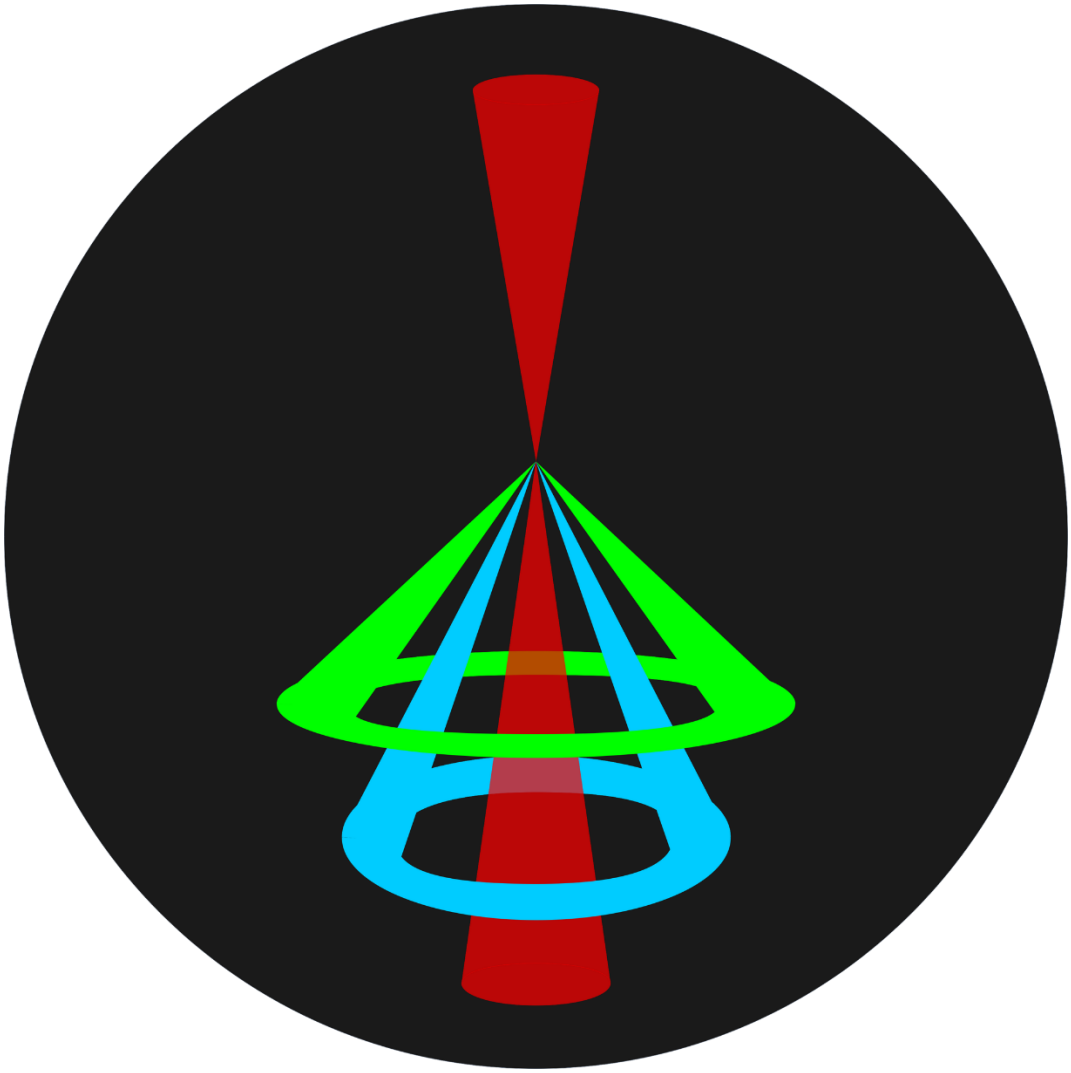


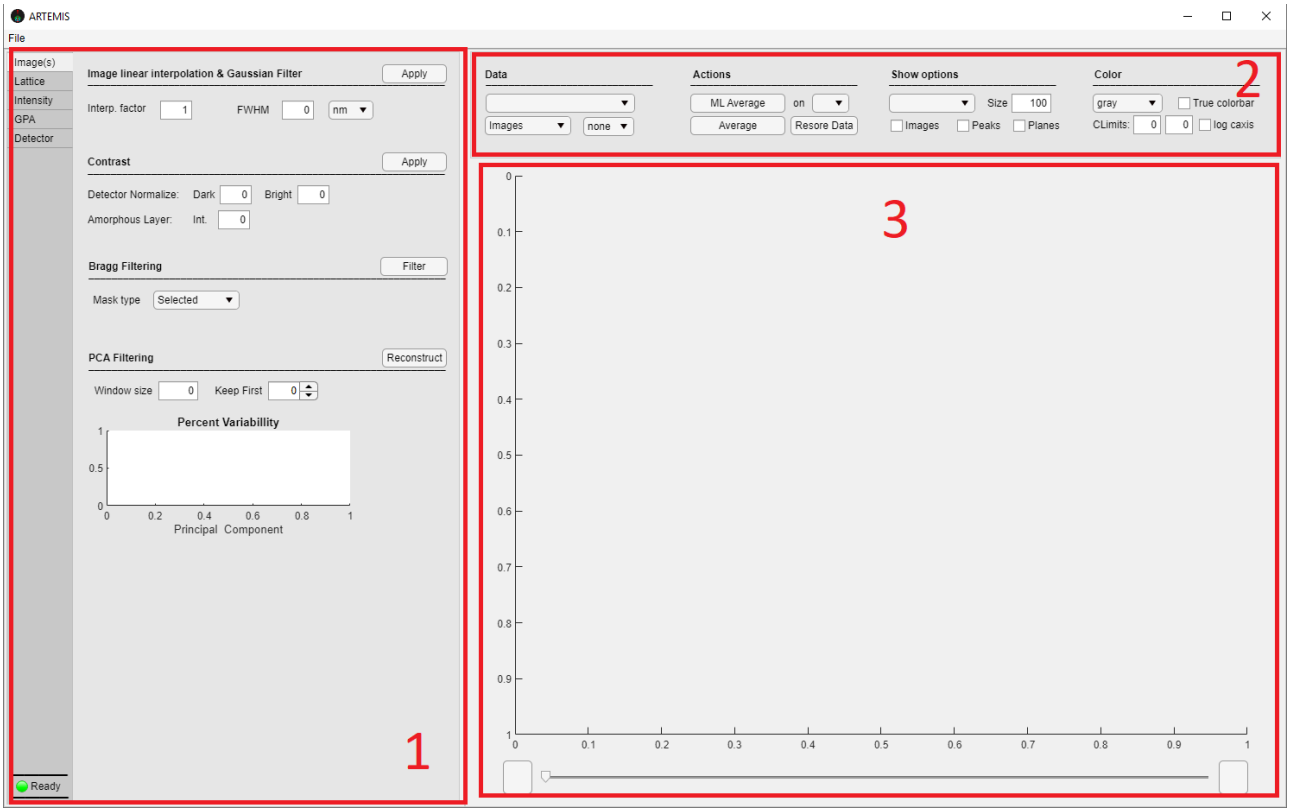
ARTEMIS 1.0 Documentation

By Athanasios Gkotinakos

1. About ARTEMIS

ARTEMIS is a Matlab app for performing analysis of atomic resolution electron microscopy images. It delivers a variety of tools to process images, make quantitative measurements, construct, and analyze atomic lattices, all in an intuitive graphical user interface (GUI).

2. GUI Overview



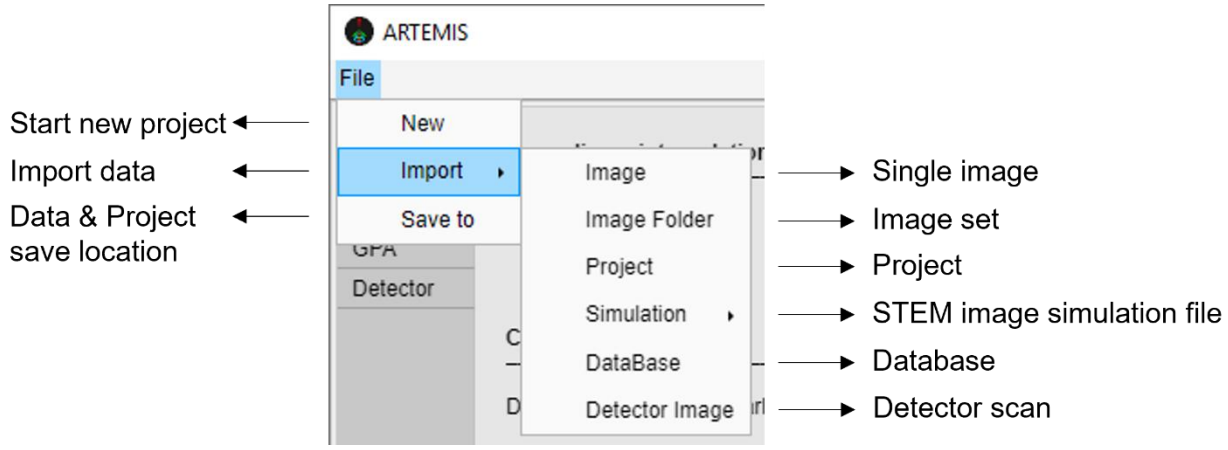
GUI consists of three main sections:

1. Analysis tab group on the left.
2. Data and Figure options Panel on the top right.
3. Figure axes on the right.

The user can setup and deploy various tools in the analysis tab group. Imported data (images, simulation files etc.) and generated data appear on figure axes. You can setup figure options, revisit and manipulate generated data on section 2.

2.1 File menu

Access "***File***" menu to start new projects, import data, and to choose saving locations.



Start a new project from ***New***. Output data, figures and the project itself is saved in a specific location provided at ***Save to***.
Projects are saved in .mat file format after a new project is initialized. Project files can be imported to revisit the project from ***Import -> Project*** option.

To import a single STEM or ADF detector image go to ***Import -> Image***, ***Detector Image***, respectively. To import a set of STEM images go to ***Import -> Image Folder***. The supported image file format is .tif. After you successfully import image(s), the figure section will refresh and display the imported file. Also, Fast Fourier Transform is applied to imported image(s) to derive a frequency-filtered version of an image(s), necessary for later analysis and visual inspection.

To import a database, select ***Import -> DataBase*** and choose the database's folder location. Databases will be explained in more depth later, here we outline the general form a database should have:
A database should consist of atomic column's intensity - thickness and ratio – thickness – composition values, in txt format with the following name and form:

1. Name: Intensity-Thickness -> Txt file that contains atomic column's intensity at different thicknesses (data column 1: intensity values, data column 2: thickness values)
2. Name: Ratio## -> File that contains atomic column's intensity ratio at different thicknesses (data column 1: ratio values, data column 2: thickness values). ## is the corresponding specimen composition, e.g., Ratio0.2, Ratio22.38.

Data columns should be comma delimited. We suggest looking at the exemplar database of *"Examples"* folder.

STEM image simulation output files are loaded from ***Import -> Simulation***, and used to construct and/ or import the contained simulation images.
Supported simulation outputs:

1. STEMSalabim: Output file holds simulation information, parameters, and simulated images data. When images are discretized to frozen phonon configurations the average image is

calculated. When images are discretized to scattering angles the real detector's angular sensitivity can be considered. Imported detector's angular sensitivity should be in .txt format (data column 1: sensitivity, data column 2: angle / comma delimited, see Examples Folder file: Detector_Sensitivity_Example.txt).

2.2 Image(s) Tab

On Image(s) tab user can process and cleanup images prior to quantitative analysis.

Image linear interpolation & Gaussian Filter

Enter an interpolation factor and/or a FWHM value and click *Apply* to perform linear interpolation and/or convolution of the image with a gaussian surface.

- Linear interpolation factor determines the output size of the image i.e., for an image of size 1024x1024 pixels and Interp. factor = 2 the output image will be 2048x2048 pixels wide. Generated pixels values are the weighted average in the nearest 2 by 2 neighborhoods.
- FWHM of the gaussian surface is provided in pixels or nanometers. Convolution of the image and the gaussian surface is performed in frequency domain.

Contrast

Enter *Dark, Bright, Amorphous Layer intensity* and click *Apply* to perform image normalization with the detector's characteristic intensity values and to subtract a constant amorphous layer intensity.

- Images intensities are normalized as $\frac{I_{raw}-I_{Dark}}{I_{Bright}-I_{Dark}}$. Amorphous Layer intensity is subtracted after normalization.

Bragg Filtering

Create *ROIs* to mask regions of FFT Images or FFT Denoised Images, select a *Mask type* and click *Filter* to perform Bragg filtering.

- To create ROIs access *figure's menu* (right click on the figure section).
- From *Mask type* select *Selected* or *Reverse* to clip unmasked or masked areas respectively.

PCA Filtering

Enter a *Window size*, the number of significant principal components to keep (*Keep First*) and click *Reconstruct* to perform PCA filtering. If *window size* equals to 1, each row of the image matrix is considered as an individual observation.

- *Window size* determines the size of the moving window that copies observations of the image.
- *Keep First* determines the number of most significant principal components to be considered in the reconstruction process.

Theory

Quantitative analysis of STEM images usually requires image processing before implemented. In the case of experimental images, intensity normalization of the image is essential for direct comparison with intensity measurements from simulation images. Image normalization is

discussed in detail in the following reference [1]. Using experimental ADF detector's dark and bright intensity, images are normalized as:

$$I_{Norm}(x,y) = \frac{I(x,y) - I_{Bright}}{I_{Bright} - I_{Dark}}$$

where $I(x,y)$ is the image's raw intensity as a function of x, y, horizontal and vertical pixels respectively, $I_{Norm}(x,y)$ is the normalized intensity and I_{Bright}, I_{Dark}, the ADF detector's surface mean intensity and mean vacuum intensity respectively.

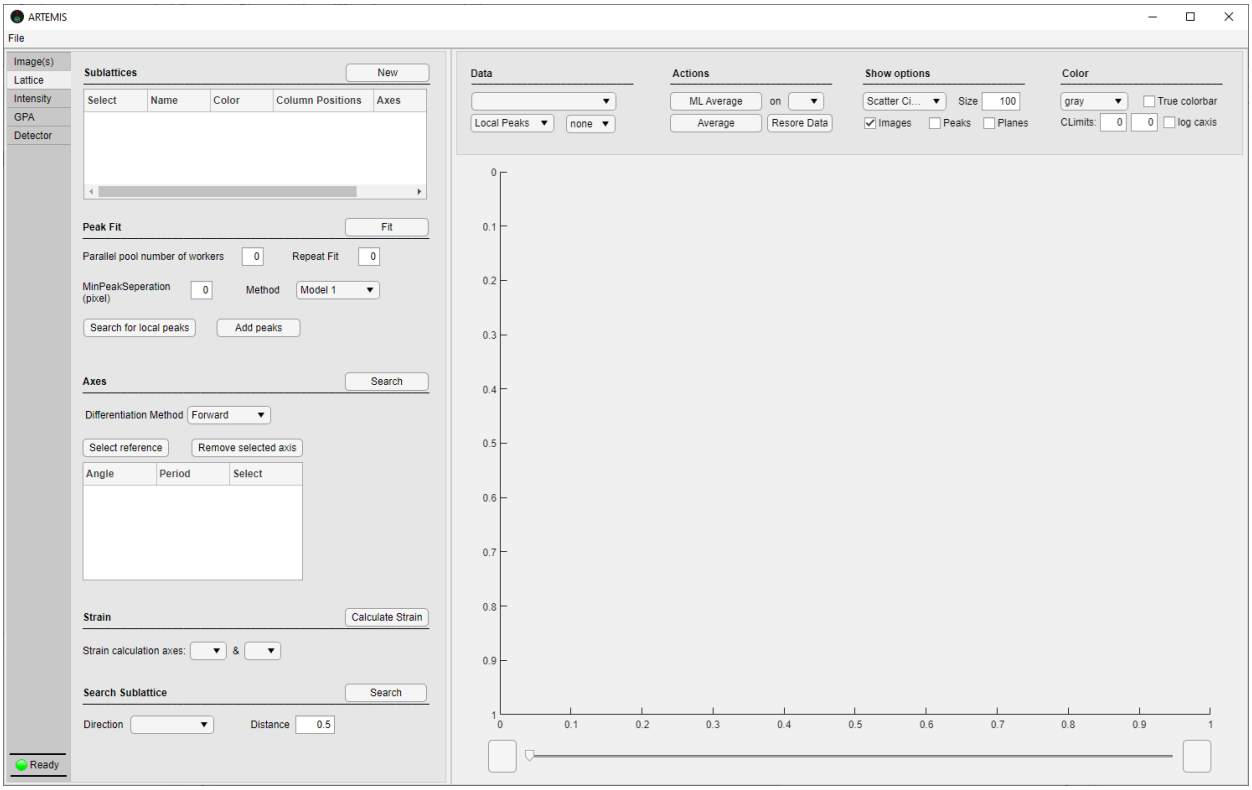
To properly construct simulation images the effect of the experimental ADF detector's angular sensitivity and the finite electron source should be considered. STEM image simulations usually output the angular intensity that reaches the ADF detector (an array of intensities binned over the angular range of the ADF detector) for each scan point. The detector's sensitivity can be considered by multiplying each scan point angular intensity array with the detector's angular sensitivity. Images are then formed by summing the intensities array over the detector's angle range. The finite electron source size effect on the image intensity is considered by convoluting the simulated images and a gaussian distribution with FWHM that corresponds to the source's diameter.

Bragg filtering is performed in reciprocal space. The Fourier image of crystals consists of bright spots emerged from the periodicities of crystal planes superimposed with a background intensity due to random noise. The first step to implement Bragg filtering is to correctly identify and selectively mask from the rest of the image the Fourier components that correspond to the periodic plane spacing of the imaged crystal. The Fourier image that consists only of true lattice symmetry components is then inverted back to real space, creating the denoised image.

PCA filtering is a more sophisticated method to denoise images, and its implementation requires a lot of computer memory and computational power. It is usually deployed on sets of observations to extract the dominant characteristics and there are different methodologies described in literature to filter STEM images. A detailed description of the method ARTEMIS implements is provided here []. Observations of the image are copied by a moving window of suitable size. An observations matrix is constructed that contains copied observations flattened to arrays. The principal components of the matrix are calculated and are sorted according to their significance. Least significant components correspond to random intensity noise and are filtered out. The image is then reconstructed following the reverse procedure utilizing only the most significant principal components.

2.3 Lattice Tab

On Lattice tab user can construct and analyze lattices.



Sublattices

Click *New* to create a sublattice object. Different sublattices are analyzed independently. To delete selected sublattice(s), right click on *Sublattices Table* to open the menu and click *Delete selected*. To change the name of a sublattice double click the corresponding name cell from the *sublattices Table* and enter a new name. To change sublattice peaks color click the corresponding *Color cell* and pick from the available colors.

Peak Fit

To locate local peaks, enter a *MinPeakSeperation* value and click *Search for local peaks*. To manually insert peaks from the figure, click *Add peaks* push button. When you finish inserting peaks click *Add peaks* push button again. Choose a gaussian model from the *Method* dropdown and provide the number of repetitive fits on the *Repeat Fit* edit field. Click *Fit* to complete peak fit.

- *Parallel pool number of workers* determines the number of CPU cores used by ARTEMIS to parallelize peak fit calculations.
- *MinPeakSeperation* defines a minimum distance between atomic columns. Try Searching for local peaks with different MinPeakSeperation values and see what works best.
- *Repeat Fit*: If > 1, column's intensity distribution is modeled as a sum of gaussians (see Theory).

Axes

Zone axes analysis is available after peak fit. Choose a *Differentiation Method* and click *Search* to identify axes of high symmetry. Axes angles and mean atomic spacing appear on *Axes Table*. Create ROIs (right click figure -> *Roi*) to highlight unstrained reference areas and click *Select reference* to update *Axes Table* with the refence spacing and axes angles. Select axes and click *Remove selected axis* to remove misidentified axes.

Strain

To calculate sublattice's strain tensor in real space, choose two of the axes. It is important to Select a reference region before continuing to strain calculations.

Search Sublattice

Provide a *direction* and the appropriate *distance fraction* to search for sublattices of intermediate, unidentified atomic columns along the planes of the selected direction.

- *Distance* edit field refers to the distance of intermediate peaks from the current sublattice's peaks along the chosen axis.

Theory

Peak finding is used to locate column's positions with subpixel precision. In the simple case, where the atomic columns are easily distinguished and there is no overlapping between neighboring columns, the gaussian peak fit is an efficient method to locate columns' positions. The peak finding steps implemented by ARTEMIS are the following:

1. Intensity local maxima of the image(s) are first located considering the *MinPeakSeperation* value.
2. Local maxima are refined using the center of mass around each atomic column.

$$x'_{peak} = \frac{\sum_{x_i,y_i \in C} I(x_i,y_i)x_i}{\sum_{x_i,y_i \in C} x_i}$$

$$y'_{peak} = \frac{\sum_{x_i y_i \in C} I(x_i,y_i)y_i}{\sum_{x_i,y_i \in C} y_i}$$

where x'_{peak}, y'_{peak} are the refined peak's position coordinates, C is a circular locus around the original peak position with radius equal to a percentage of the column's first neighbor distance, x_i, y_i are the pixel coordinates that belong to C, $I(x_i, y_i)$ is the pixel intensity at (x_i, y_i). Locus C is also refined after peak refinement.

3. 2D gaussian surface model is fitted on the refined locus C'. Two different types of gaussian models are available:

$$G(x,y) = G_0 + A \exp\left(-\left(\frac{(x-x_0)+(y-y_0)}{\sigma}\right)^2\right)$$

and

$$G(x,y) = G_0 + A \exp\left(-\left(\frac{(x-x_0)cos\theta + (y-y_0)sin\theta}{\sigma_x}\right)^2 - \left(\frac{(x-x_0)sin\theta + (y-y_0)cos\theta}{\sigma_y}\right)^2\right)$$

where $G(x,y)$ is the gaussian intensity as a function of the two spatial directions, A is the amplitude of the gaussian peak, σ, σ_x, σ_y are constants proportional to the distribution variance, x_0, y_0 is the

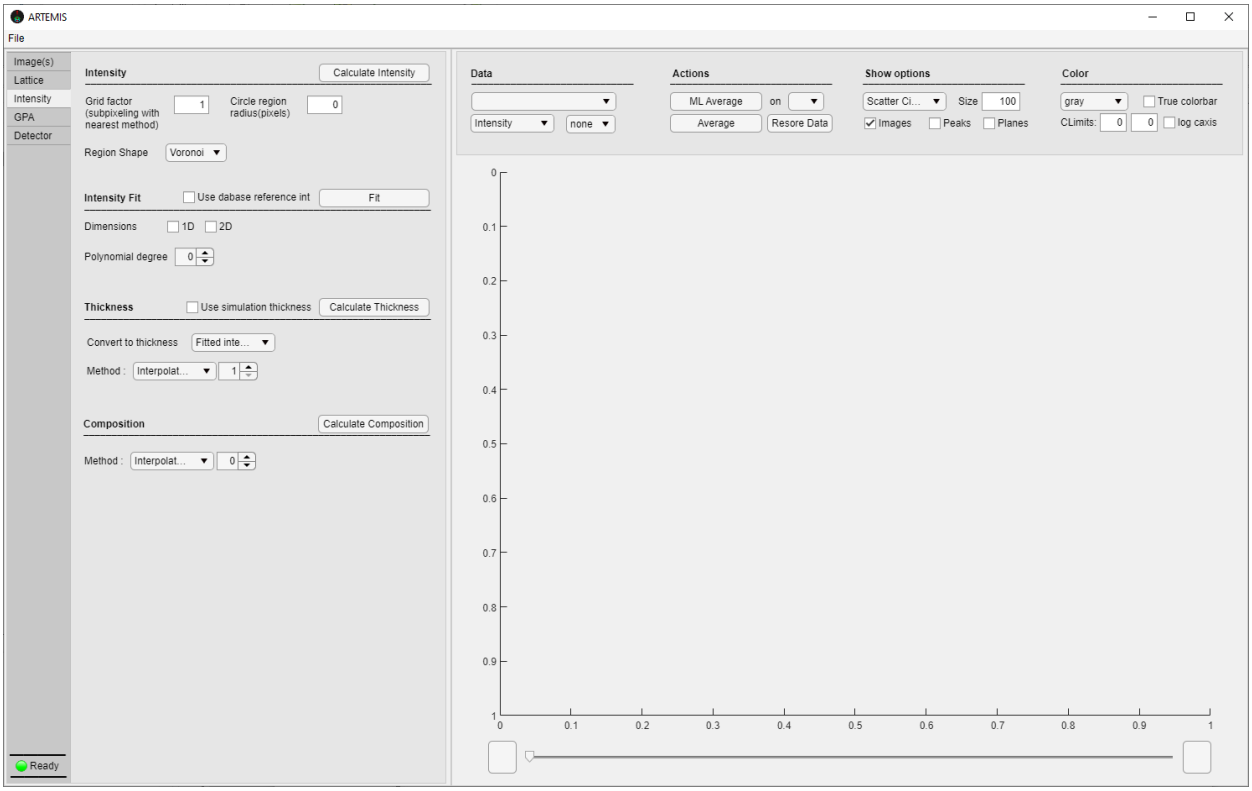
peak position, θ is the rotation of the gaussian in the counterclockwise direction on xy plane and G_0 is a constant intensity background. The first gaussian plot should be used in cases where atomic columns appear circular, and the second should be used in cases where atomic columns appear elongated.

In cases where the image is subjected to significant aberrations or neighboring atomic columns slightly overlap, the intensity distribution around columns deviates from the simple gaussian distribution and a more robust method is required to accurately measure column's positions. In such case, as discussed in the reference [**2**], column's intensity is more accurately described as a sum of overlapping gaussians, and a good approximation can be derived using a subset of these gaussians. For this cause repetitive gaussian fits are preformed were after each fit, the gaussian surface is subtracted from the image intensity. Column positions are then calculated from the mean value of a subset of the gaussian peak positions weighted by the peak intensity. The subset contains peaks that fall inside a circular area with a radius equal to a percentage of the first neighbor distance.

Zone axes analysis is implemented to locate high symmetry axes, extract the spatial spacing periods of atomic columns along axes planes, distances, and angles between neighboring columns. ARTEMIS zone axes analysis algorithm uses as input the positions of the first 40 neighbors of each atomic column. Distances and angles are calculated from the neighbor's positions, for the whole set of peaks. Statistics on these values provides the angles of high symmetry and column's periodicity along planes. Groups are then created for each located axis, that contains subgroups of peaks that belong to the axis planes. Planes are sorted in a counterclockwise fashion and peaks are sorted along each plane in the direction of the axis. First neighbor distances and angles are calculated along the plane direction by forward, backward or central differentiation. Bond length variation is also calculated as the distance mismatch from the mean spacing of atomic columns of a reference (unstrained) region.

2.4 Intensity Tab

On Intensity Tab user can quantify atomic column's intensity and compare intensity values with simulations to convert measured intensity to thickness and composition. To execute thickness and composition calculations, the simulation database is required (*File -> Import -> DataBase*).



Intensity

Enter a *Grid factor*, choose a tessellation shape from *Region Shape* dropdown and click *Calculate Intensity* to quantify and map atomic columns' intensities.

- When *grid factor* is > 1 the image is interpolated via the box method (new pixels have the same value as the pixels they fall in). This improves the precision of the tessellation process and the intensity measurements' accuracy overall at the expense of computational power.
- *Circle region radius* value determines the radius or the box size of circle or box regions respectively.
- Voronoi, Circle and Box shapes are available for intensity integration.

Intensity Fit

Choose analysis *dimension* and draw a *ROI* on figure to highlight areas that should NOT be considered for intensity fit. Select the appropriate *Polynomial degree* and click *Fit* to perform intensity fit.

Thickness

Choose an *interpolation* or *fitting* method for the intensity – thickness database data and click *Calculate Thickness* to compare with the database values and convert fitted intensity values or *Selected intensity* values to thickness.

Composition

Choose one of the available methods to treat database ratio-thickness-composition values of the database and click *Calculate Composition* to compare with the database and convert intensity ratio and thickness values to composition.

Theory

To perform thickness and composition quantification on experimental Z-Contrast HAADF STEM experimental images, an extensive series of image simulations is required on carefully modeled and relaxed supercells with varying thickness and composition. Thickness can be estimated on experimental images by comparing atomic columns' intensity values in regions of known composition with corresponding column intensity values from image simulations of supercells with varying thickness. Composition on the other hand can also be estimated on experimental images in areas with known thickness, by comparing columns' intensity ratio values with corresponding values from image simulations of varying supercell thickness and composition. So, a simulation database is required to translate intensity and intensity ratio to thickness and composition. A database should consist of column intensity values at different supercell thicknesses, and of column intensity ratio at different supercell thicknesses and compositions. In the case of thin film specimens containing 2D or 1D structures (Quantum Wells or Quantum Dots) the procedure of composition quantification includes fitting columns' intensity from regions with known composition, over the regions of unknown composition. ARTEMIS uses 1D or 2D polynomial functions to execute this step. The intensity ratio of measured intensity and fitted (reference) intensity can then be calculated. Thickness and ratio values are then compared against the database values and columns' composition is determined.

ARTEMIS offers all necessary tools to execute the procedure. Provided a database, Intensity and Thickness database values are linear interpolated or fitted with a polynomial function to produce a continuous function which can be used to convert measured columns' intensity to thickness. Ratio – Thickness-Composition values are also transformed to continuous function(s):

1. "*Interpolate Linear*": 2D linear interpolation of ratio, thickness, and composition database values.
2. "*Fit Polynomial*": Ratio and thickness database values for different compositions are polynomial fitted. Ratio vs composition values are determined from measured columns' thickness and are also fitted with polynomial functions to convert measured ratio and thickness to composition. The degree of the polynomials used is defined by the user and in case where there are not enough ratio - composition points, is automatically reduced on the second step.
3. "*Fit 2D polynomial*": 2D polynomial fit of ratio, thickness, and composition database values.
4. "Interpolate Linear + Fit 2^{nd} deg inverse poly": Ratio and thickness database values for different compositions are linearly interpolated. Ratio vs composition values are determined

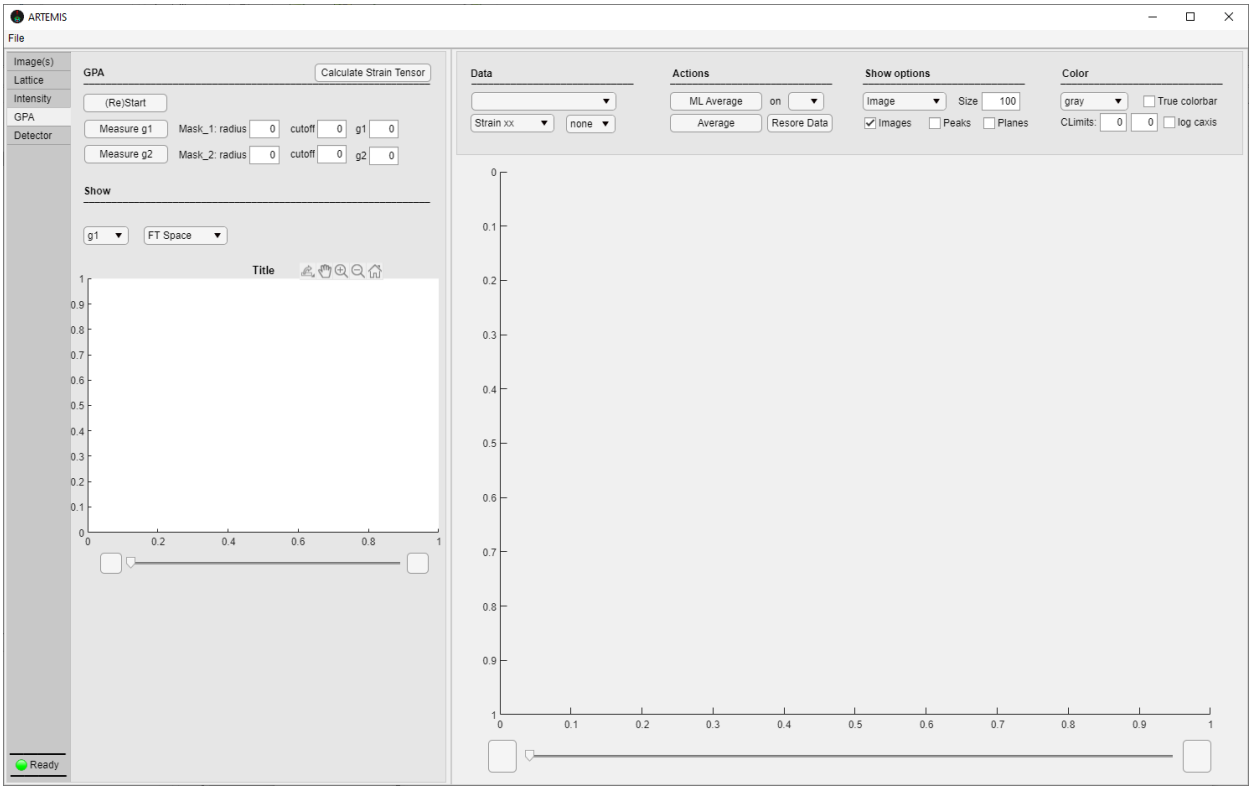
from columns' thickness and are fitted with 2nd degree inverted polynomial functions to convert measured ratio and thickness to composition.

5. "*Fit poly* + *Interpolate Linear*": Ratio and thickness database values for different compositions are fitted with polynomials. Ratio vs composition values are determined from measured columns' thickness and are linearly interpolated to convert measured ratio and thickness to composition.
6. "*Fit poly* + *Fit 2nd deg inverse poly*": Ratio and thickness database values for different compositions are fitted with polynomials. Ratio vs composition values are determined from measured columns' thickness and are fitted with 2nd degree inverted polynomial function to convert measured ratio and thickness to composition.

Composition and thickness calculations can be performed either in one or two dimensions. Measured 1D ML averaged intensity data are used for thickness/composition calculations performed in one dimension. Measured 2D columns' intensity data are used for two dimensional calculations of thickness and composition.

2.5 GPA Tab

On GPA Tab user can perform GPA analysis.



GPA

Click *(Re)Start* to initialize GPA analysis. Enter *mask* and *cutoff radius* and click *Measure g1/g2* to manually measure g vectors from FT Space image (click the desired Fourier g components *on Show figure*). To refine a g – vector Create *ROIs* in Show figure and select unstrained reference areas on phase image, right click *ROIs* and select *refine g vector*. Click *Calculate Strain Tensor* to complete GPA.

Show

In *Show figure* user can visualize the FT Space of the image, g1-g2 phase and Bragg images.

Theory

Geometric phase analysis is also included in ARTEMIS app. The procedure followed is described in detail in the resources [3,4]. An image can be represented with a Fourier series as:

$$I(\boldsymbol{r}) = \sum_{\boldsymbol{g}} A_{\boldsymbol{g}} e^{iP_{\boldsymbol{g}} + 2\pi \boldsymbol{g}\boldsymbol{r}}$$

where $I(\boldsymbol{r})$ is the intensity as a function of the position $\boldsymbol{r}$ (pixel coordinate), $\boldsymbol{g}$ is a reciprocal vector and $A_{\boldsymbol{g}}$, $P_{\boldsymbol{g}}$ are the amplitude and phase of the $\boldsymbol{g}$ vector. By masking a g component of the Fourier transformed image and then applying an inverse Fourier transformation, a complex image is produced. $\boldsymbol{g}$ vectors with significant $A_{\boldsymbol{g}}$ correspond to lattice, plane periodicities. If the planes are

distorted, then the resulting masked inverse Fourier transformation of the reciprocal image will lead to a space dependent $A_{\boldsymbol{g}}(\boldsymbol{r})$ and $P_{\boldsymbol{g}}(\boldsymbol{r})$ component. Strain tensor calculations require the phase images of two noncolinear $\boldsymbol{g}$ vectors, $\boldsymbol{g_1}$, $\boldsymbol{g_2}$. $P_{\boldsymbol{g_1}}(\boldsymbol{r})$, $P_{\boldsymbol{g_2}}(\boldsymbol{r})$ are calculated by extracting the $2\pi \boldsymbol{gr}$ term from the complex images phase and the distortion tensor can be derived as:

$$E_{xx}(\boldsymbol{r}) = -\frac{1}{2\pi}\left(\frac{\frac{\partial}{\partial x}\left(P_{g_1}(\boldsymbol{r})\right) g_{2y} - \frac{\partial}{\partial x}\left(P_{g_2}(\boldsymbol{r})\right) g_{1y}}{g_{1x}g_{2y} - g_{1y}g_{2x}}\right)$$

$$E_{yy}(\boldsymbol{r}) = -\frac{1}{2\pi}\left(\frac{\frac{\partial}{\partial y}\left(P_{g_2}(\boldsymbol{r})\right) g_{1x} - \frac{\partial}{\partial y}\left(P_{g_1}(\boldsymbol{r})\right) g_{2x}}{g_{1x}g_{2y} - g_{1y}g_{2x}}\right)$$

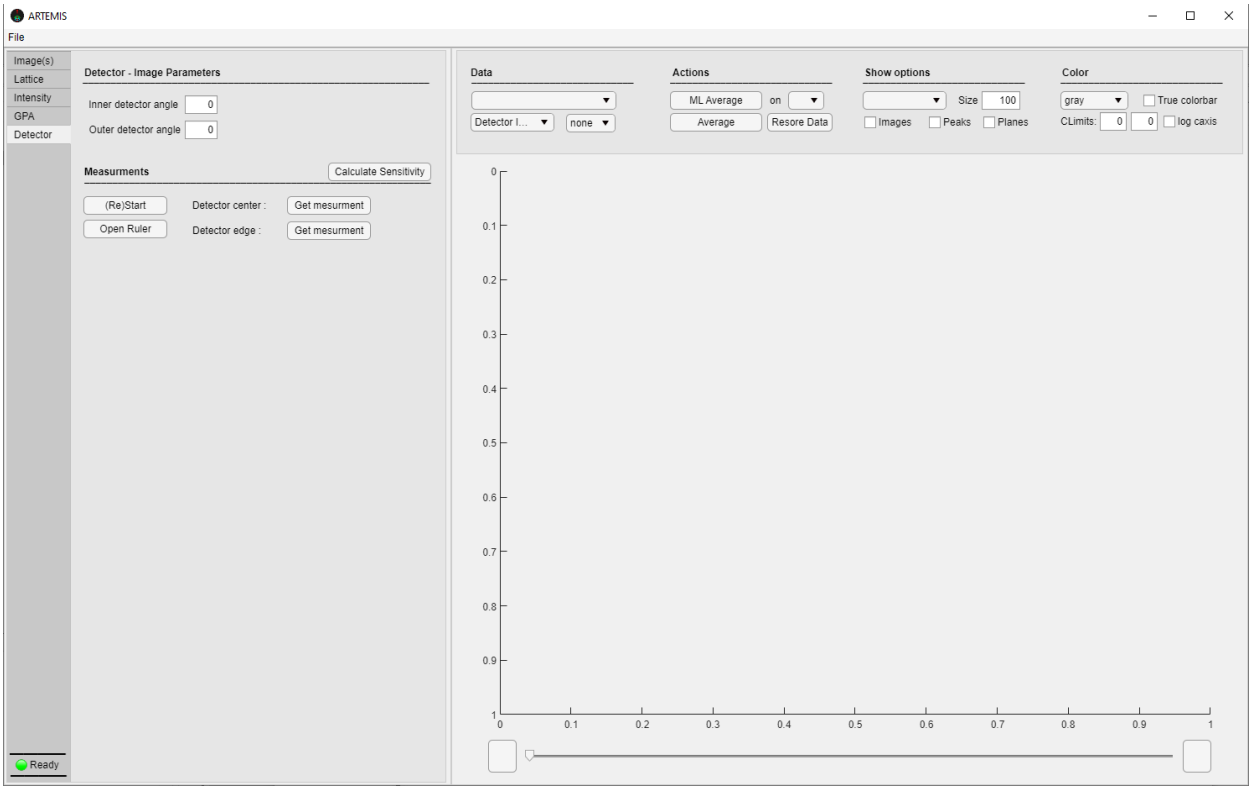
$$E_{xy}(\boldsymbol{r}) = -\frac{1}{2\pi}\left(\frac{\frac{\partial}{\partial y}\left(P_{g_1}(\boldsymbol{r})\right) g_{2y} - \frac{\partial}{\partial y}\left(P_{g_2}(\boldsymbol{r})\right) g_{1y}}{g_{1x}g_{2y} - g_{1y}g_{2x}}\right)$$

$$E_{yx}(\boldsymbol{r}) = -\frac{1}{2\pi}\left(\frac{\frac{\partial}{\partial x}\left(P_{g_2}(\boldsymbol{r})\right) g_{1x} - \frac{\partial}{\partial x}\left(P_{g_1}(\boldsymbol{r})\right) g_{2x}}{g_{1x}g_{2y} - g_{1y}g_{2x}}\right)$$

Strain tensor, ε_{ij}, and local rigid rotation, ω_{ij} is obtained from the distortion's tensor symmetric and antisymmetric terms as $\varepsilon_{ij}(\boldsymbol{r}) = \frac{1}{2}\left(E_{ij}(\boldsymbol{r}) + E_{ij}(\boldsymbol{r})^T\right)$, $\omega_{ij}(\boldsymbol{r}) = \frac{1}{2}\left(E_{ij}(\boldsymbol{r}) - E_{ij}(\boldsymbol{r})^T\right)$. The mean ramp of an unstrained region on the phase image, $P_{\boldsymbol{g}}(\boldsymbol{r})$, can be used to refine any of the selected $\boldsymbol{g}$ vectors as: $\Delta \boldsymbol{g} = \frac{1}{2\pi}\overline{\nabla P_g(\boldsymbol{r})}$

2.6 Detector Tab

On Detector Tab user can measure ADF detector's angular sensitivity, bright and dark detector intensities.



Detector – Image Parameters

Provide the detector's image (*File -> Import -> Detector Image*). Insert the nominal *inner* and *outer* detector angles in degrees.

Measurements

Click *(Re)Start* to initialize detector measurements. Click "*Open Ruler*" to create a ruler on the detector's image. Use the ruler to make measurements on the detector's image figure:
1) Measure the diameter of the detector's inner hole.
2) Measure the outer detector's diameter.
When enough measurements are acquired click "*Sensitivity Calculation*" to calculate and display the detector's binned intensity and angular sensitivity profile, dark and bright intensity.

Theory

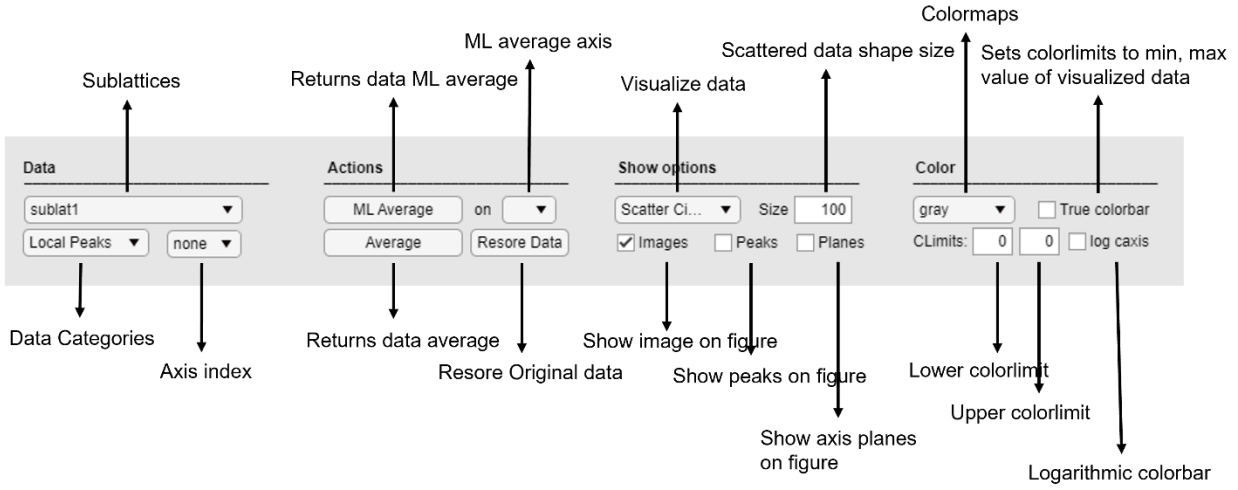
Detector sensitivity is essential in constructing accurate STEM simulation images. ARTEMIS provides a way to measure detector bright, dark intensity and an estimation of angular sensitivity. The detector is assumed to be a disk with perfectly circular inner and outer boarders. Enough measurements around the perimeter of the inner and outer detector boarders should give an acceptable estimation for the detector's center, inner and outer radius. The mean intensity value

of the detector surface is then extracted, referred as bright or detector intensity, as well as the mean intensity outside detector's surface, referred as dark or vacuum intensity, by averaging the pixels that fall between and outside, respectively, of the measured inner and outer detector radius. Angular sensitivity is calculated using:

$$S(\theta) = \frac{I(\theta) - I_{Bright}}{I_{Bright} - I_{Dark}}$$

where $S(\theta)$ is the angular sensitivity as a function of the angle from the optical axis.

2.7 Data & Figure Options Panel



Data

Use the dropdowns to visualize data categories.

- Different analysis tabs give access to different categories of data once analysis is completed. You can navigate through sublattices, data categories and axes dependent data from *Data* subsection. Once a selection is made, the *figure* is updated, and the respective data are visualized.

Actions

Choose an axis direction and click *ML Average* to average scattered data to monolayers. Click *Average* to average data. Click *Restore Data* to retrieve the data that was originally generated.

- When Zone axes analysis is performed successfully, axes can be used to average data to monolayers (ML) along the selected axes direction. ML separation is calculated as the distances of the midst points of neighboring planes and is used as the horizontal axis of the ML averaged data.

Show options

Choose plot types, superimpose visualized data with images and/or peaks and/or axis planes.

There are different plot types available to visualize data of the selected data category:

- *Scatter Circle* -> Data are visualized on column's positions as colored circles.
- *Scatter Box* -> Data are visualized on column's positions as colored boxes.
- *Maps* -> (Available after voronoi tessellation) Data are visualized at tessellated regions.
- *ML* -> Scatter plot of ML averaged data.
- *Averages* -> Scatter plot of averaged data with respect to image slice.
- Image -> Visualizes the image of matrix data.

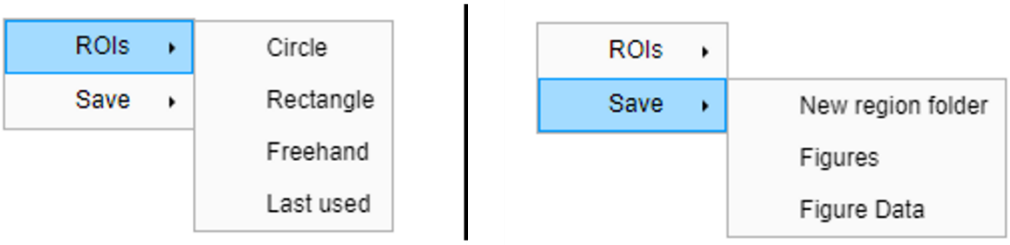
Color

Control the appearance of the visualized data. Select from the available Colormaps, manually define colorbar limits and/or change colorbar to logarithmic.

By default, the figure's colorbar is set to the minimum and maximum value of the visualized data category. If you are analysing more than one image (slice) you can set the colorbar limits to the minimum and maximum value of current slice by clicking *True colorbar*.

2.8 Selecting and saving data and figures

As ARTEMIS is designed to simultaneously analyze many images and offers a wide range of analytical tools that produce huge amount of data, a simple yet not restrictive method to selectively handle and save meaningful output is also provided. Right click on the figure area to bring up the figure menu.



ROI

To select data from the figure, expand the *ROI* menu and select a type of ROI. Highlight the region of interest and right click on ROI to bring up the ROI menu. Select *Keep* to remove the unselected data or *Remove* to remove the selected data.

Save

To save data and figures:

1. **Provide output folder path (*File -> Save To*).**
2. **Make data selections with *ROI(s)*. (optional)**
3. ***ML Average* or/and *Average Selected*.**
4. **To save data select *Scatter Box* or *Scatter Circle* or *Maps* or *Image* option from the Show options dropdown and click *Figure Data*. To save ML Averaged data select *ML* option in *Data Show* dropdown and click *Figure Data*. To save averaged data select Averages in *Data Show* dropdown and click *Figure Data*.**
5. **To restore original raw data and collect data from a new region click *Restore Original* and *New region folder* and repeat steps 2-4.**
6. **To save figures right click *Figures*.**

Figures and data are saved inside the respective data category folder.
If you want to collect data from more than one region, a new output folder is created by clicking *New region folder*.
All data are saved in .txt files organized in folders on the provided output folder path. All images and figures are saved in .tif format. Every time a new session is initiated (*File -> New*), a main output folder is created on the provided output folder path named Data(Date, Time), e.g. Data07-23-2024 13-46-22.

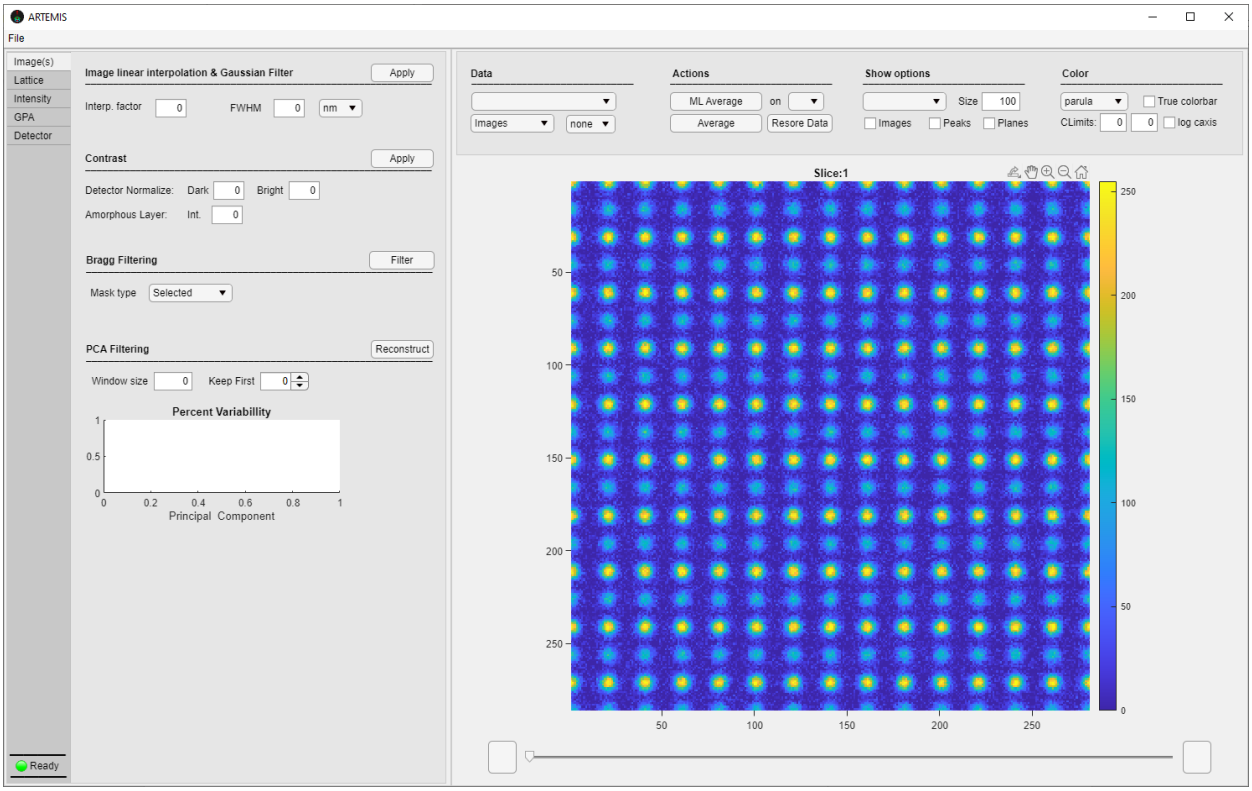
3. Example

3.1 Extracting the lattice structure

In this example a typical project workflow is presented. Our aim is to denoise the case image, build and extract information from the projected lattice structure.

Importing

First create a new, clean project and provide a save location. Locate and import this case image (see 2.1).

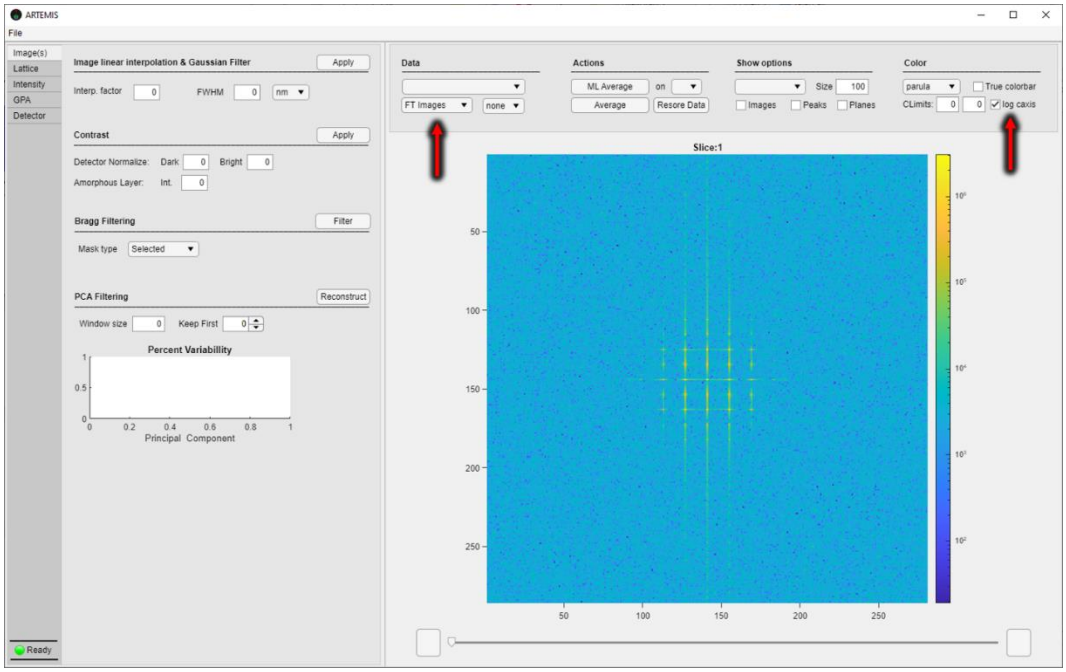


Denoising

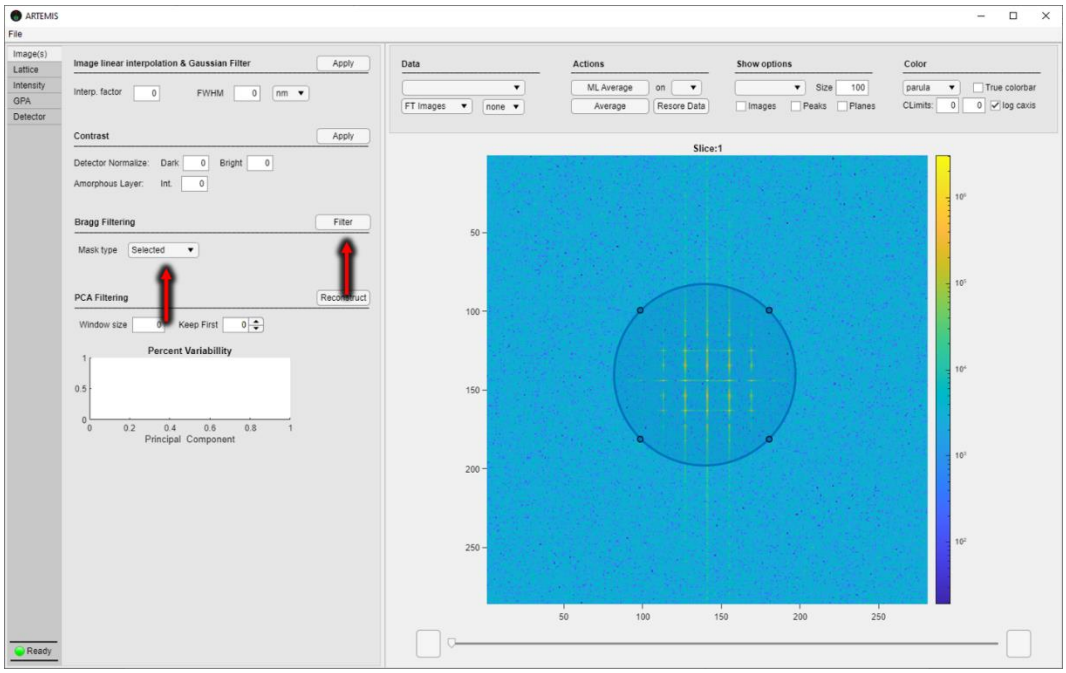
We will first try to dampen the image noise without obscuring the general image characteristics. Different denoising procedures are demonstrated.

Bragg Filtering

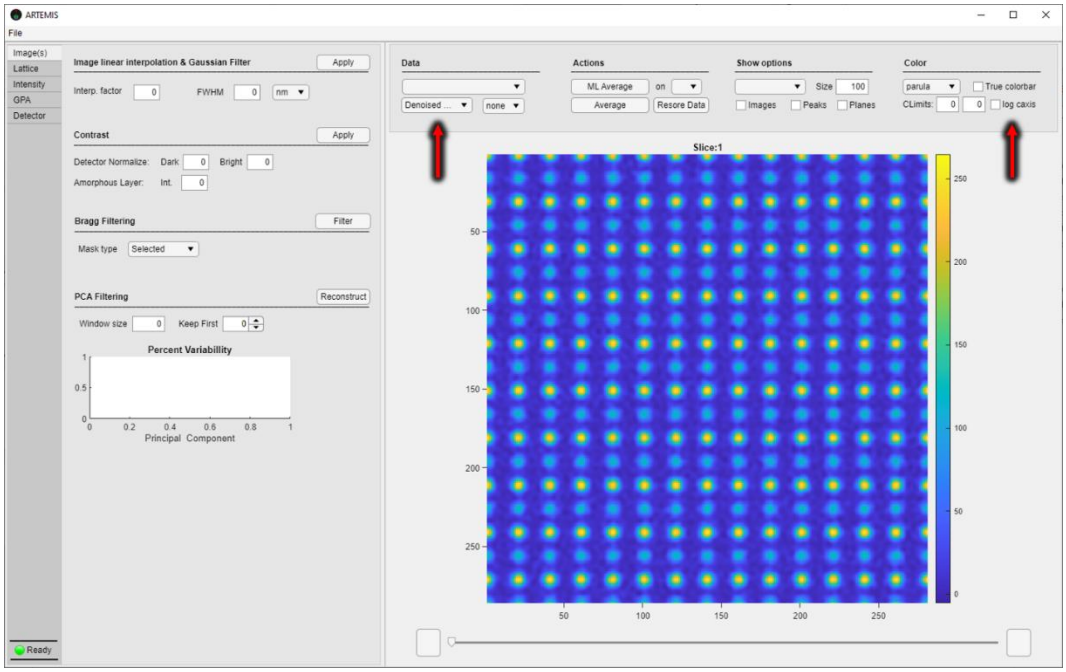
1. From the Data & Figure options extend the Data categories drop down and click *"FT images"*. FT images are displayed in the Figure section. Tick *"log caxis"* to make the image more comprehensive.



2. Next, we opt to create a low pass filter to isolate the spot's area. Right click on the figure, go to *"ROI" -> "Circle"*. Left click and drag to place your mask. We want to isolate the masked area so expand Mask type drop down and click *"Selected"*. Click *"Filter"* to perform inverse Fourier transformation of the image and complete the Bragg filtering process.

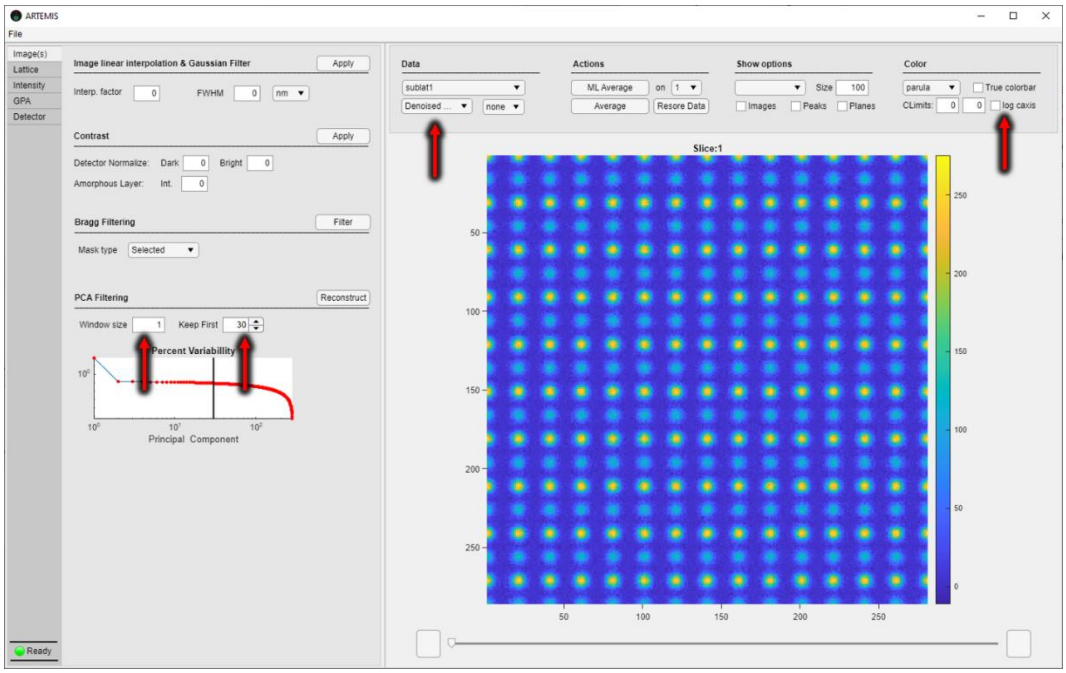


3. Access *“Denoised Images”* from Data categories drop down and untick *“log axis”* to visualize the resultant image.



PCA Filtering:

1. Select a “window size” of 1 to keep the calculations simple and fast. In this case, 30 significant PCA components are adequate to denoise the image and preserve the shape of the atomic columns, so use this value to *“Keep First”* field and hit *“Reconstruct”*. If this procedure performed after the Bragg filtering, ARTEMIS will know that the raw case image is already denoised via Bragg filtering and a warning will appear asking if you would like to restart (and proceed with the PCA filtering on the raw case image) or continue denoising (apply the PCA filtering over the Bragg filtered image). Here we proceed with *“Restart”* so that the PCA filtering effectiveness on the raw image is clear.



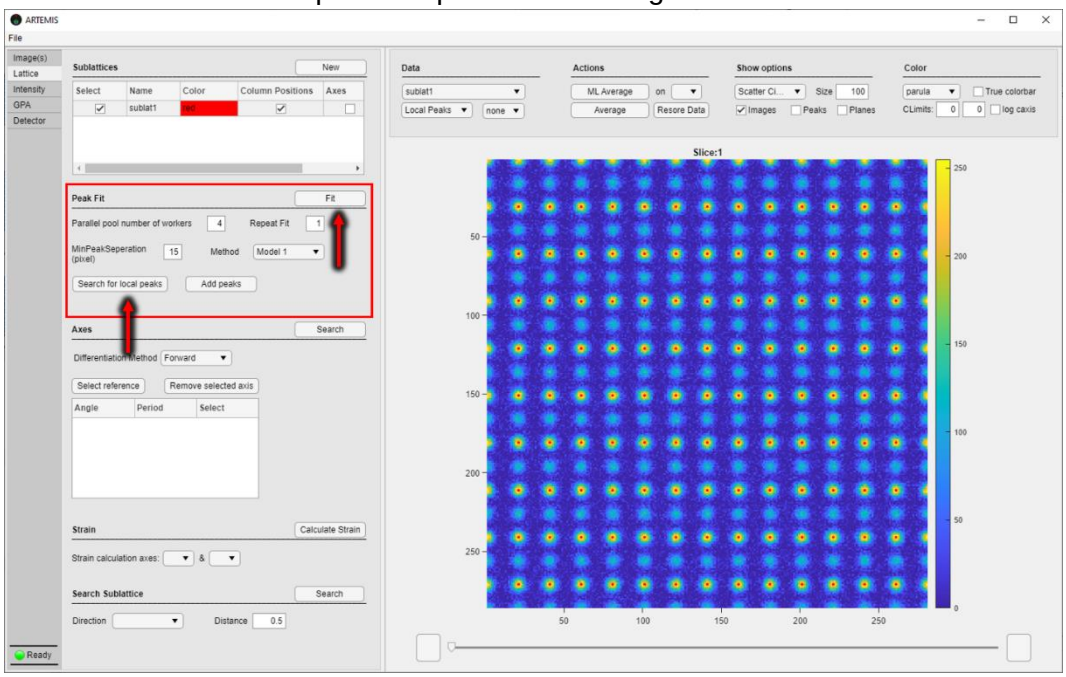
2. The *Percent Variability Plot* shows principal components significance and should be used as an indicator to decide the number of significant components to keep for the image reconstruction.

Lattice construction

Next, we try to identify the sublattices present on the image and deduce the atomic distances, angles, search for symmetries and group atoms to atomic planes. Change to *"Lattice"* tab to continue.

Peak Finding

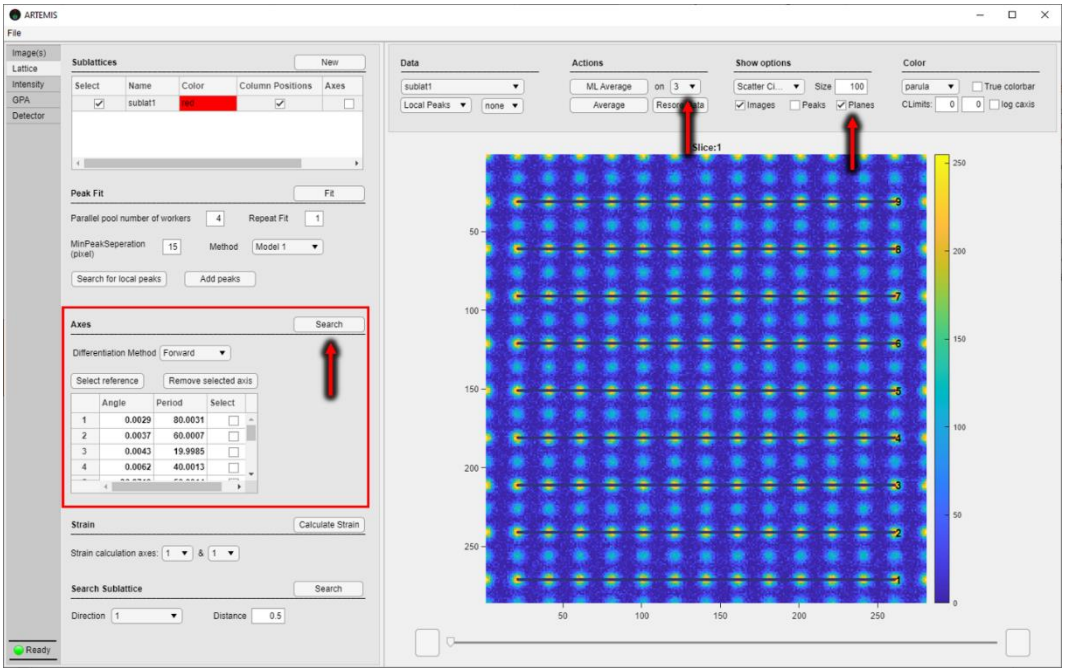
1. As the image consists of two sublattices, the aim here is to isolate, identify and analyze those separately. Click new to initiate a sublattice object. Make any changes to the name and/or color of the sublattice (see 2.3) at any time. Insert the number of *"Parallel pool number of workers"* your computer can spawn (for a pc with 4 physical CPU cores, write 4 here). Fill a *"MinPeakSeperation"* value and click *"Search for local Peaks"*. Repeat the search for different values of *"MinPeakSeperation"* (go for 5, 10 and 15). A value of 15 works fine at isolating one of the sublattices. We now can proceed with the actual intensity gaussian fitting. As the columns in this case are clearly round sized without overlapping, set *"Repeat Fit"* to 1 and choose the symmetric gaussian model (Expand *"Method"* dropdown and pick *"Model 1"*). Click *"Fit"* to start the CPU pool and perform the fittings.



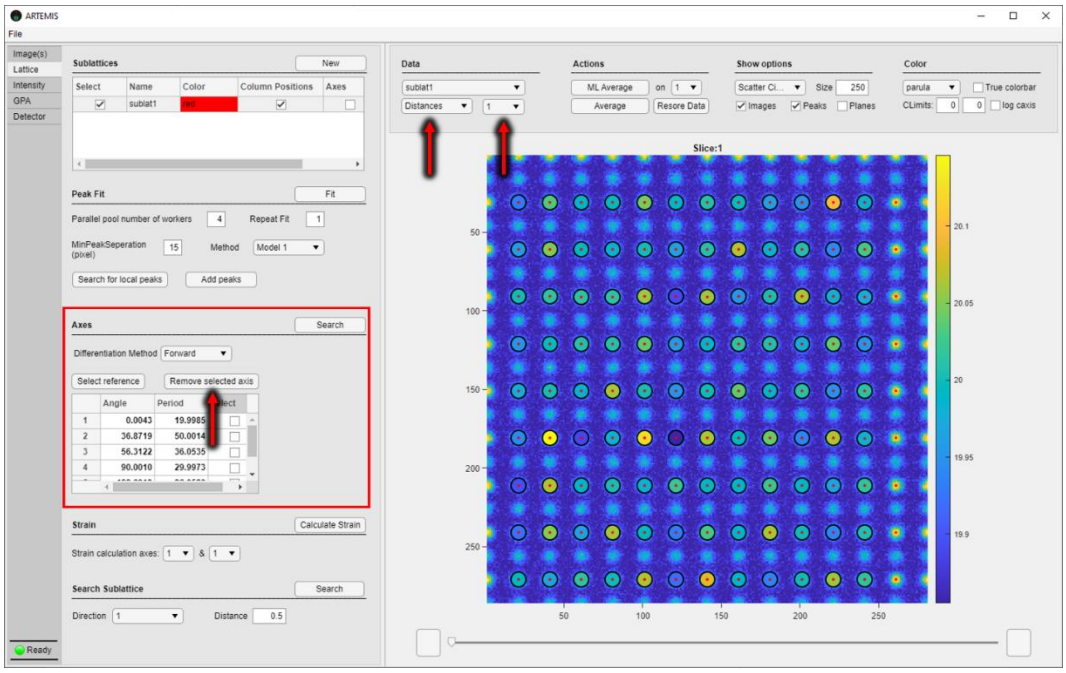
Axes

Here we identify axes of layered atomic columns, group them, calculate the interatomic distances, angles and bond length variations.

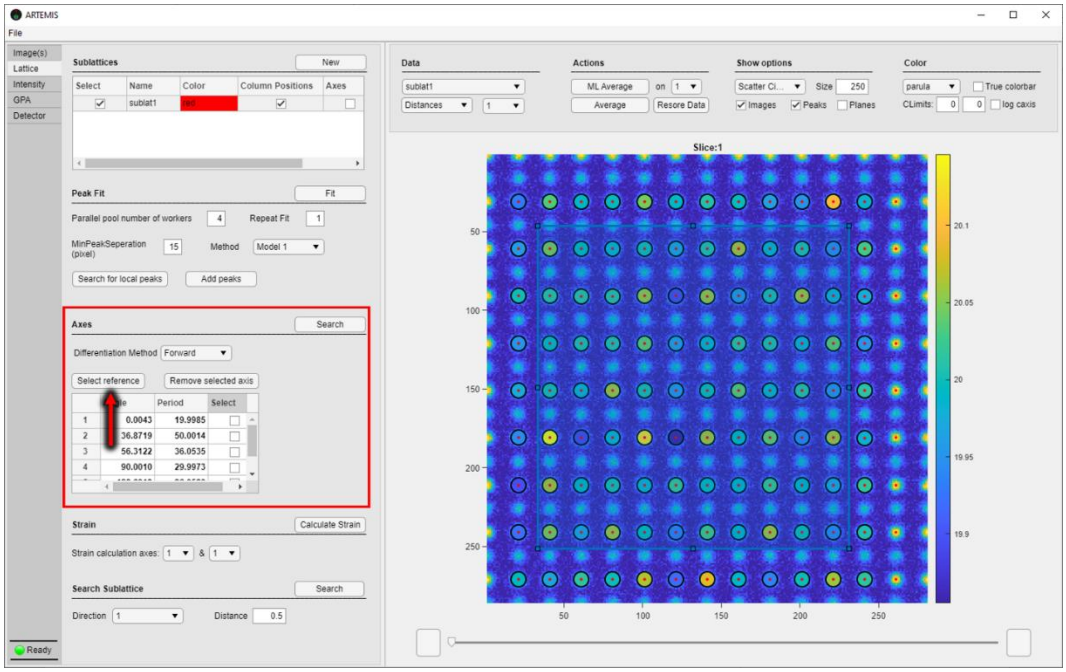
Select a differentiation method and click *"Search"*. To visually inspect the acquired layered column grouping, and the direction, tick *"Planes"* from Show options and expand Actions dropdown to choose a desired direction.



Remove any misidentified direction by selecting the from the Axes information table and clicking *"Remove selected axis"*. At this point, the atomic columns are grouped into planes, and the atomic distances and angles are calculated in the background. To access and visualize distances and angles data expand Data category drop down and select *"Distances"* or *"Angles"* and go through directions from directions drop down. Configure Scattered data appearance from Show options.



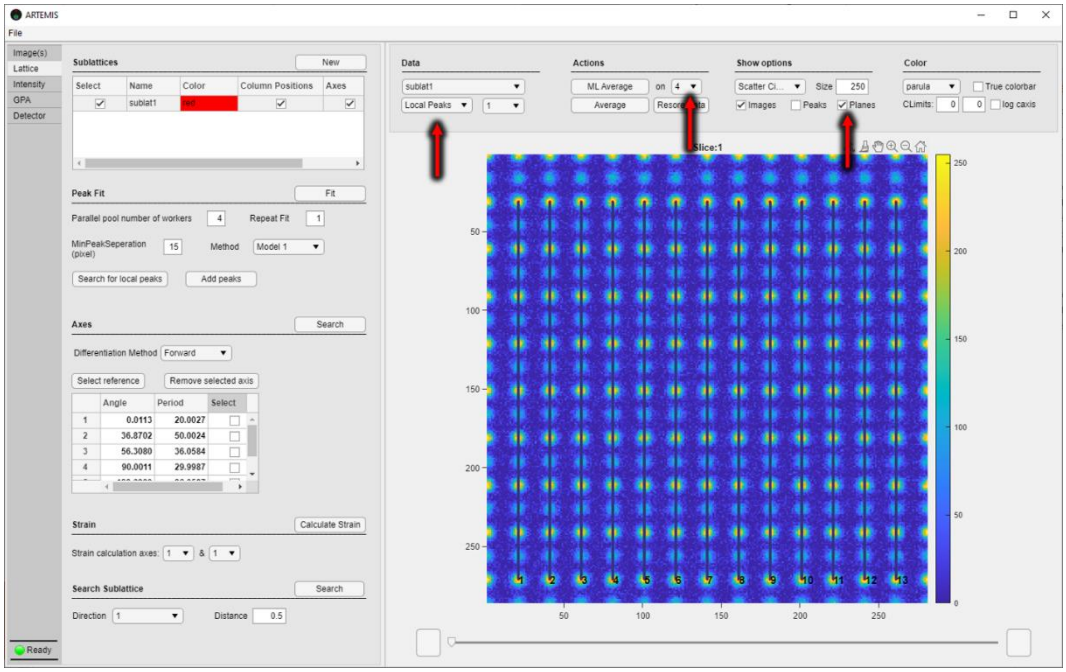
Next step is to calculate bond length variations. To do so we have to refine the angle and spacing period of the grouped columns with a reference area. Here the projected crystal is designed to be unstrained, as seen from Distances data (deviation ~ 0.1 pixels), and we can select as reference the whole image, otherwise the area of choice in this step is particularly important for lattices containing strained areas. To highlight a reference area, open and draw a ROI (right click figure), click *"Select reference"* to update the Axes information table and perform the bond length variation calculations.



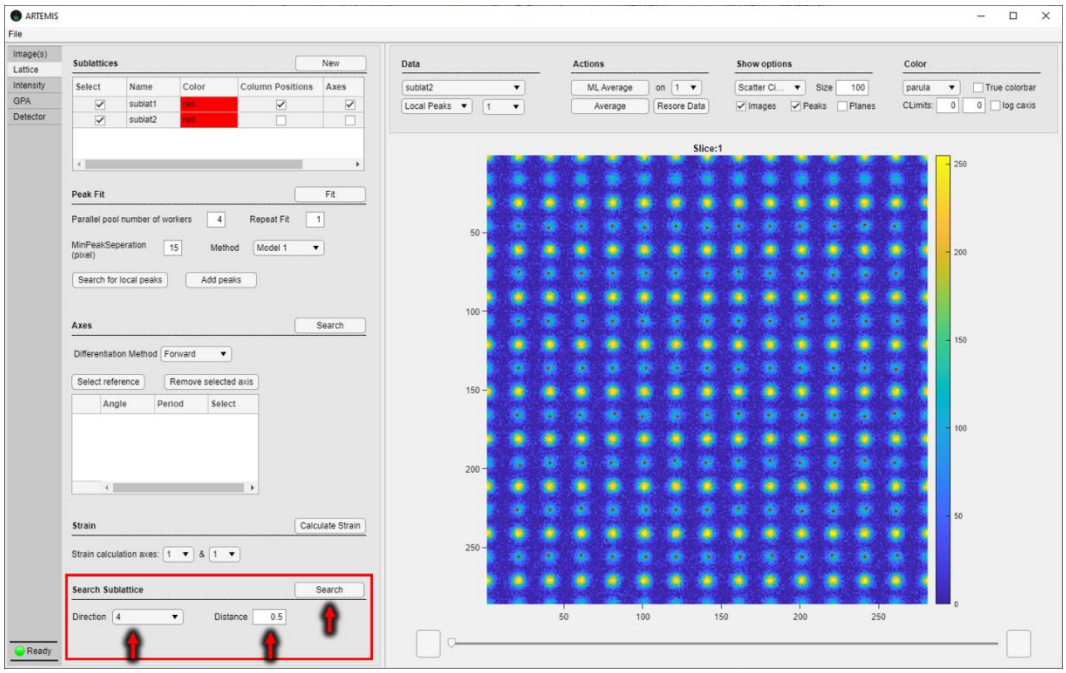
Visualize data from Data category drop down by selecting *"Bond length variation"*.

Search Sublattice

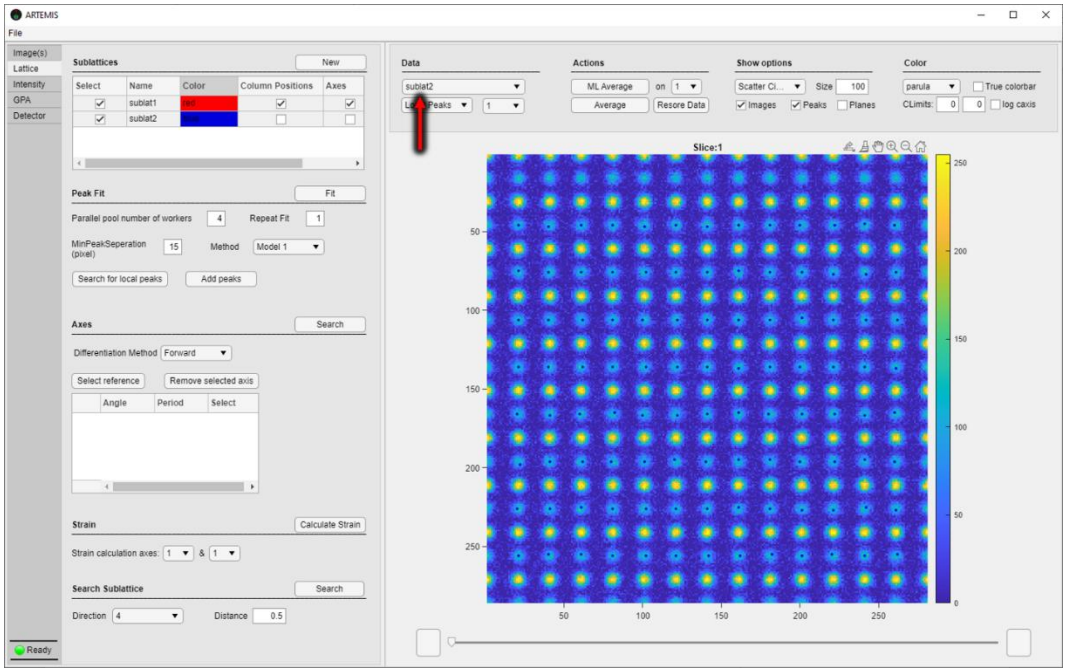
Our aim here is to locate and capture the second superimposed sublattice. To do so we need to select an appropriate direction that overlaps with the unidentified columns. Visually inspect different directions to locate a suitable one.



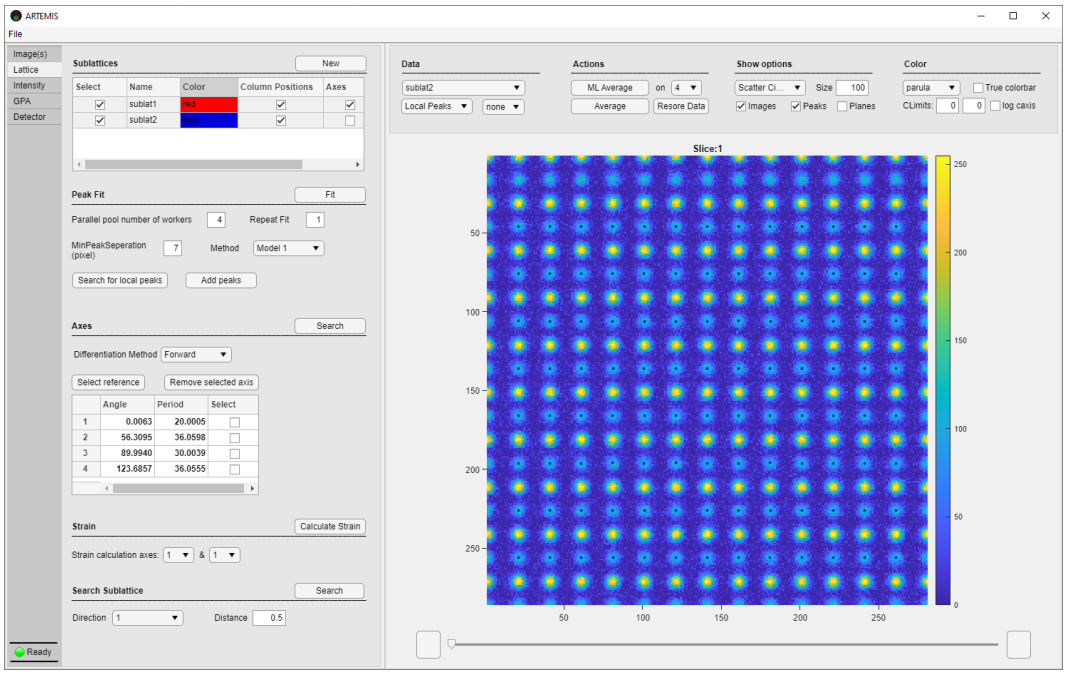
Direction 4 from Axes table passes through most columns of the second sublattice. To identify those, we have now to input a rough estimation of their relative distance from columns of sublat1 in this direction. A good value is 0.5 (*"Distance"* edit field). To search for the columns, click *"Search"*.



A second sublattice object has been generated and the newly identified columns are assigned to it. The whole *"Lattice Tab"* is refreshed to the input options of the user on the current sublattice. To switch between different sublattices simply select the target sublattice name from Data dropdown as shown below.

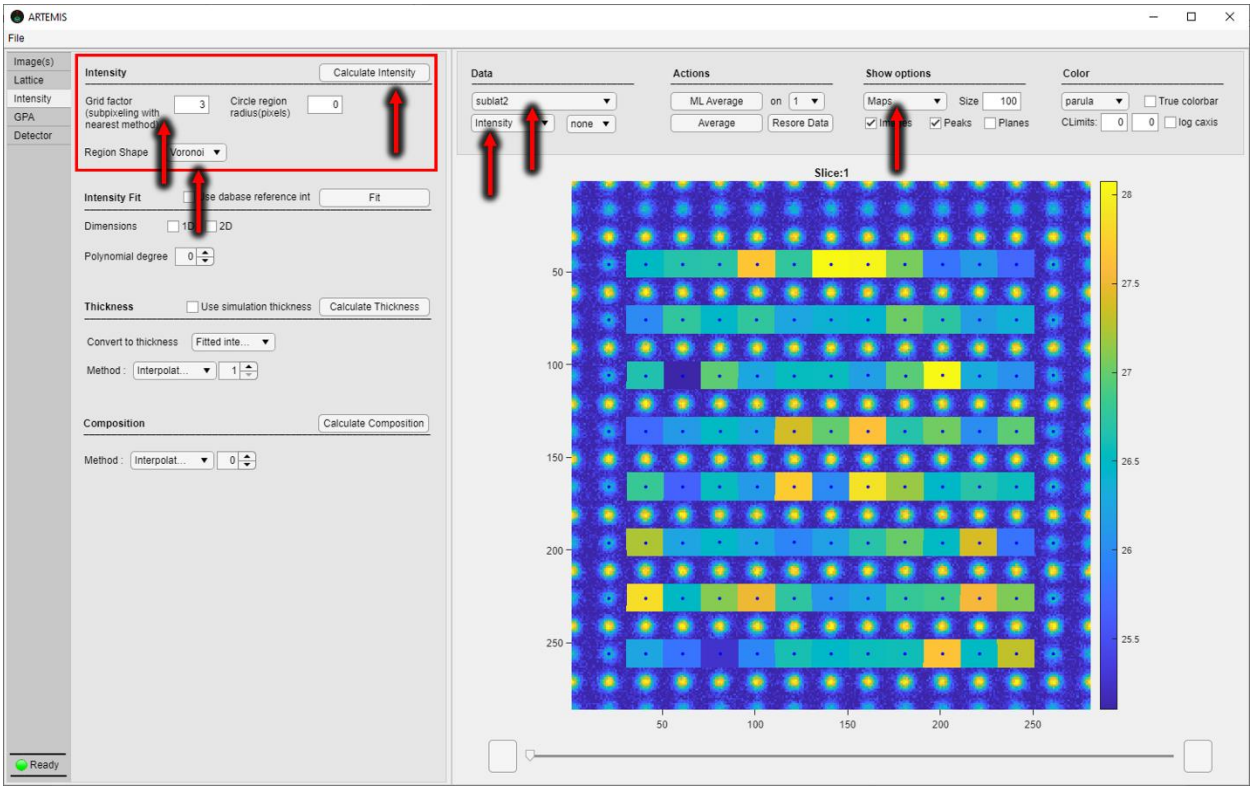


To get all the structural information we have to repeat the same steps followed for the first sublattice. We summarize the steps here: 1. Perform peak fit (Model 1, *"Repeat Fit"* = 1) 2. Search for axes, 3. remove misidentified, 4. select a reference region.



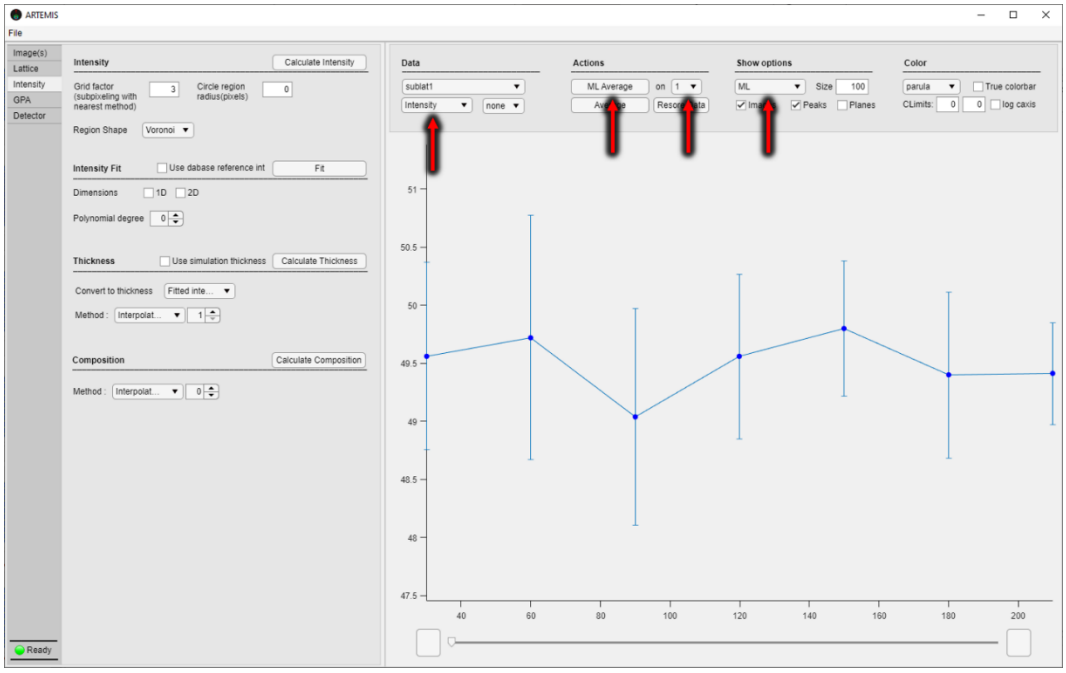
Column Intensities

On to intensity tab, we integrate column intensities. Select *"Voronoi"* in *"Region Shape"* and provide a grid factor (grid factor dictates the resolution of the integration areas and does not affect the intensity of the image, so higher values return more accurate integration results at the expense of computational power). Click *"Calculate Intensity"* to perform the calculations. This process generates Voronoi Maps with columns intensity values. Access and visualize intensity data, for the respective sublattice, from Data and Figure options panel.



Data

You may obtain average value along monolayers for specific regions and collect data from any of the populated data categories. A simple exercise is to average intensities from sublat1 across the horizontal axes. To do so, select the direction with angle ~ 0 (here is direction 1) from the Actions options and click *"ML Average"*. Note that Show options drop down has changed to *"ML"*.



Save intensity data by right-clicking the figure *"Save"* -> *"Data"*. Same logic for saving data applies to all generated data (2D, ML Averaged and Averaged). First you visualize them and then right click and save data and/or figures (See 2.8).

4. Acknowledgments

[1]A. Rosenauer, K. Gries, K. Muller, A. Pretorius, M. Schowalter, A. Avramescu, K. Engl, and S. Lutgen, Ultramicroscopy **109**, 1171 (2009).

[2] Mukherjee, D., Miao, L., Stone, G. *et al., Advanced Structural and Chemical Imaging* **6,**1 (2020).

[3] E Snoeck, B Warot, H Ardhuin, A Rocher, M.J Casanove, R Kilaas, M.J Hÿtch, Thin Solid Films **319**, 157 (1998).

[4] Martin J. Hÿtch, Microsc. Microanal. Microstruct. **8**, 41 (1997).