a semi-angle of 5mrad. The algorithm is supplied with amplitude images of each of these probes, an initial guess of -600nm, -300nm, 0nm and +300nm for the defocus values with a search range of +/-50nm for each value and was run over 200 generations.

Table 1 shows the results of these tests. When the modelled probes have no aberrations the algorithm has incorrectly fixed on an underfocused value rather than an overfocused one for probe 3. This is not surprising as without the presence of any aberrations the probe intensity will be identical at the same distance either side of focus. If the search range for the defocus value of probe 3 is reduced or spherical aberration is introduced then this problem no longer occurs. The presence of spherical aberration or any other aberration who’s effects differ either side of focus will eliminate this problem in experimental data.

| Probe characteristics                | Probe 1 | Probe 2 | Probe 3 | Probe 4 |
|--------------------------------------|---------|---------|---------|---------|
| True defocus values                  | -600.00 | -330.00 | 30.00   | 345.00  |
| Unaberrated probes                   | -617.33 | -347.33 | -47.30  | 327.70  |
| Probe 3 restricted to overfocus      | -606.15 | -336.15 | 23.83   | 338.82  |
| Spherical aberration, Cs=0.5mm      | -607.45 | -337.46 | 22.61   | 337.62  |
| Extra 50 generations with full size probe images | -595.52 | -325.53 | -25.54  | 349.47  |
| Cs=0.5mm and Poisson noise           | -594.66 | -331.24 | 25.58   | 345.96  |
| Cs=0.5mm and salt and pepper noise   | -598.36 | -290.93 | -14.60  | 307.75  |

The modelled probes were originally generated as images of 1024x1024 pixels. To speed up the running time of the algorithm these were reduced to 128x128 pixels. Using the original larger images.
for an extra 50 generations at the end of the algorithm did give a slight improvement in accuracy although the results showed only the first 5 of these generations were necessary.

The effect of noise was also simulated. Surprisingly the results generated in the presence of Poisson noise are actually the most accurate results in Table 1 however the results from salt and pepper noise simulations show that large quantities of noise can reduce the accuracy of the algorithm.

Interestingly the result for the unaberrated probes in Table 1 seems to accurately predict the difference in defocus between consecutive probes (except for the erroneous result for probe 3) but the defocus values are incorrect by -17.3nm. This indicates that the algorithm may be most effective in calculating the difference in defocus between two consecutive probes. With this in mind the algorithm was run again with the same test data but was altered to output the difference in defocus between consecutive probes rather than specific defocus values. The algorithm was given an initial guess that all probes were separated by 300nm +/- 100nm. The results of this are shown in Table 2.

Again we can see that without any aberrations probe 3 produces erroneous results however with some spherical aberration introduced the results appear to be more accurate than those in Table 1. Unfortunately this approach does seem to be less resilient to noise.

| Table 2 – Results from GS-GA searching only for difference in defocus between consecutive probes. |
|---------------------------------------------------------------|
| | Probe characteristics | Defocus change (nm) |   |
| | | Probe 1-2 | Probe 2-3 | Probe 3-4 |
| True defocus changes | 270.00 | 360.00 | 315.00 |
| Unaberrated probes | 270.03 | 299.91 | 375.00 |
| Probe 3 search range 350+/-50nm | 269.97 | 300.03 | 375.00 |
| Probe 3 search range 375+/-25nm | 269.97 | 360.02 | 315.02 |
| Spherical aberration, Cs=0.5mm | 269.97 | 359.99 | 315.02 |
| Cs=0.5mm and Poisson noise set1 | 263.23 | 282.26 | 390.29 |
| Cs=0.5mm and Poisson noise set2 | 282.08 | 365.62 | 217.21 |
| Cs=0.5mm and salt and pepper noise | 306.50 | 268.11 | 317.22 |

4.2. Experimental data

Data was obtained on a JEOL 2010F microscope at the University of Sheffield. Sequences of probe images were obtained, beginning with a highly underfocused probe and incrementing the defocus in steps of 32nm or 64nm, recording the image at each point until a highly overfocused probe was reached. Selections of four probes were taken from these sequences to form data sets. An estimate for the defocus of a probe in the sequence was obtained from the relative change in defocus between that image and the image obtained closest to focus. These estimates were supplied to the algorithm to form an initial guess around which to search.

Table 3 shows the results from a typical pair of overlapping data sets. Probe 1 and 2 in data set A are the same as probes 3 and 4 in data set B. By including a pair of probes in two different data sets we were able to test whether the results were consistent with one another. Each data set was put through two versions of the algorithm, one to find specific defocus values for each probe and the other to look at only the difference in defocus between probes. We can see from Table 3 that the results from both versions of the algorithm are self consistent for each data set. We found this to be true of most but not all data sets tested so far. Unfortunately the results from the two data sets are not consistent with each other despite containing two of the same probes.
The convergence of the algorithm was monitored over each generation. For modelled data the error metric decreased rapidly over the first 50 generations then further small decreases were observed up to the 200th generation when the algorithm terminated. For the experimental data however a decrease was only observed over the first 20 generations then the error varied up and down a little each generation with the overall trend being a plateau. This indicates that the Gerchberg-Saxton part of algorithm is struggling to estimate the phase. The most likely cause of this and the inconsistent results shown in Table 3 is that when the defocus is changed by altering the current through the objective lens this changes the convergence angle significantly enough to disrupt the algorithm [3]. Subsequent testing with modelled data showed the algorithm convergence would be significantly impaired by only a 1% change in convergence angle over a defocus range of 950nm. Further work is required to determine the scale of this effect experimentally.

Table 3 – Typical experimental results obtained from GS-GA.

| Data set | Defocus (nm) | Probe 1     | Probe 2     | Probe 3     | Probe 4     |
|----------|--------------|-------------|-------------|-------------|-------------|
| A        | Estimated defocus | 2080.00     | 2240.00     | 2400.00     | 2560.00     |
|          | Result       | 2001.14     | 2210.92     | 2386.23     | 2569.09     |
| B        | Estimated defocus | 1760.00     | 1920.00     | 2080.00     | 2240.00     |
|          | Result       | 1690.83     | 1887.37     | 2055.90     | 2236.61     |

| Data set A | Defocus change (nm) | Probes 1-2 | Probes 2-3 | Probes 3-4 |
|------------|---------------------|------------|------------|------------|
| Estimated difference | 160.00     | 160.00     | 160.00     |
| Result | 206.61     | 176.45     | 180.38     |
| Difference from result above | 209.78     | 175.31     | 182.86     |

| Data set B | Defocus change (nm) | Probes 1-2 | Probes 2-3 | Probes 3-4 |
|------------|---------------------|------------|------------|------------|
| Estimated difference | 160.00     | 160.00     | 160.00     |
| Result | 197.68     | 168.95     | 177.42     |
| Difference from result above | 196.54     | 168.53     | 180.71     |

Figure 1 shows an example of the final phase retrieved for probe 1 in data set B. It also shows the final amplitude calculated after the last propagation in the algorithm and how this compares to the recorded amplitude for this probe.

Figure 1 – Real amplitude of probe 1, data set B (left), retrieved amplitude (centre) and phase (right) produced by the GS-GA algorithm.

5. Further work
We hope to improve this algorithm by being able to incorporate a varying convergence angle, perhaps as a free variable for each probe which can be outputted by the algorithm together with the defocus values.

References
[1] Gerchberg R W and Saxton W O 1972, Optik, 35 237-246
[2] Chakrabarti S et al 2008, Data Mining: Know It All, Morgan Kaufmann, 221-292
[3] Atkinson K M, Sweeney F and Rodenburg J M 2008, J.Phys: Conf. Ser., 126 012092