

PHYS3042
Research Topics Physics

Helium nano-bubble matching in defocused TEM
images of plasma exposed Tungsten lamella

Kamil Zietara

A Report submitted for the degree of
MPhys Particle Physics and Cosmology
The Australian National University
Research School of Physics & Materials Physics



April 2023

© Copyright by Kamil Zietara, 2023

All Rights Reserved

Disclaimer

I hereby declare that the work undertaken in this report has been conducted by me alone, except where indicated in the text. I conducted this work between February 2023 and May 2023, during which period I was an undergraduate exchange student at the Australian National University. This report, in whole or any part of it, has not been submitted to this or any other university for a degree.



Kamil Zietara

24 May 2024

Abstract

In the pursuit of a carbon-neutral energy source, fusion power stands out as a viable solution to stabilize rising CO₂ levels. Achieving this requires addressing the challenges faced by plasma-facing materials in fusion reactors, such as tungsten (W). One of the significant issues is the formation of helium (He) nano-bubbles under high-flux, low-energy plasma conditions, which can compromise the material's integrity. This report focuses on optimizing the detection and matching of these He nano-bubbles in W from $\pm 1\mu\text{m}$ defocused transmission electron microscopy (TEM) images in different post-plasma exposure conditions (pre/post-annealing). Two matching techniques, direct and relative angle, were employed to enhance the accuracy of bubble detection. The study revealed precise bubble identification when compared to the post-annealing TEM images. Relative angle matching successfully accounted for distortion during image processing. However, it introduced a bias that removed bubbles systematically nearest to reference points. Automated bubble matching methods identified nearly half as many bubbles as manual counting and provided a more consistent and narrow distribution of bubble sizes. The accurate identification of bubbles is crucial for studying the material behavior under reactor conditions, ultimately aiding in the development of safer and more efficient fusion reactors for mass energy production. Future work will focus on refining these detection methods and comparing them with GISAXS imaging to provide a comprehensive analysis of sub-surface features, further contributing to the advancement of fusion technology.

Contents

List of Figures	iii
0 Background and Motivation	2
0.1 Importance of Fusion	2
0.2 Why Tungsten?	3
0.3 He Exposure and Nano-Bubbles in W	4
0.4 Bubble Analysis	5
0.5 Bubble Detection and Matching	6
1 Methods	8
1.1 Bubble Matching	8
1.2 Direct Bubble Matching	9
1.3 Relative Angle Bubble Matching	11
2 Results and Discussion	14
2.1 Direct Bubble Matching	15
2.2 Relative Angle Bubble Matching	18
3 Conclusions	21
Bibliography	23
A Appendix	25
A.1 Modelling fusion environments	25
B Appendix	26
B.1 Grazing Incidence Small Angle X-ray Scattering	26
C Appendix	27
C.1 Python code snippets	27

List of Figures

1	The graphs show monthly mean carbon dioxide measured at Mauna Loa Observatory, Hawaii. The carbon dioxide data on Mauna Loa constitute the longest record of direct measurements of CO ₂ in the atmosphere. The red lines and symbols represent the monthly mean values, centered on the middle of each month. The black lines and symbols represent the same, after correction for the average seasonal cycle. Source: (NOAA 2024a)	2
2	Electron backscatter diffraction maps showing grain size increase with increasing annealing temperature, with significant recrystallization evident for all samples after annealing at 1400°C. This is seen as the detail vanishes from left to right and the colours become a defined region showing well-defined grain boundaries. Source: (Thompson et al. 2022)	4
1.1	Over-focused, \vec{v}_o , and underfocused, \vec{v}_u , vectors used for the transformation of the underfocused sample onto the overfocused one. Where v_{1o} and v_{2o} are the first and second points that join the \vec{v}_o vector together. The magenta cross is the halfway point on the \vec{v}_o vector that defines the pivot of where the rotation takes place.	10
1.2	Calculation of angles using vectors from under-focused (black) and over-focused (white) reference points.	13
2.1	Source: (Teo et al. 2024). The Figure contains TEM images of a W lamella after pure He plasma exposure at a fluence of 10^{25} m^{-2} - one at RT pre-annealing (left) and RT post-annealing (right) from 998K for 20 minutes. The scale of 40nm allows one to depict the small bubble sizes and the white text indicates the W layer and the protective gold (AU) layer. The two yellow circles indicate the features of interest that were included in the cropped images.	14

2.2	The figure shows the AC cropped TEM images with their corresponding BC blob form. (a) and (b) are the $+1\mu\text{m}$ focus and (c) and (d) are the $-1\mu\text{m}$ focus. From the cropped images, (a) and (c), one can clearly observe similar distinct features near the top and bottom. The condition for generating blobs was $\sigma_{min} = 3$ and $\sigma_{max} = 9$, where σ represents the standard deviation for the Gaussian kernel - smooths the image to reduce noise and enhance the detection of blobs. The overlap was $\pi = 0.8$ (fraction of how much two blobs overlap) - these settings allowed for a focus on predominantly common features and clusters to be detected i.e. ignoring the super small/big points.	15
2.3	Plot (a) shows the original blob positions from the $\pm 1\mu\text{m}$ AC images as well as the transformation vectors and the pivot point (magenta cross). Plot (b) shows the mapping with the parameters used and highlights 3 points of interest whose coordinates can be identified from the real TEM image. They align exceptionally well with the mapped points.	16
2.4	Overlap threshold conditions applied ranging from 2px to 8px on the AC lamella.	16
2.5	The figure contains a comparison of the 8px overlap threshold direct bubble match on top of the $-1\mu\text{m}$ (a) and $+1\mu\text{m}$ (b) AC TEM sample.	17
2.6	Figure containing histograms from the $+1\mu\text{m}$ focus AC (a) and RT (b) lamella comparing the frequency of counts in manual counting from Teo et al. (2024) and automated counting using the direct bubble matching.	18
2.7	Graph (a) shows lenient conditions of an angle and distance threshold $\phi = 2^\circ$ and $\delta = 10\text{px}$ respectively. The thin red diamonds are the reference points (where N is the index of the bubble position in the data file) from which the angles are calculated. The yellow diamonds are features of interest that appear in the real TEM sample. Graph (b) performs 2 restrictions. It decreases the $\phi = 0.5^\circ$ which eliminates only bubbles around the reference points and the $\delta = 5\text{px}$ which removes bubbles around the entire sample but mainly towards the top - the bubbles that are close but not overlapping. Graphs (c) and (d) are exactly the same other than the reference points are chosen to be on opposite ends of the image.	19

Chapter 0

Background and Motivation

0.1 Importance of Fusion

Carbon dioxide (CO₂) emissions are up +228.48% globally since 1965 with the highest 5-year increase of +32.84% between 2005 and 2010 graphed in Energy Institute (2023). Under the business-as-usual scenarios of the International Panel on Climate Change (IPCC), these levels are forecast to rise as high as 600 ppm by 2050. Even under the most optimistic scenarios (representative concentration pathway 2.6 in IPCC (2019)), the atmospheric CO₂ concentration is not expected to stabilise until 2100. The atmospheric CO₂ levels were 425ppm in March 2024 with no real sign of decreasing. This can be seen in Figure 1 - a graph of the trend in atmospheric CO₂ from the Mauna Loa Observatory (NOAA 2024b). This rising increase in CO₂ emission will potentially increase the average global temperatures by 1.5°C pre-industrial level where international agreements, such as the 2015 Paris Agreement, aim to limit global warming to well below 2°C pre-industrial levels. An increase of 1.5°C will destabilize climate patterns significantly (IPCC 2018).

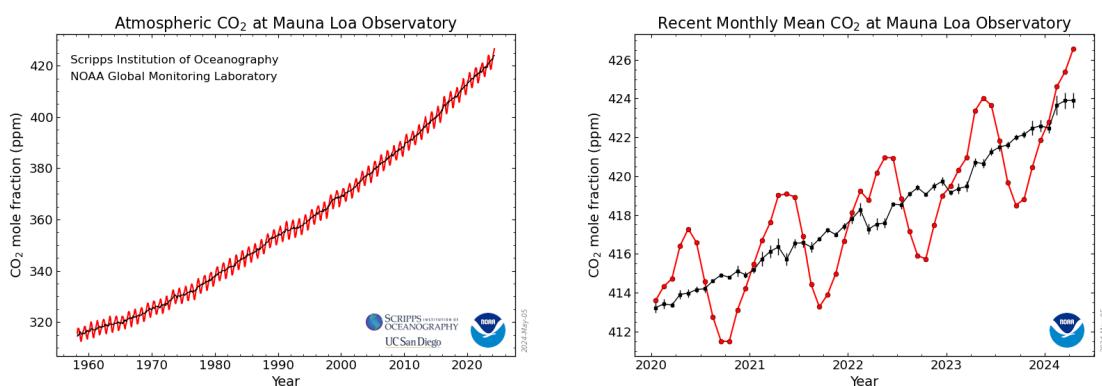


Figure 1: The graphs show monthly mean carbon dioxide measured at Mauna Loa Observatory, Hawaii. The carbon dioxide data on Mauna Loa constitute the longest record of direct measurements of CO₂ in the atmosphere. The red lines and symbols represent the monthly mean values, centered on the middle of each month. The black lines and symbols represent the same, after correction for the average seasonal cycle. Source: (NOAA 2024a)

The proposed idea of fusion nuclear power through the International Tokamak Experimental Reactor (ITER) is the answer to a carbon-neutral energy source (Giuliani et al. 2023).

Fusion technology is based on the principle of combining light atomic nuclei to form a heavier nucleus, a process that releases a tremendous amount of energy. The primary fuel for fusion reactions is isotopes of hydrogen, such as deuterium and tritium. These isotopes are abundant and can be extracted from water and lithium, making fusion a potentially limitless energy source.

In a fusion reactor, such as ITER, the fuel is heated to extremely high temperatures (about 150 million degrees Celsius) to form a hot plasma. This plasma is a state of matter where the electrons are separated from the nuclei. The plasma must be confined long enough and at a sufficient pressure to allow the nuclei to overcome their electrostatic repulsion and fuse.

The tokamak is a doughnut-shaped device that uses powerful magnetic fields to confine the plasma. These magnetic fields are generated by superconducting coils surrounding the tokamak. The goal is to achieve conditions where the fusion reactions enter the self-sustaining burning plasma state, known as ignition. Once ignition is reached, the energy produced by the fusion reactions can be used to generate electricity (Morse 2018).

ITER aims to reach the Q=10 goal, where the reactor would produce 10 times the amount of thermal energy from fusion than required for external heating, achieving a significant net energy gain, which is a crucial step toward practical energy production from fusion. By the end of 2019, ITER had completed about 67% of the overall work required to achieve First Plasma. First Plasma refers to the first time the reactor will successfully create a plasma inside the tokamak—the heart of the reactor where fusion reactions take place. This milestone is critical for testing and validation of the reactor’s design and operational integrity and not intended for sustained energy production. The primary goal during the First Plasma phase is to achieve a plasma state and maintain it for a short duration to ensure that all systems are functioning as designed (Bigot 2021).

0.2 Why Tungsten?

In 2013 a crucial design decision was made by ITER to switch to a fully Tungsten (W) armored divertor from a Carbon fibre composite (CFC) and W mixture, with the main reasons being lifetime and safety. Carbon-based materials in a reactor environment lead to an increase of carbon impurities in the plasma reducing its overall fusion efficiency caused by radiative cooling. The high sputtering yield at high temperatures of carbon causes erosion of the divertor material wall, leading to a build-up of radioactive waste exposed to a neutron-rich environment. W with its high melting point (3695 K) and excellent thermal conductivity becomes a more robust choice for the high heat-flux divertor environment and more sustainable for long-term operations as outlined by Griffith (2013). Initially, the plasma-facing components of W were not well understood and there was a fear that W would be

too brittle to withstand the thermal shocks occurring during edge localized modes (ELM's) causing cracking and material failure (Wirtz et al. 2013). The switch to a full W divertor happened as an iterative process of technological advancement, with more research the confidence in W increased. The initial use of CFC's was due to the sheer fact that they were very well understood and could be more easily engineered. However, using purely W also has drawbacks (Pitts & Bonnin 2019). One problem is that W has a high atomic number of 74, which means its atoms can be partially ionised even at very high temperatures. These ionized atoms emit radiation in the form of photons. When electrons in the ionized W recombine with ions or transition between different energy levels, energy is released as electromagnetic radiation. This process removes energy from the plasma, leading to cooling and potential quenching, thereby reducing the plasma's operational efficiency. Another issue is recrystallisation in W - recrystallisation leads to a potential loss of hardness, as well as mechanical properties, due to the formation of new large defect-free grains as the material releases stress going through repeated thermal shocks of a fusion reactor. This was demonstrated in an experiment by Thompson et al. (2022) where samples exposed to helium (He) plasma and then annealed showed significant recrystallization after annealing at 1400°C shown in Figure 2. Another major effect observed in that experiment was the formation of He nano-bubbles under the surface of W, discussed in more detail in the next section.

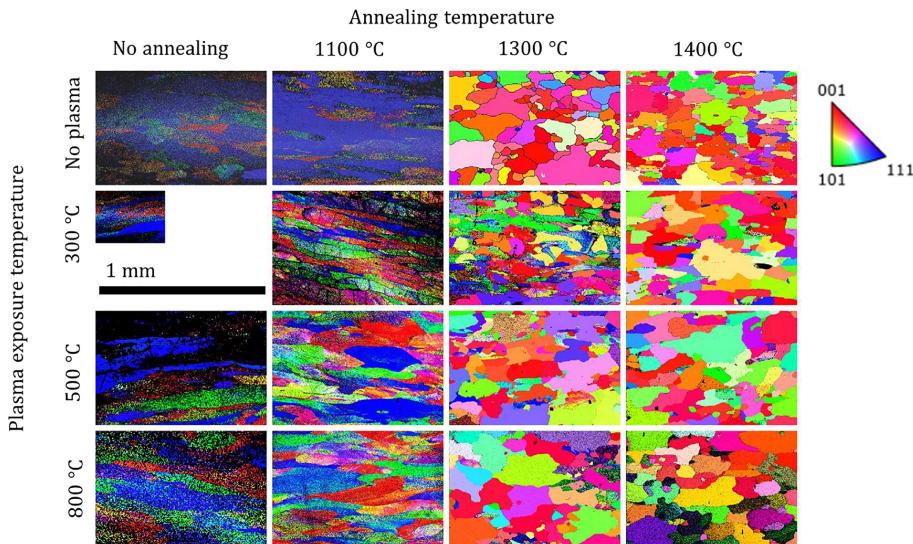


Figure 2: Electron backscatter diffraction maps showing grain size increase with increasing annealing temperature, with significant recrystallization evident for all samples after annealing at 1400°C. This is seen as the detail vanishes from left to right and the colours become a defined region showing well-defined grain boundaries. Source: (Thompson et al. 2022)

0.3 He Exposure and Nano-Bubbles in W

Nano-bubbles of He formed when W was exposed to high-flux low energy He plasma (A scenario that reflects the plasma conditions in the divertor). The bubble forma-

tion was likely due to the diffusion and clustering (via its self-trapping property) of He in the tungsten matrix (Henriksson et al. 2005). The nano-bubble formation wasn't observed for deuterium or tritium plasmas as their atoms are much smaller in size than He atoms and likely bind into existing defects of the tungsten matrix. Recrystallisation and nano-bubble formation happen to interact with each other. For low-temperature He exposure (< 700 K), nano-bubbles approximately range from 1 – 2 nm in diameter meaning that they would engulf multiple tungsten atoms. This caused retardation in the recrystallisation as grain boundaries would more preferably align through the bubbles rather than around them. This produced a more energetically favourable configuration by decreasing the surface area of interfaces such as grain boundaries (Raabe 2014). For a similar reason, there was a high rate of coalescence in He nano-bubbles as being together reduced the total volume they occupied putting them in a preferable configuration. The analysis of He nano-bubble formation at different conditions in conjunction with recrystallisation is crucial for understanding the properties of W in the ITER reactor for safe and effective operation (Song et al. 2019).

0.4 Bubble Analysis

At this stage, He nano-bubble formation is inevitable in the situation that the plasma-facing components of W are exposed to. Therefore, it is imperative that we understand the nucleation, how they evolve throughout a range of conditions, and what this means for the integrity of the material. Modelling can be a great tool for this and is outlined in Appendix A.

The most suited experimental techniques for analysing these sub-surface structures are Transmission Electron Microscopy (TEM) (Thompson 2018) and Grazing Incidence Small Angle X-ray Scattering (GISAXS) described in Appendix B.

0.4.1 Transmission Electron Microscopy

TEM is a powerful analytical tool used for visualizing structures at the atomic to nanoscale. This technique operates by transmitting a beam of high-energy electrons (between 100-300 keV for heavy metals) through an ultra-thin sample that is focused by a series of electromagnetic lenses. The emphasis on the ultra-thin size of the sample is to allow the electrons to pass through entirely to reduce scattering inside the sample that would otherwise reduce the image resolution. Apertures are used to control beam intensity and reduce imaging defects such as spherical aberration from imperfect lenses. Once passed through the sample the electrons pass through a forward scattering projector lens before hitting the viewing screen forming diffraction patterns used for analysis (Carter & Williams 2016).

TEM enables the direct visualization of nanobubbles, offering insights into their distribution, size, and interactions with the surrounding matrix. However, TEM is constrained by the limited volume it can image, which imposes certain limitations on the conclusions drawn from its images. Previous studies have shown that grain

boundaries and orientation significantly impact bubble growth and nanostructure formation. To address this, numerous specimens from different regions of the same sample are required. This necessity complicates the process, making it difficult to account for discrepancies and requiring a statistical approach to overcome the sampling limitations (Ohno et al. 2013).

0.5 Bubble Detection and Matching

The motivation behind this report is to explore the feasibility of automated bubble counting over manual counting in the different samples of W. Automated counting involves a systematic selection criteria increasing precision and reducing operation time significantly. Pre-annealed samples in the thinnest part have up to a few hundred features which could be classified as He nano-bubbles. Bubbles are not simple to differentiate from other features like surface pits and nano-fuzz. Fringes from Airy discs can also cover nearby features. Manual counting involves identifying numerous grey-scale features on a complex image, which inherently introduces random errors, such as omitting extremely small bubbles with low contrast and resolution discussed in Teo et al. (2024). Increasing brightness and contrast too much can exacerbate this issue by causing faint features to merge into surrounding ones. Both manual counting and automated blob generation share the drawback of systematically overestimating/underestimating bubble size compared to their true size. The overestimation arises from the assumption that bubbles are circular, which is generally not the case, and the underestimation can come from a poorly defined pixel intensity condition for determining bubble edges.

This report focuses on the optimisation of bubble matching methods and interpreting TEM samples of W exposed to 1013K pure He plasma at a fluence of 10^{25} m^{-2} . The samples were measured via TEM and then annealed in-situ up to 998 K, before being returned to room temperature for a final measurement, as described in Teo et al. (2024). The W lamella was measured with a $\pm 1\mu\text{m}$ ('Over/under focused' respectively) defocus.

The basis of the bubble matching algorithms work on a 'Bubble Convolution' (BC) Python code, created by dingk (2024), that would identify all objects on the samples that could represent physical features, like bubbles, as circular blobs using the scikit-image package. Two approaches — direct matching and relative angle matching — were devised to interpret these samples, identify real physical features, eliminate false positives and double counting, and enhance process efficiency to reduce computational demand.

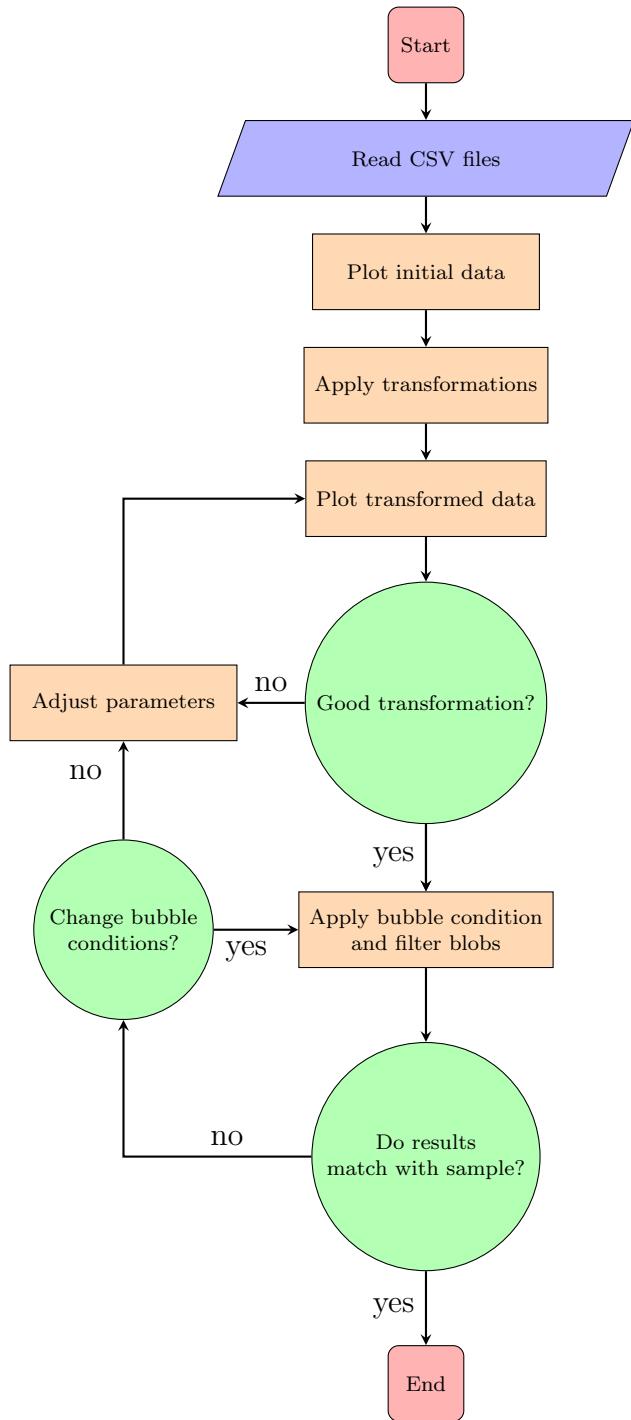
A systematic approach in conjunction with human verification is necessary to be able to confidently analyse these complex samples. Before using the BC code the TEM samples were modified, (using ImageJ) by increasing their contrast and brightness, to make it easier for the BC code to identify bright spots. They were also cropped to approximately the same size focusing on an area that appeared to have lots of bubbles. Due to the different focus' on the W lamella, it wasn't obvious where the

bubbles should begin. The region of interest defined captured the same distinct features. The cropping may cause one sample to have more visible bubbles than the other, however, this should be mitigated during the filtering part of the matching algorithm. The BC code uses a package called scikit-image which allows for greyscale image processing using Laplace of Gaussian (LoG) method to detect edges and features (scikit image 2024). The combination of Gaussian smoothing and the Laplacian operator result in enhancing edges and features while suppressing noise. By thresholding the resulting image (separating regions of interest from the background based on a pixel intensity) edges can be detected and extracted efficiently (LoG 2024). Based on a threshold pixel intensity the BC code will define circular blobs of different sizes around bright regions in the greyscale image. The blobs are constrained by a standard deviation sigma which affects the central position and radius of the bright feature. It needs to be noted that the BC code doesn't differentiate the shapes of these features, which can be spheroidal or clusters and draws them all as circles. There is also no indication of the depth of the bubbles (relative to the direction of the electron beam), as the resulting image is a projection of the sample's internal structure onto a two-dimensional plane. The view of the sample is a cross-section of the thinnest part after He plasma exposure. The code when processing the over-focused sample first inverts it so that features go from black to white spots - to allow the code to search for light regions in both images. During the first run of the sample, the BC code will also crop a small line width around all 4 edges of the samples to eliminate any bubbles that were cut in half or poorly captured during the initial cropping leaving only edge blobs that could be shown as complete circles. Once the sample is analyzed for bright features it's inverted back and the BC code produces a PNG image of the TEM sample, a PNG image of the circular blobs, and a text file of positions and radii. An identical process is followed for the under-focused sample without the inversion. Using the two methods mentioned (direct and relative angle matching) the algorithm attempts to match blobs from both samples to identify real bubbles. The process then involves observing these filtered samples to determine if they correspond to actual physical TEM images. Various samples will be tested to evaluate how bubble nucleation and clustering, induced by annealing, impact the detection code's performance and to verify if the matching code effectively filters out false positives.

Methods

1.1 Bubble Matching

This report explores two methods for bubble matching, a direct one and a relative angle one. Both of these methods follow the same underlying structure. This structure can be seen in Flowchart 1.

**Flowchart 1:**

This flow chart illustrates a much-simplified process of how the algorithm works. It begins by importing the relevant data (x,y coordinates, and radii) from the under-focused and over-focused samples and plotting them for visual aid. Then through a series of equations, which define the translation, scaling, and rotation, it maps one data set onto the other. After assessing the quality of the transformation against a stringent set of criteria, it determines whether to adjust the transformation parameters or proceed to apply matching conditions that distinguish between blobs and bubbles. Finally, by plotting the filtered bubbles onto the real TEM image one can decide if the bubbles match up with physical features of the sample. If not, then it returns to either change the matching conditions or the transformation parameters. This process would then be applied to W lamella that were measured at different conditions (i.e. Pre-annealing and post-annealing) to see how robust it is at identifying He nanobubbles in samples of W.

1.2 Direct Bubble Matching

The direct approach is the simplest process we can perform for this investigation. It uses two vectors to carry out a transformation. The transformation includes translation, scaling, and rotation. After the mapping, the algorithm attempts to filter out any blobs that aren't overlapping with other blobs from the opposite sample within a certain radius of pixels through Euclidean distance calculations

using 'cdist' from 'scipy.spatial.distance'.

$$\text{Euclidean Distance} = \sqrt{\sum_{i=1}^N (x_{-1i} - x_{+1i})^2} \quad (1.1)$$

Where N is the number of blobs, x_{-1i} and x_{+1i} are the i-th blobs of the under and over-focused points respectively.

The transformation only depends on two vectors. They must represent a line connecting the same physical features and ideally span a large length of the sample. For example, two distinct bright features are chosen in the overfocused sample - one at the top and one at the bottom of the sample. These features, although may not be bubbles, clearly appear as very dark spots in the opposite underfocused sample. This would be an ideal scenario of the vector criteria for our transformation allowing to have a precise mapping of points throughout the sample. The accuracy of the transformation depends on how closely these vectors depict the same real features of the differently focused images. Figure 1.1 illustrates how the transformation operates.

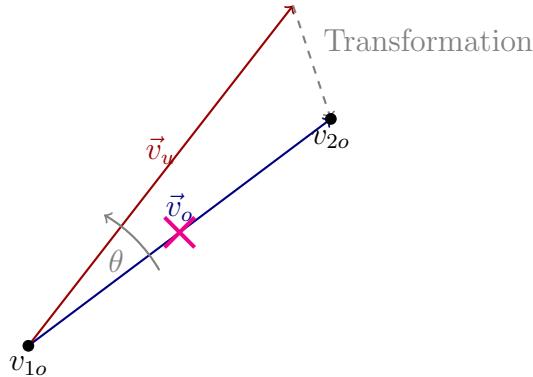


Figure 1.1: Over-focused, \vec{v}_o , and underfocused, \vec{v}_u , vectors used for the transformation of the underfocused sample onto the overfocused one. Where v_{1o} and v_{2o} are the first and second points that join the \vec{v}_o vector together. The magenta cross is the halfway point on the \vec{v}_o vector that defines the pivot of where the rotation takes place.

The transformation is performed by a translation, rotation and scaling.

Translation: Determined by the difference in the position of the vectors

$$\delta s = v_{1o} - v_{1u} \quad (1.2)$$

Where v_{1o} and v_{1u} are the first over-focused and under-focused points respectively.

Scaling: Determined by the ratio of the distances between points v_{1o} to v_{2o} and v_{1u} to v_{2u} . This scales the under-focused data to match the over-focused data. Computed as

$$\delta l = \frac{\|v_{2o} - v_{1o}\|}{\|v_{2u} - v_{1u}\|} \quad (1.3)$$

Where v_{1o} and v_{2o} are the first and second points that join the \vec{v}_o vector together.

Rotation: Determined by the angle between the two vectors calculated using the tangent of the angle formed by them, which can be computed using the arctan2 function:

$$\delta\theta = \arctan 2(y_{v_o}, x_{v_o}) - \arctan 2(y_{v_u}, x_{v_u}) \quad (1.4)$$

Where y_{v_o} and x_{v_o} are the y and x coordinates of the over-focused vector.

The arctan2 function (unlike the arctan function which returns angles only between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$) returns angles in the range $-\pi$ to π . This allows it to cover all four quadrants, making it more suitable for calculating angles in various geometric contexts like this one.

After the transformation is complete the direct matching method filters out any non-overlapping pairs of blobs using equation 1.1 to create a mask for any points that are less than the overlap threshold. It does this for both sets of the over-focused and under-focused blobs. To filter out false positives or many-to-one mapping it uses 'np.unique' to only leave unique elements in the array. Plotting these newly matched bubble positions and radii one can overlay the plot onto the real TEM image for comparison and the robustness of the filtering.

1.3 Relative Angle Bubble Matching

Relative angle (or angular) bubble matching is more complex than direct bubble matching. While the direct method relies solely on global spatial transformations, angular matching adds a second layer of constraints by identifying corresponding bubbles based on the relative angles formed by their positions with respect to an initial known point. This dual-layer approach supplements distance filtering with angular relationships, providing a more robust and refined matching process - especially in scenarios where global transformations and distance calculations alone may not suffice.

For this approach, the transformation process for mapping the blobs is adjusted. Similarly, the algorithm uses translation, scaling, and rotation. All 3 functions are defined with respect to a 'center' which is the mean position of all x and y coordinates of the over-focused sample. Once the transformation is complete for an initial guess an optimisation is applied to find a better translation, scaling factor, and angle of rotation.

Translation: Determined by the difference of each x and y coordinate for every point and the center of the overfocused sample.

$$\begin{aligned} x' &= x - x_c \\ y' &= y - y_c \end{aligned} \quad (1.5)$$

Where x' and y' are the shifted points and x_c and y_c are the coordinates of the center.

Scaling: Each point is multiplied by an initial guess scaling factor, δl .

$$\begin{aligned} x' &= x \times \delta l \\ y' &= y \times \delta l \end{aligned} \quad (1.6)$$

Rotation: Determined by a rotation matrix in Cartesian coordinates with an initial guess of the angle θ .

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (1.7)$$

Where x' and y' are the coordinates for the rotated points.

Once these transformations are applied (in the same order as stated above), using 'scipy optimize differential evolution' all the transformation parameters are optimised to minimise the distance between all individual points essentially finding a global minimum for the set of parameters provided. Differential evolution is a stochastic population-based method that is useful for global optimization problems. At each pass through the algorithm mutates each candidate solution by mixing with other candidate solutions to create a trial candidate. A crossover check is performed between the original and mutated solution. If the mutated solution outperforms the original it replaces the solution. The code used for this optimisation is described in Appendix C.1.1

The strategy for creating trial candidates used was "best1bin". The "best1bin" strategy in differential evolution is a variant that selects the best individual from the current population, along with another randomly chosen individual, to generate a trial candidate. This strategy emphasizes exploitation by favoring the best-performing individuals while maintaining some level of exploration through random selection. This stochasticity helps explore the search space more effectively and avoid getting stuck in local optima (SciPy 2024).

After getting an adequate transformation the bubble criteria are applied by the angular matching algorithm. 4 parameters are set - over-focused blob index, under-focused blob index, nearest neighbour radius and angle threshold. Each dataset is initialised with a 'Tested' = -1 and 'Bubbles' = 0 columns. The tested column will flip to a 1 when the blobs have been iterated over and the bubble column will either take a 1 or a -1 depending on whether or not it satisfies the bubble condition. The angular matching process involves the following sections:

Angle: The angle is calculated simply using the dot product between 2 vectors. Vector 1 is from the reference point to the chosen over-focused blob and vector 2 is from the reference point to the chosen under-focused blob. This is done twice for each pair of points from the under-focused and over-focused reference points to obtain 2 angles.

$$\theta = \arccos \left(\frac{\vec{v}_1 \cdot \vec{v}_2}{|\vec{v}_1| \cdot |\vec{v}_2|} \right). \quad (1.8)$$

where θ is the angle between the vectors \vec{v}_1 and \vec{v}_2 .

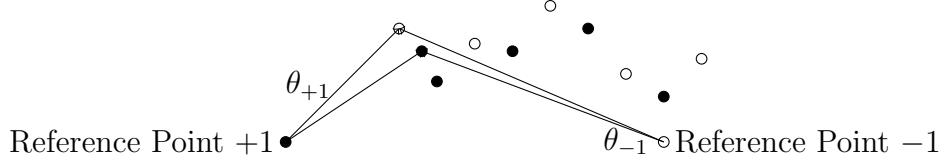


Figure 1.2: Calculation of angles using vectors from under-focused (black) and over-focused (white) reference points.

Bubble condition: This condition is met when the absolute difference in the angle, formed by the same blobs from different reference points is, less than the angle threshold.

$$|\theta_{-1} - \theta_{+1}| < \phi. \quad (1.9)$$

Where θ_{-1} and θ_{+1} are the under and overfocused angles from their respective reference point and ϕ is the angle threshold.

Nearest Neighbours: The algorithm uses `NearestNeighbors` class in scikit-learn which is a powerful tool for finding the nearest neighbors of a point in a dataset efficiently. It constructs a spatial index structure, such as a KD Tree or Ball Tree, to enable fast querying of nearest neighbors - index construction becomes a one-time computational cost. In this specific case 'auto' is called which chooses the most appropriate for the given dataset. The algorithm uses the 'radius neighbors' method that returns an array of the indices of all points within a specified radius of the query point.

The Loop: The data points are ordered from smallest to largest y coordinate so the bubbles are tested from the bottom up with only the nearest neighbours being evaluated at one time. Once all of the nearest neighbouring blobs have had their tested column initialised to a '1' the next blob to be tested is the one with the smallest y value that has a '-1' in the tested column. The blobs that were checked by the previous neighbours are not re-tested. This loops until there are no more blobs in the over-focused sample left to initialise in the tested column.

The angular bubble matching offers a geometric approach to bubble correspondence, leveraging relative angles to establish associations between bubbles in different samples. While its effectiveness relies on accurate angle calculations and suitable parameter selection, it provides a robust method for bubble matching. One additional motivation to point out is that the direct transformation method requires that the images are not distorted on a large scale and that one can accurately identify equivalent points. The angular method relaxes this requirement, as you are using smaller vectors so these large baseline issues don't arise. Once the loop is finished, just like the direct method, the matched bubbles are plotted and compared against the real TEM samples. A snippet of this loop can be seen in Appendix C.1.2.

Results and Discussion

In this investigation two conditions of a W lamella were analysed - one before annealing at room temperature (RT) and one after annealing from 998K to RT (taking 20 minutes) labeled 'After Cool' (AC). This was to test the robustness of the bubble matching in 2 very different conditions. Each lamella had a defocus version of $\pm 1\mu\text{m}$ (Over and under focused)- this resulted in 4 images. The $+1\mu\text{m}$ focused images showed features as dark spots whereas the $-1\mu\text{m}$ focus had them as light spots.

Figure 2.1 below shows the TEM images of the RT and AC lamella with $-1\mu\text{m}$ focus.

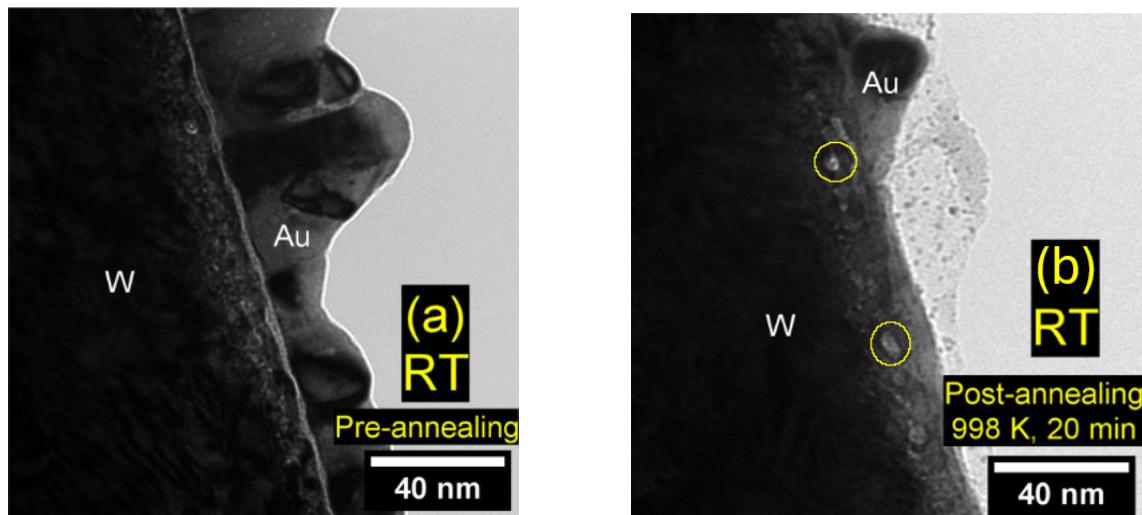


Figure 2.1: Source: (Teo et al. 2024). The Figure contains TEM images of a W lamella after pure He plasma exposure at a fluence of 10^{25} m^{-2} - one at RT pre-annealing (left) and RT post-annealing (right) from 998K for 20 minutes. The scale of 40nm allows one to depict the small bubble sizes and the white text indicates the W layer and the protective gold (AU) layer. The two yellow circles indicate the features of interest that were included in the cropped images.

Then the images were cropped and edited by increasing brightness and contrast in both the $\pm 1\mu\text{m}$ versions. The BC code then generated an image of blobs for

all features that satisfied the conditions. An example of this is shown in Figure 2.2.

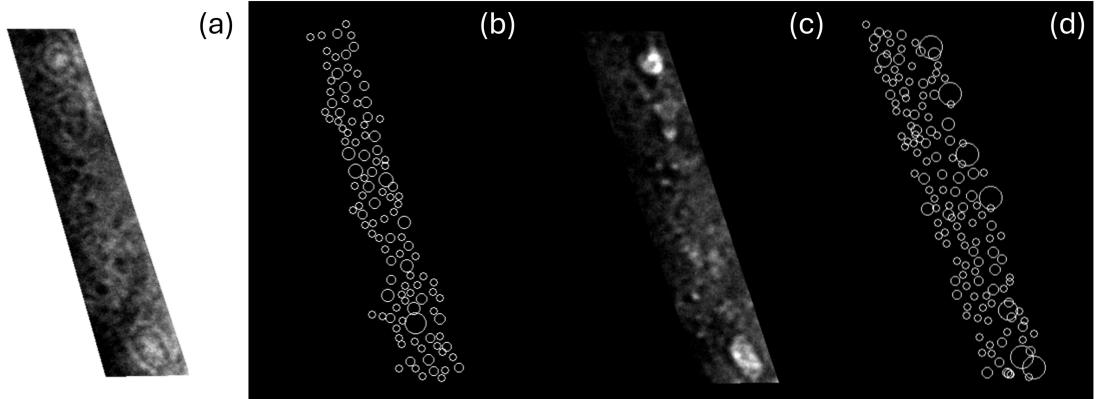


Figure 2.2: The figure shows the AC cropped TEM images with their corresponding BC blob form. (a) and (b) are the $+1\mu\text{m}$ focus and (c) and (d) are the $-1\mu\text{m}$ focus. From the cropped images, (a) and (c), one can clearly observe similar distinct features near the top and bottom. The condition for generating blobs was $\sigma_{min} = 3$ and $\sigma_{max} = 9$, where σ represents the standard deviation for the Gaussian kernel - smooths the image to reduce noise and enhance the detection of blobs. The overlap was $\pi = 0.8$ (fraction of how much two blobs overlap) - these settings allowed for a focus on predominantly common features and clusters to be detected i.e. ignoring the super small/big points.

The cropping of the $\pm 1\mu\text{m}$ AC images in Figure 2.2 are reasonably similar with many comparable features, however, the BC blob forms do not look much alike with differences in bubble size and distribution. Looking at the coordinates of distinct features there are false positives from the BC circles in the $-1\mu\text{m}$ lamella (right) - around the bottom right where there are large overlapping blobs. The bubble matching aimed to find the similarities that the BC blob forms did have and compare them to the cropped TEM images.

2.1 Direct Bubble Matching

Direct bubble matching, as described in Section 1.2, maps one vector onto the other - where the vectors represent lines in each image that connect the same physical features. The original points with their vectors and the mapping are shown in Figure 2.3.

The mapping resulted in a Translation of [61,-1] pixels in x and y, a scaling of 0.9990, and a rotation of 0.776°. Identifying coordinates of points of interest from the real samples (using ImageJ) confirms that the mapping is accurate. After the mapping, the direct bubble matching conditions described in Section 1.2 were applied.

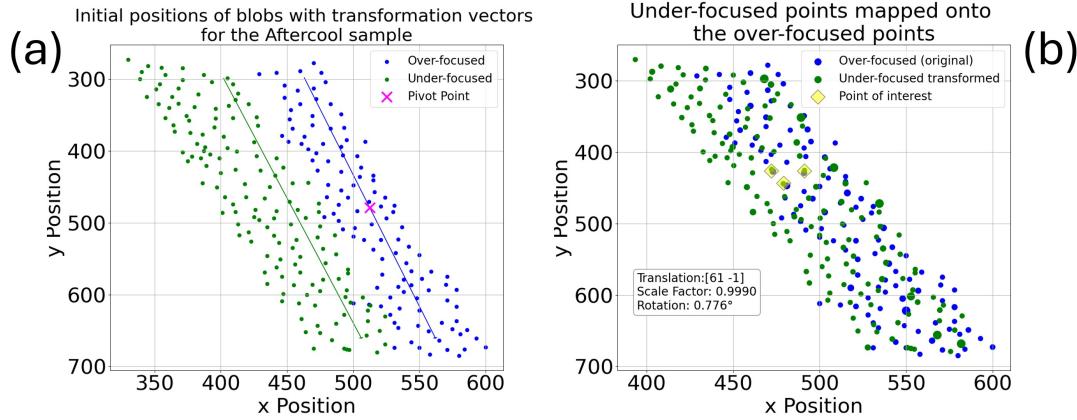


Figure 2.3: Plot (a) shows the original blob positions from the $\pm 1\mu\text{m}$ AC images as well as the transformation vectors and the pivot point (magenta cross). Plot (b) shows the mapping with the parameters used and highlights 3 points of interest whose coordinates can be identified from the real TEM image. They align exceptionally well with the mapped points.

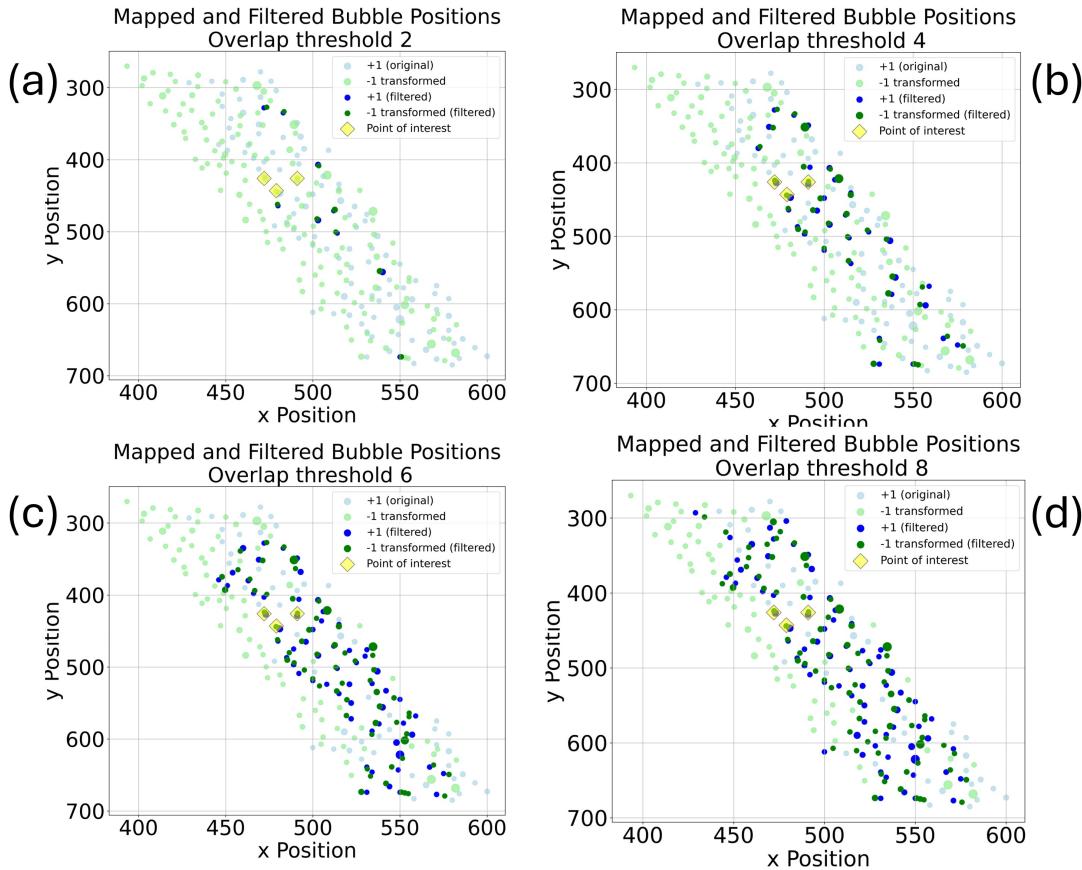


Figure 2.4: Overlap threshold conditions applied ranging from 2px to 8px on the AC lamella.

Figure 2.4 shows 4 graphs with different overlap thresholds, which is the length of pixels that the blobs can be away from each other, ranging from 8px to 2px. This

range spans from the smallest, 2px, to more than the largest bubbles, 6px. The pixel-to-nanometers ratio is 6.6673px/nm. Anything over the overlap of 8px would result in blobs that are too far away from each other to represent the same bubble. An overlap threshold of 4px mostly leaves blobs that are mapped one to one and are physically overlapping. Whereas an overlap threshold of 2px results in too many false negative features, including the highlighted ones identified previously. To test these bubble-matched samples they were overlayed onto the TEM cropped images.

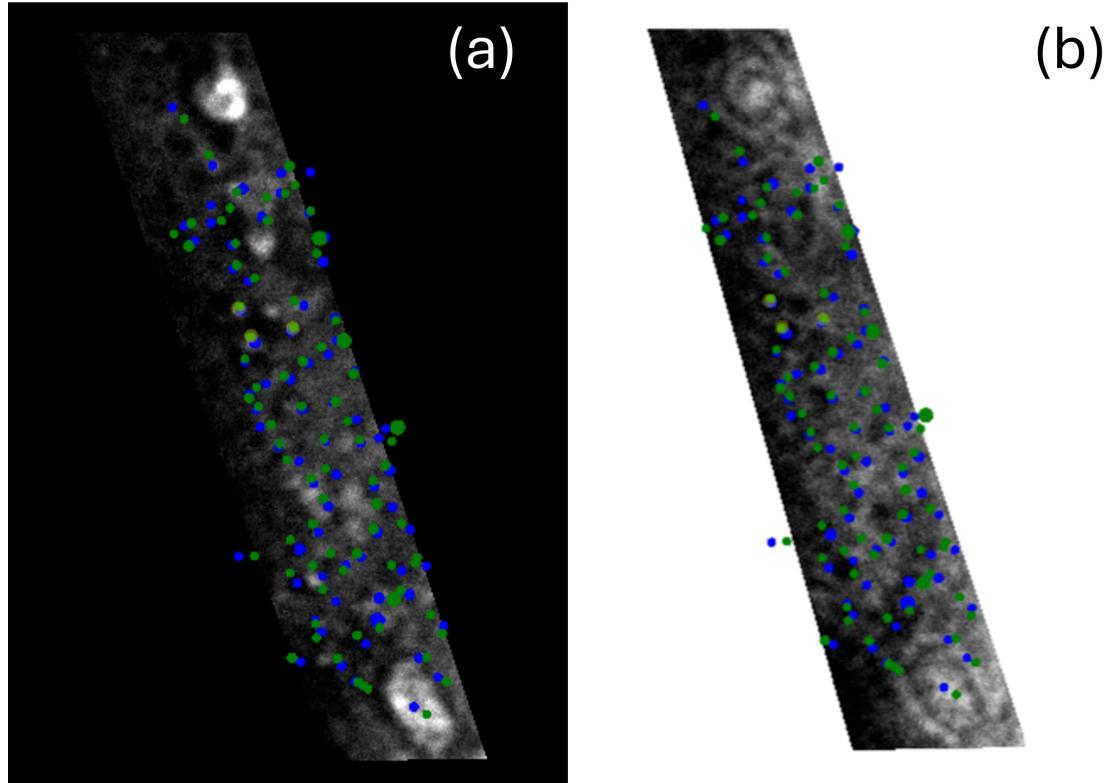


Figure 2.5: The figure contains a comparison of the 8px overlap threshold direct bubble match on top of the $-1\mu\text{m}$ (a) and $+1\mu\text{m}$ (b) AC TEM sample.

The 8px overlap threshold contains false positives mainly around the edges of the sample which can be eliminated by carefully lowering the threshold. However, the accuracy of these bubble positions heavily depends on the actual BC blob generation for those blobs to depict physical features. Nonetheless, there are multiple points that have matched bubbles and appear to lay on a feature of interest in the TEM image.

Another area of interest of this report was to show the effectiveness of automated counting over manual counting as described in Section 0.5. Manual counting can underestimate bubble counts, which may arise from omitting extremely small bubbles exhibiting low contrast and resolution. The range of radii for manual counting is also greater than in automated counting as the restrictions are stricter for detection in automated counting.

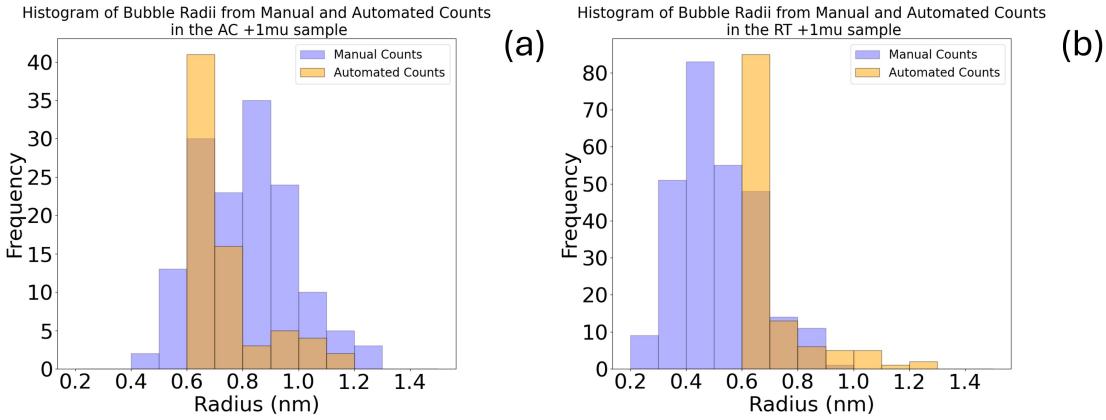


Figure 2.6: Figure containing histograms from the $+1\mu\text{m}$ focus AC (a) and RT (b) lamella comparing the frequency of counts in manual counting from Teo et al. (2024) and automated counting using the direct bubble matching.

These trends are noticed in the histograms in Figure 2.6. The total number of bubbles in automated counts compared to manual counts in Teo et al. (2024) was 73 vs 146 for the AC and 127 vs 274 for the RT sample (with a 6px overlap threshold for automated bubbles). The number of automated counts was almost half as much as the manual ones. This could be due to the lack of matches of bubbles near the edges and the potential difference in cropping and/or brightness of the images. The automated counts also show a majority of a single bin of bubble radius of 0.6nm - 0.7nm in the $+1\mu\text{m}$ image showing the strict criteria of automated counts and the potential overestimation of 'true size' in the manual counts. However, automated counting introduced a minimum size cut-off defined by the blob detection criteria of $\sigma_{min} = 3$, corresponding to the detection of features no smaller than 0.6 nm in radius. This cut-off seemed to limit the results for the Room Temperature sample, as smaller features should have been detected (as seen in the lamella in Figure 2.1(a)) to reveal the over-estimation of bubble size in manual counts, as observed in the Aftercool histogram 2.6 (a) and described in Teo et al. (2024).

The RT lamella, as expected, had much more and smaller bubbles than the AC lamella showing the effect of He bubble coalescence and smoothing of grain boundaries during annealing of the Aftercool lamella.

2.2 Relative Angle Bubble Matching

The angular bubble matching technique builds on the direct approach by adding more layers of complexity and thorough operations to increase accuracy. Firstly, The vector transformation was [61, -1]px, 0.9990 and 0.776° , and the optimised angular transformation was [61.5, -1.3]px, 1.0013 and 0.044° for the translation, scaling, and rotation respectively. The transformation optimisation described in Section 1.3 appears to work exceptionally well as it is almost identical to the transformation described by the vectors in Figure 2.3 (a), which were carefully chosen to represent real features. It is also clear to see that more points overlap in Figure 2.7 (a) from

the optimised transformation than in Figure 2.3 (b).

The angular approach introduces more degrees of freedom with 3 parameters that, when changed, affect the outcomes - namely the radius of nearest neighbours, angle threshold, and distance threshold. The dependence of reference points for angle calculation also has a noticeable effect. If the reference points represent the same physical feature in both the $\pm 1\mu\text{m}$ focused images then the points closest to them are filtered out as the angle they form is much larger than if the same points were further away. This is illustrated in Figure 2.7 as well as the effect of decreasing the distance threshold. The radius is fixed in this situation.

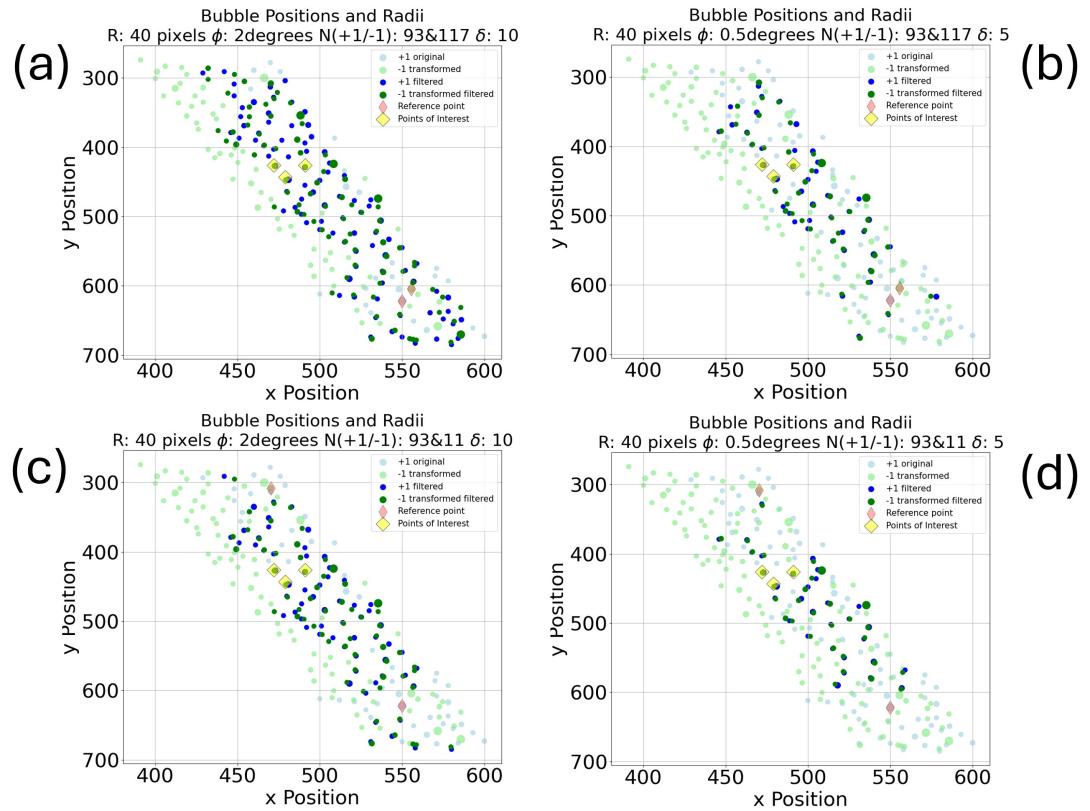


Figure 2.7: Graph (a) shows lenient conditions of an angle and distance threshold $\phi = 2^\circ$ and $\delta = 10\text{px}$ respectively. The thin red diamonds are the reference points (where N is the index of the bubble position in the data file) from which the angles are calculated. The yellow diamonds are features of interest that appear in the real TEM sample. Graph (b) performs 2 restrictions. It decreases the $\phi = 0.5^\circ$ which eliminates only bubbles around the reference points and the $\delta = 5\text{px}$ which removes bubbles around the entire sample but mainly towards the top - the bubbles that are close but not overlapping. Graphs (c) and (d) are exactly the same other than the reference points are chosen to be on opposite ends of the image.

From Figure 2.7 we can see the biased effect caused by the reference points depicting the same feature due to the lack of matches close to the reference points. Next, the investigation was repeated with 2 different reference points - an under-focused point near the top and an over-focused point near the bottom of the lamella whilst keeping

the same radius (40px) and the same change in angle and distance thresholds from 2° to 0.5° and from 10px to 5px respectively.

Repeating this investigation it is noticed that the bias has not gone away entirely. Although some points near the reference points are accepted there is still removal of bubbles near the reference points. For future investigation, more reference points could be added, for example in 4 quadrants of the sample, with a condition that the angle calculation is performed on blobs only by reference points that are furthest away from them.

Increasing the radius of nearest neighbours and keeping everything else fixed yielded no differences other than a longer computational time. The angle threshold acts as a second-layer restriction after the distance threshold is passed. From the investigations above it can be observed that the process is sensitive to both angle and distance threshold changes. They should be optimised for each case individually to allow the bubble matching to represent the real sample accurately. However, the focus of this report is on the effectiveness of bubble matching. The direct approach doesn't account for large deviations in sample orientation from the imaging process. The angular approach accounts for this but introduces biases around the reference points that remove bubbles nearby and themselves. The accuracy of matching increases further away from the reference points the blobs are

Conclusions

This report provides valuable insights into the optimization of helium nano-bubble matching in differently focused TEM images of tungsten lamella exposed to pure helium plasma. It demonstrates a simple and effective direct approach to bubble matching, yielding approximate results that allow for clear bubble identification for further analysis. The approach was refined to account for image distortions using a relative angle method. This method supplements distance filtering with angular relationships, considering sample orientation to provide a more accurate process.

However, the relative angle approach introduced a reference point bias by systematically removing blobs closest to the reference points. Despite this limitation, the report qualitatively demonstrated good precision in bubble matching when compared to TEM images. Quantitatively, it compared the number of bubbles matched in automated counting versus manual counting, as conducted in Teo et al. (2024). Manual counting identified nearly twice as many bubbles as automated counting, with counts of 127 versus 73, respectively, in the $+1\mu\text{m}$ focused Aftercool image. This trend was consistent in the Room Temperature image, suggesting that the automated counting misses some of the bubbles that have multiple diffraction fringes or complex shapes because of the single brightness selection criteria. It would be expected that the manual process is accurate as the human eye/brain have phenomenal object recognition capabilities, despite some of the other limitations. Automated counting used a minimum size cut-off defined by the blob detection criteria of $\sigma_{min} = 3$, corresponding to the detection of features no smaller than 0.6 nm in radius. This cut-off limited the results for the Room Temperature sample, as smaller features should have been detected to reveal the overestimation of bubble size in manual counts, as observed in the Aftercool sample and described in Teo et al. (2024). The automated results displayed a sharp and narrow Gaussian-like distribution with small deviations for larger radii, while the manual results showed a much wider Gaussian distribution. This highlights the strict selection criteria in automated approaches, increasing precision.

Future work includes refining the blob detection criteria in the Laplace of Gaussian method within the Bubble Convolution code to accurately encompass smaller features. Additionally, introducing more reference points into the relative angle

approach (e.g., four points in four quadrants) and calculating angles only from the furthest reference points could mitigate reference point bias. Given the limited volume (information) provided by TEM images, GISAXS presents a promising future for analyzing sub-surface features like helium nano-bubbles in tungsten. An extension to this project would involve comparing results from GISAXS experiments on bubble counts over a larger volume.

These advancements are crucial for understanding defect nucleation, which is essential for developing robust plasma-facing materials and ensuring the safe and effective operation of fusion reactors. The study's results contribute to the broader understanding of material behavior in extreme environments, paving the way for future innovations in fusion energy technology.

Bibliography

- Bigot, B. 2021, Fusion Engineering and Design, 164, 112207 (Link)
- Carter, C. B., & Williams, D. B. 2016, Transmission electron microscopy: Diffraction, imaging, and spectrometry (Springer)
- dingk. 2024, Bubble Convolution, created on Thu Feb 22 14:50:05 2024
- Energy Institute. 2023, Statistical Review of World Energy (Link)
- Giuliani, U., Agostini, M., Bustreo, C., & Zollino, G. 2023, IEEE Access, 11, 131178
- Griffith, I. O. S. 2013, Fusion Engineering and Design, 88, 1234 (Link)
- Henriksson, K. O., Nordlund, K., Krasheninnikov, A., & Keinonen, J. 2005, Applied Physics Letters, 87
- IPCC. 2018, Global Warming of 1.5°C, edited by Valérie Masson-Delmotte, Panmao Zhai, Hans-Otto Pörtner, et al. (Link)
- . 2019, Representative Concentration Pathway — Wikipedia, The Free Encyclopedia, online; accessed 12-May-2024 (Link)
- Knaster, J., Moeslang, A., & Muroga, T. 2016, Nature Physics, 12, 424 (Link)
- Langridge, R., & Wilson, H. 1960, J. Mol. Bioi, 2, 38
- LoG. 2024, Blob detection — Wikipedia, The Free Encyclopedia, online; accessed 12-May-2024 (Link)
- Morse, E. 2018, Nuclear Fusion (Springer)
- NOAA. 2024a, Monthly average Manua Loa CO₂, online; 09-May-2024 (Link)
- . 2024b, Trends in CO₂, CH₄, N₂O, SF₆, Tech. rep., Mauna Loa Observatory
- Ohno, N., et al. 2013, Journal of Nuclear Materials, 438, S879
- Pitts, R., & Bonnin, X. 2019, Nuclear Materials and Energy, 20, 100696
- Raabe, D. 2014, Physical Metallurgy, 2291
- Renaud, G., Lazzari, R., & Leroy, F. 2009, Surface Science Reports, 64, 255 (Link)
- scikit image. 2024, scikit-image: Image processing in Python, online; accessed 12-May-2024 (Link)
- SciPy. 2024, Differential evolution, online; accessed 12-May-2024 (Link)

- Song, K., Thompson, M., Temmerman, G. D., & Corr, C. 2019, Nuclear Fusion, 59, 096031 (Link)
- Teo, S. H. B., Thompson, M., Bilokur, M., Bhattacharyya, D., & Corr, C. 2024, Materialia, 102110
- Thompson, M. 2018, Helium Nano-Bubble Formation in Tungsten: Measurement with Grazing-Incidence Small Angle X-Ray Scattering (Springer)
- Thompson, M., et al. 2022, Journal of Nuclear Materials, 559, 153448
- Wirth, B. D., Hu, X., Kohnert, A., & Xu, D. 2015, Journal of Materials Research, 30, 1440–1455
- Wirtz, M., Linke, J., Pintsuk, G., De Temmerman, G., & Wright, G. 2013, Journal of nuclear materials, 443, 497
- Yager, K. 2014, GISAXS, this page was last edited on 30 April 2019, at 14:35. (Link)

Appendix

A.1 Modelling fusion environments

Modeling is essential for understanding the environmental impacts of a fusion reactor over its lifetime. However, the challenge lies in the vast timescales involved, spanning over 22 orders of magnitude, from less than a femtosecond onward (Wirth et al. 2015). No single modeling method can cover this range, so researchers use "multi-scale modeling," linking various methods optimized for specific temporal or spatial scales (Knaster et al. 2016). Errors in these models can accumulate exponentially, as each step relies on the previous one. Therefore, empirical validation is crucial to ensure their accuracy.

Appendix B

Appendix

B.1 Grazing Incidence Small Angle X-ray Scattering

Similar to Transmission Electron Microscopy (TEM), Grazing Incidence Small-Angle X-ray Scattering (GISAXS) is used to study the surface and near-surface structure of materials at nanoscale resolution. This technique operates by illuminating a material with a highly collimated, monochromatic X-ray beam at a very low angle relative to the sample surface. This minimizes penetration into the bulk material and focuses on the surface layers (Renaud et al. 2009). The X-rays scatter off the material due to fluctuations in electron density, which can be either an increase, such as from high-density inclusions, or a decrease, such as from voids or pores. According to Babinet's principle, both types of density fluctuations scatter X-rays similarly, allowing GISAXS to effectively analyze various features like nanoparticles, thin films, and surface roughness (Langridge & Wilson 1960).

The scattered X-rays are detected at small angles opposite to the incident beam. By analyzing these scattering patterns, detailed information about the size, shape, and distribution of the nanostructures within the material surface can be extracted. For example, structures like He nanobubbles in W, where the electron density within the bubbles is significantly lower than in the surrounding matrix, result in distinctive scattering patterns. This technique is non-destructive and capable of real-time characterization, making it invaluable for monitoring material structures during processing or under different environmental conditions (Thompson 2018).

One advantage of GISAXS is that samples do not need to be ultra-thin because various methods, such as using a smaller incident angle or a lower refractive index film, can increase the X-ray path length to penetrate the bulk of the material. Additionally, by selecting a specific incident angle, precise illumination of features of interest (like He nanobubbles) can be achieved, excluding signals from the rest of the sample (Yager 2014).

Appendix

C.1 Python code snippets

C.1.1 Differential Evolution code

```

1
2 # Optimization
3 #row +-1 need to be (,2) shape
4 data_plus1_no_rad = data_plus1.drop(columns=['radius'])
5
6 # Define objective function to minimize distance between
7 # corresponding points
8 def objective(params, data_plus1_no_rad,
9               data_minus1_transformed):
10    shift_y, shift_x, angle, factor = params
11
12    # Apply transformation to data_minus1_transformed
13    transformed_positions = []
14    for _, row in data_minus1_transformed.iterrows():
15        transformed_pos = transform([row['y'], row['x']], [
16            shift_y, shift_x], angle, factor)
17        transformed_positions.append(transformed_pos)
18
19    # Calculate total distance between corresponding points
20    total_distance = 0
21    for i, row_plus1 in data_plus1_no_rad.iterrows():
22        min_distance = 60
23        for j, row_minus1 in enumerate(transformed_positions):
24            :
25            distance = np.linalg.norm(np.array(row_plus1) -
26                                      np.array(row_minus1))
27            if distance < min_distance:
28                min_distance = distance
29        total_distance += min_distance
30
31    return total_distance

```

```
28 # Initial guess for optimization
29 initial_guess = [-10, -70, -0.05, 1.05]
30
31 # Define bounds for parameters
32 bounds = [(-100, 100), (-100, 100), (-np.pi/4, np.pi/4),
33           (0.9, 1.1)]
34 # Define the callback function to print the progress
35 def callback(xk, convergence):
36     print("Iteration:", len(xk), "Best Objective Value:",
37           convergence)
38
39 # Perform optimization with callback
40 result = differential_evolution(objective, bounds, args=(
41     data_plus1_no_rad, data_minus1_transformed), callback=
42     callback)
43
44 # Get optimized parameters
45 shift_y, shift_x, angle, factor = result.x
46
47 print(result)
```

Listing C.1: Differential Evolution method for optimisation of transformation parameters.

C.1.2 Relative angle bubble matching loop

```

1 #Load in all the data and manually select parameters
2
3 ...
4
5 # Iterate through each bubble in the over-focused dataset
6 for index_plus1, bubble_plus1 in data_plus1.iterrows():
7     # Find nearest neighbors in both datasets
8     neighbors_plus1 = find_neighbors(bubble_plus1,
9         nn_index_plus1, data_plus1)
10    neighbors_minus1 = find_neighbors(bubble_plus1,
11        nn_index_minus1, data_minus1_transformed)
12
13    # Check angles between each pair of neighbors
14    for neighbor_plus1 in neighbors_plus1:
15        for neighbor_minus1 in neighbors_minus1:
16            if np.linalg.norm(np.array([neighbor_plus1['y'],
17                neighbor_plus1['x']]) - np.array([
18                    neighbor_minus1['y'], neighbor_minus1['x']]))[
19                        <= distance_threshold:
20                            if is_bubble(neighbor_plus1, neighbor_minus1,
21                                reference_plus1, reference_minus1):
22                                data_plus1.at[neighbor_plus1.name, 'bubble'] = 1
23                                data_minus1_transformed.at[
24                                    neighbor_minus1.name, 'bubble'] = 1
25
26 ...
27
28 #Plot results

```

Listing C.2: Relative Angle matching Loop for identifying bubbles.

C.1.3 Python code

The rest of the code for both direct and relative angle matching can be found at:
[Bubble Code](#)