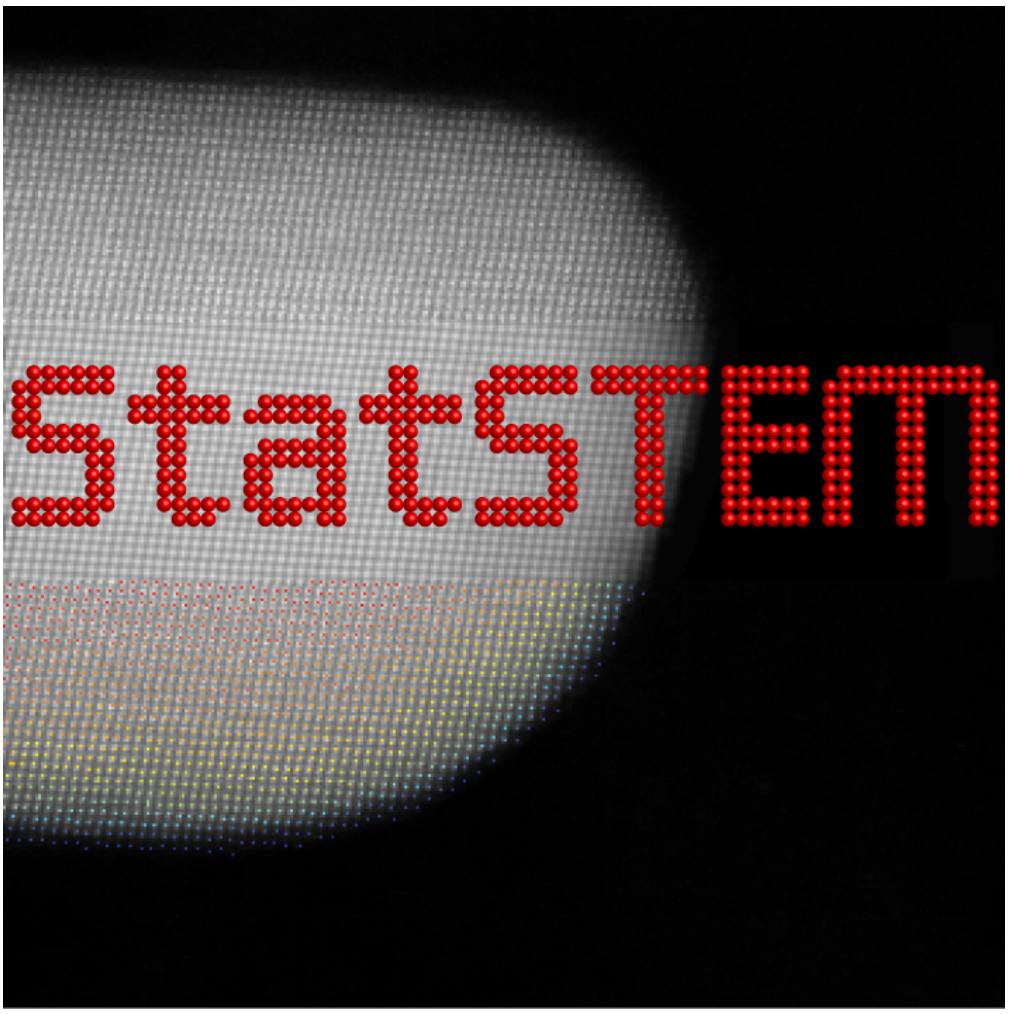

StatSTEM
User Guide
For StatSTEM v2.1

Contents

1 Getting started

The program can be downloaded from the StatSTEM website. One can choose between running StatSTEM from MATLAB or installing a compiled version (executable).

1.1 MATLAB 2015

Once all StatSTEM files are copied to your computer, start MATLAB and go to the folder containing the StatSTEM files. The program can be started by running the main script *StatSTEM.m*.

1.2 Compiled version for Windows

By installing the compiled version, a MATLAB compiler will be installed together with the program. After the installation, the program can be started by running the executable *StatSTEM.exe*. Be aware that during the first run, Windows can ask for some permission, which should be granted.

2 A first look

In StatSTEM different options are offered for quantifying your images. In general, the quantification consists of 3 steps:

- Preparation
- Modelling
- Analysis

For each step a tab is made in the left tab panel. On the right, a tab panel is made in which loaded images and generated results can be shown. In this panel, multiple image can be loaded. At the bottom left, two buttons enable one to load and store files. A message panel will inform you when computations are finished, changes are made, or things went wrong. A progress bar indicates the status of the computation. As computations are often iterative, the progress bar cannot be used for estimating how much time computations will take.

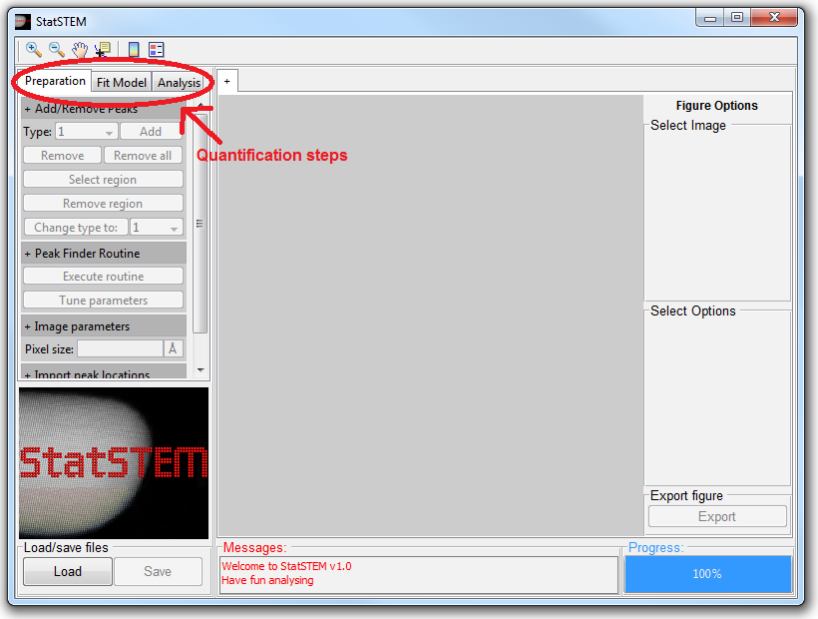

Figure 2.1: In the left tab panel, a tab can be selected for each quantification step

3 Load and save files

Once StatSTEM is started, you can load your image into the program by either clicking on the *Load* button below or the addition tab in the right panel. StatSTEM supports files having the MATLAB (*.mat*) and text (*.txt*) format. After loading the image, a dialogue will appear asking for the pixel size. Images and the corresponding parameters can be saved in the MATLAB format by clicking on the *Save* button next to the *Load* button. In the saved file, all variables are stored in a structured format, see Appendix A for more details.

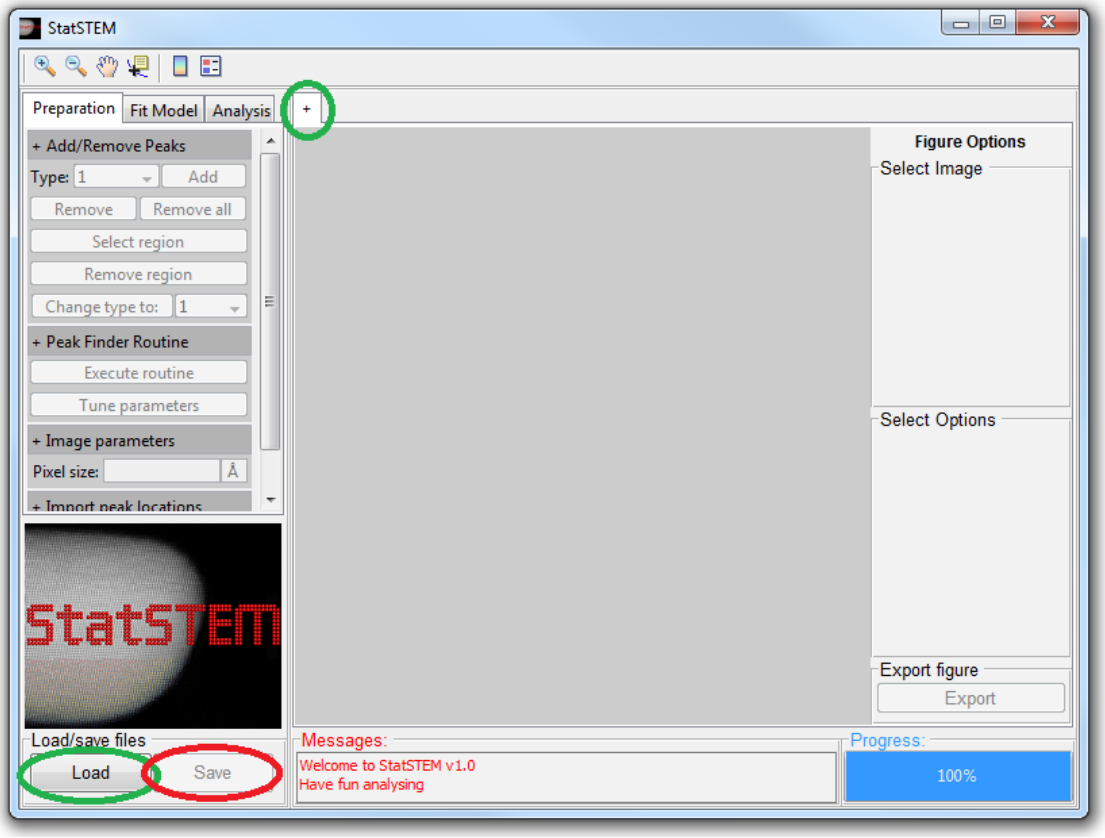

Figure 3.1: Buttons for loading and saving files indicated by respectively the green and red circle(s)

4 Figure options

StatSTEM offers different options to show the images and the parameters you want. For each loaded file, the *Figure Options* panel allows one to select different images (*Select Image*) and show different parameters in the images (*Select Options*). By hitting the *Export* button, the shown image will be opened in a new MATLAB figure which allows one to save the image in the desired format.

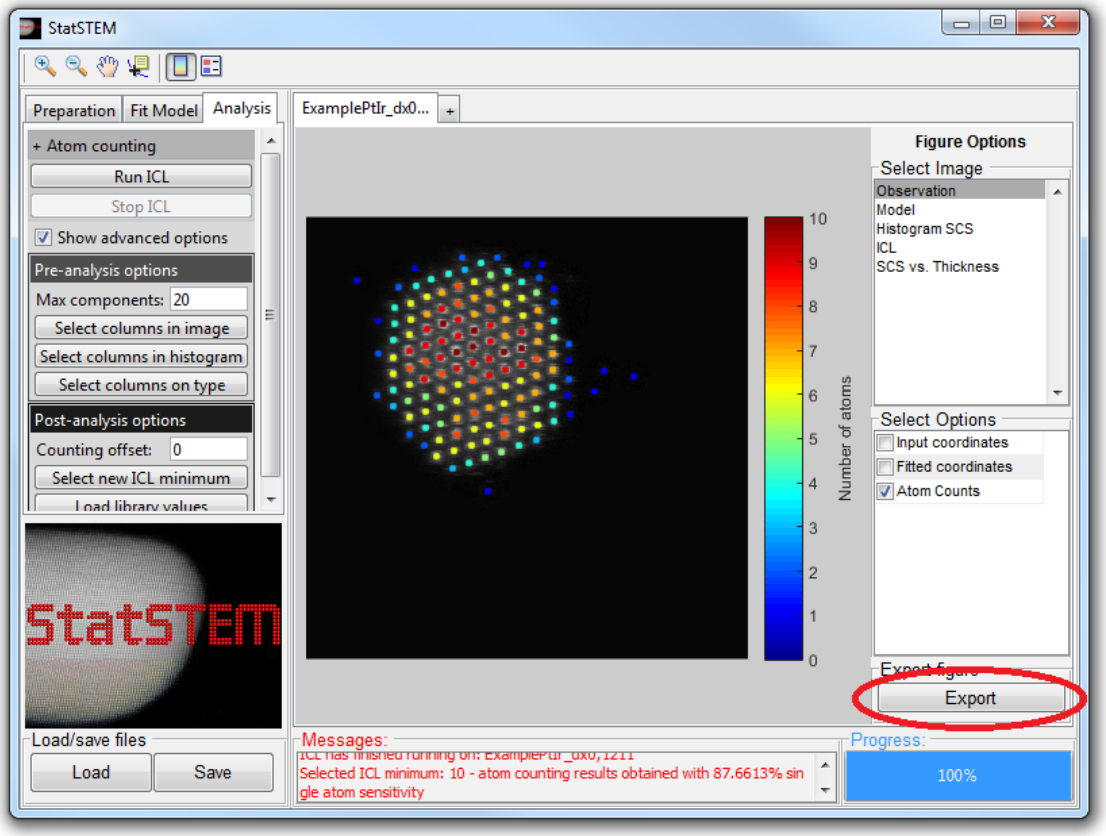

Figure 4.1: The right Figure Optional panel of StatSTEM allows one to show the an image with the desired parameters. The export button, indicated by the red circle, can be used to open the image in a new MATLAB figure.

5 Preparation

Before fitting a Gaussian model to your loaded experimental images, starting coordinates for the atomic column positions have to be defined (see Fig. 5.1). For this step, different options are available on the preparation panel.

5.1 Add/Remove peaks

In this section, different routines are available to define, remove or, change starting coordinates manually. By pressing *escape* the routines will be aborted.

5.1.1 Type

By using this option, starting coordinates can be labelled. This is particularly useful when your image consists of different column types, for example columns containing different elements. When adding column types, StatSTEM will automatically label different column types by numbers. In the fitting procedure, labelled columns may be given a special treatment. This will be explained in Section 6.2.

5.1.2 Add

This option allows one to add manually starting coordinates in the image. Each added starting coordinate will have the selected column type label (indicated next to the button). By pressing *escape* or clicking outside the image, the routine will be aborted.

5.1.3 Remove

Starting coordinates, of any column type, can be manually removed by this option. By clicking in the image, the starting coordinate closest to the cursor will be removed. By pressing *escape* or clicking outside the image, the routine will be aborted.

5.1.4 Remove all

All starting coordinates will be removed.

5.1.5 Select region

Select a region in the image in which the starting coordinates should be maintained. Outside the selected region, all starting coordinates will be removed. In this routine, the user defines the corner points

of the selected area one by one. A right click will connect the last defined point with the starting point. By pressing *escape*, the routine will be aborted.

5.1.6 Remove region

Select a region in the image in which the starting coordinates should be removed. In this routine, the user defines the corner points of the selected area one by one. A right click will connect the last defined point with the starting point. By pressing *escape*, the routine will be aborted.

5.1.7 Change type to

By using this option, a region in the image can be selected in which the label should be changed to the selected column type label (indicated next to the button). In this routine, the user defines the corner points of the selected area one by one. A right click will connect the last defined point with the starting point. By pressing *escape*, the routine will be aborted.

5.2 Peak Finder Routine

As it is much work to define each starting coordinate manually, a peak finder routine is added which searches for local maxima in the image.

5.2.1 Execute routine

Runs the peak finder routine.

5.2.2 Tune parameters

This option will open a new window in which the parameters of the peak finder routine can be tuned. As noise is often present in experimental images, filters are used for smoothing the image. In total 3 different filters can be selected, a *gaussian*, *averaging* and a *disk* filer. Furthermore a threshold value can be defined, above which the pixel values should lie. In this window, the routine can be tested for the selected parameters. Once appropriate parameters are selected, the parameters can be stored in StatSTEM by hitting the *Store and close* button. Next, the routine can be executed in StatSTEM.

5.3 Image parameters

In this section, the image pixel size can be changed. By changing the pixel size, starting coordinates and fitted parameters will be rescaled.

5.4 Import peak locations

In this section, a MATLAB (*.mat*) or text (*.txt*) file can be loaded containing starting coordinates. The starting coordinates have to be defined in Ångström in a (n $\times$ 2) vector containing the x- and y-coordinates. Furthermore, another StatSTEM file can be loaded to use its starting or fitted coordinates.

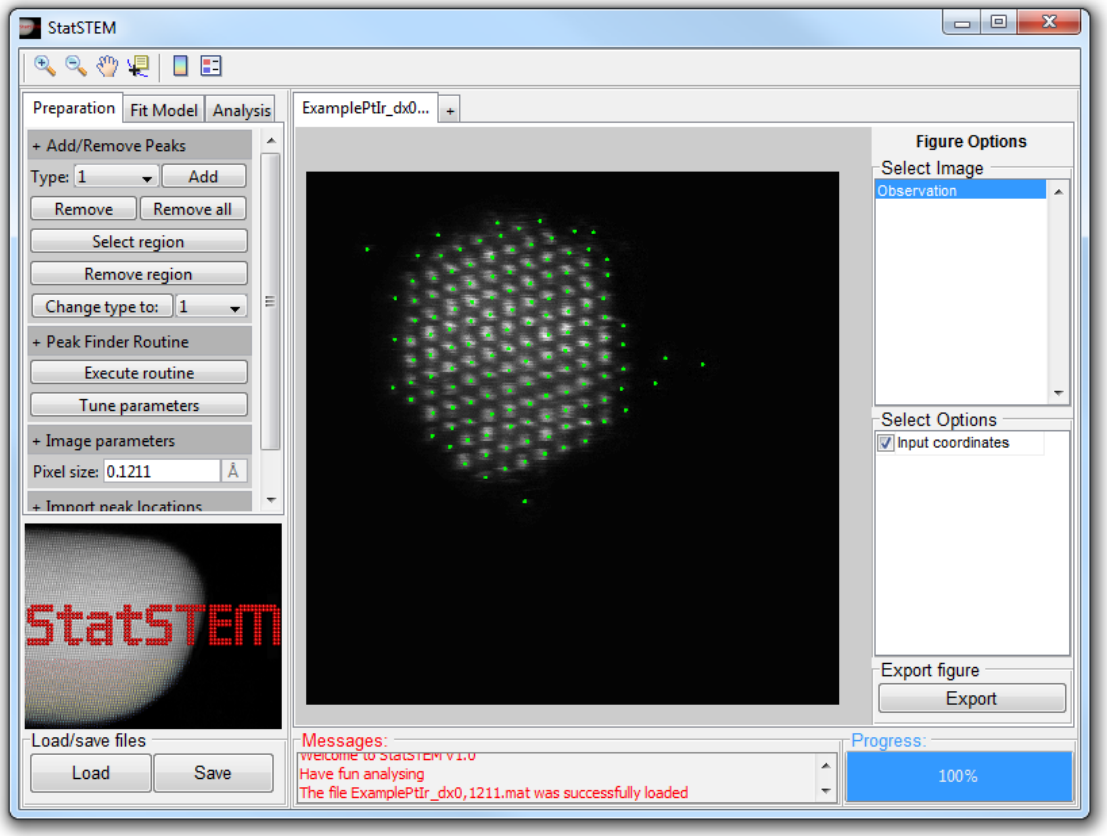

Figure 5.1: An experimental HAADF STEM image of a Pt/Ir sample with defined starting coordinates.

6 Fit Model

From this panel, a Gaussian model can be fitted to the image. In this model, each atomic column will be modelled as a Gaussian peak. Detailed information about the fitting routine can be found in Ref. [1]. The options in this panel are only available if at minimum one starting coordinate is defined. The model can be fitted by hitting the *Run fitting routine* button. After the procedure, the fitted Gaussian model with the fitted coordinates will be shown (see Fig. 6.1). The fitting procedure can be aborted by hitting the *Abort fitting routine*. This button may not always work as MATLAB sometimes cancels or postpones functions during a computation. A second (or third) hit usually works. Once the procedure will be aborted, StatSTEM will show a message.

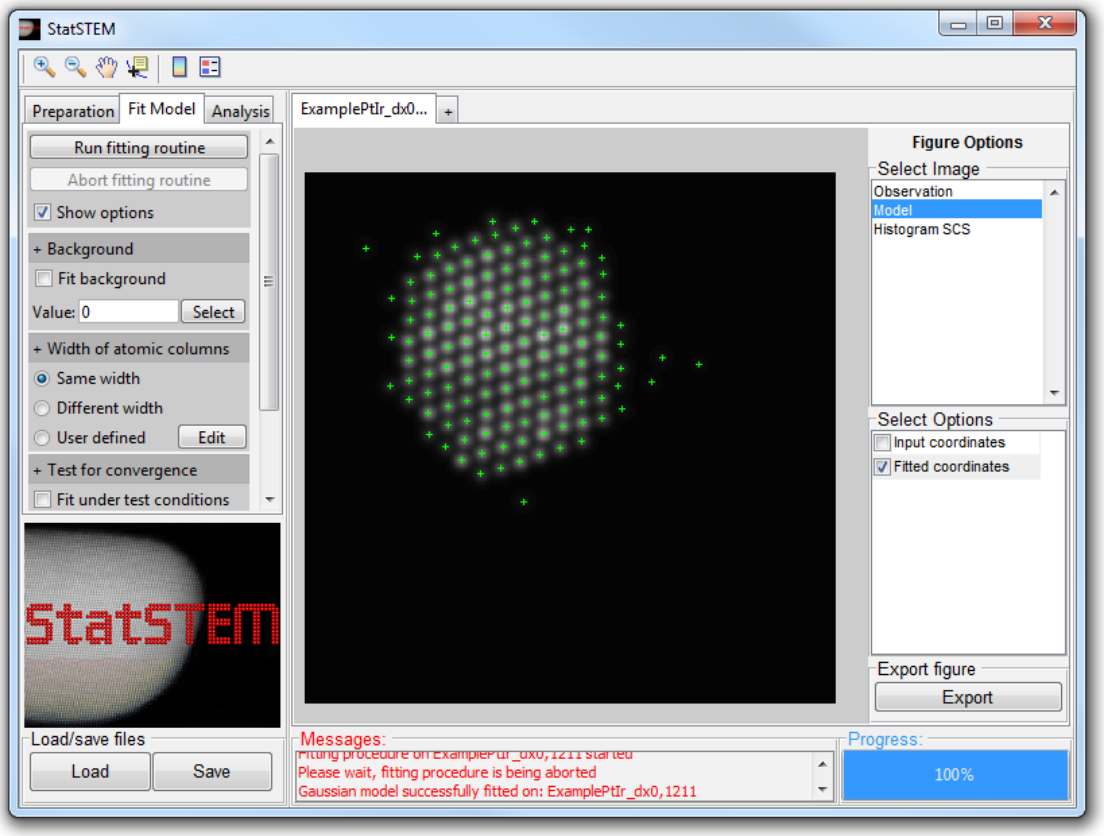

Figure 6.1: The fitted model of an experimental HAADF STEM image of a Pt/Ir sample with the fitted coordinates of the atomic columns.

Once the model is fitted, the fitted coordinates of the atomic columns and the total intensity of electron scattered by the atomic columns, the so-called scattering cross-sections, are calculated. These scattering cross-sections can be seen in the histogram of scattering cross-sections (see Fig. 6.2).

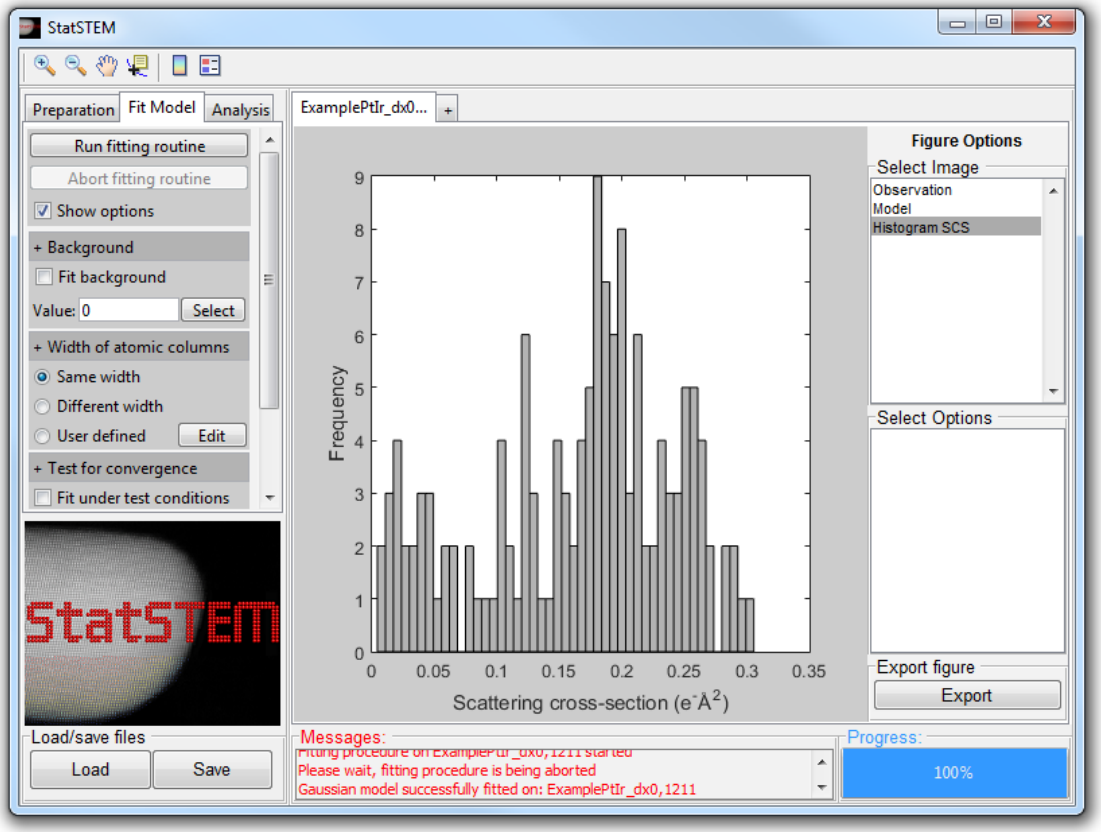

Figure 6.2: The histogram of scattering cross-sections of the atomic columns in an experimental HAADF STEM image of a Pt/Ir sample.

The modelling panel also shows different options of the fitting procedure that may be defined (by using the *Show options* button). The following sections discuss the available options.

6.1 Background

In this options, the user can choose to fit a constant background. If no background will be fitted, a constant value may be given. By using the *Select* button, the mean value of a selected region in the image can be determined.

6.2 Width of atomic columns

In the fitting procedure, one may choose to fit Gaussian peaks to all columns having the same or a different width. In the *Same width* option, the estimated Gaussian peaks will have the same width for columns of the same atom type (as can be labelled in the preparation step discussed in Chapter 5). In the *Different width* option, a different width is estimated for each Gaussian peak of an atomic column. In the *User defined* option, the user may define a width for each column type which will be used to find the atomic column positions. Usually, the *Same width* option is used as this is computationally less demanding and gives good results. Only when no information is known about the structure under study (the present elements are unknown), the *Different width* option is advisable. More information can be found in Ref. [1].

6.3 Test for convergence

This option may be used for testing the correctness of the starting coordinates and fitting parameters. The number of iterations will be limited to 4. If all starting coordinates are converging in the correct manner, the *Export fitted coordinates* button may be used to make the fitted coordinates the new starting coordinates.

6.4 Parallel computing

For improving computational speed, the fitting procedure uses parallel computing in which the calculations are divided over the different CPU cores of your computer. The number of CPU cores used for parallel computing may be reduced to lower the CPU usage during the fitting procedure (be aware that total calculation time will most likely increase).

7 Analysis

In this panel, the fitted parameters of the model may be used for analysis. In version 2.0, the statistics based method for atom counting is implemented in StatSTEM (see also Refs. [2, 3, 4]). Also atom counts can be obtained from a comparison with a simulated library. Next to atom counting, strain mapping is available. More options may become available in future versions.

7.1 Atom counting: statistical method

In HAADF STEM imaging, a statistics based method is developed to count the number of atoms based on the scattered cross-sections (the total intensities of electrons scattered by the atomic columns), which increase monotonically with thickness [5]. Be aware that this method is only reliable when only one column type is present. In this statistics based method, the scattering cross-sections are presented in a histogram. Owing to a combination of experimental noise an residual instabilities, broadened - rather than discrete - components are observed in such a histogram. Therefore, these results cannot directly be interpreted in terms of number of atoms. By evaluation of the so-called integration classification likelihood (ICL) criterion in combination with Gaussian mixture model estimation, the number of components and their respective locations can be found. From the estimated locations of the components, the number of atoms can be quantified. More information can be found in Refs. [2, 3, 4]. For evaluating the ICL different options can be used.

7.1.1 Pre-analysis

Maximum number of components

In order to evaluate the ICL, an upper limit on the number of components must be given. Up to a given number of components, Gaussian mixture models are fitted to the histogram of scattering cross-sections and the ICL criterion is determined. A rough estimate for this upper limit can be obtained by using the shape of the particle under study.

Select columns in image

Select a region in the image in which all columns should be taken into account for atom counting. The columns outside this region are neglected. By pressing *escape*, the routine will be aborted.

Select columns in histogram
This option allows one to exclude outliers in the histogram of scattering cross-sections. First the lower limit must be defined, then the upper limit. By pressing *escape*, the routine will be aborted.

Select columns on type
By using this option, columns can be selected based on the column type labels given in the preparations step (see also Chapter 5).

7.1.2 Post-analysis

After selecting atomic columns and the maximum number of components, Gaussian mixture model may be fitted to the histogram of scattering cross-sections by using the *Run ICL* button. For each number of components, the ICL-criterion is determined and displayed. For atom counting, one searches for local minima in the ICL-curve (see for example Fig. 7.1). After the computation, StatSTEM asks the user to select a local minimum. After a local minimum is selected, the experimental image with the corresponding

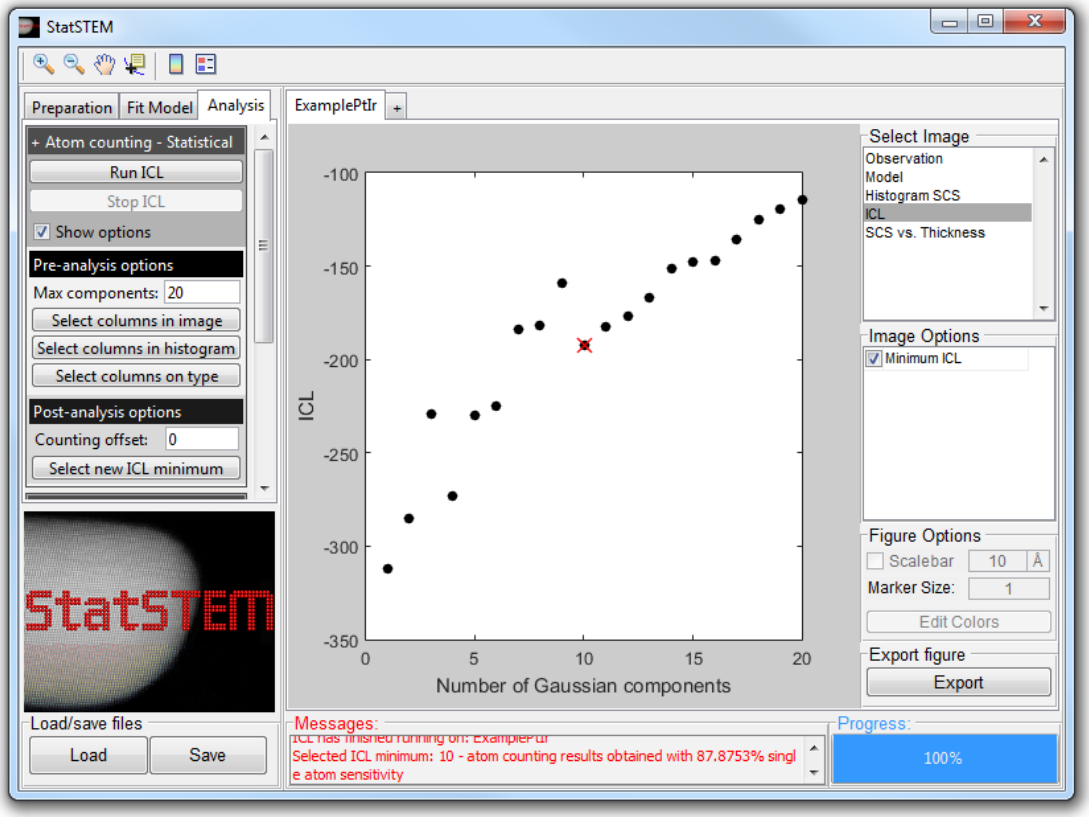

Figure 7.1: The ICL for an experimental image. A local minimum appears at 10.

atom counts will be shown (see Fig. 7.2). The computation of the ICL values can be aborted by pressing the *Stop ICL* button. The button may not always work as MATLAB cancels many functions during a computation. A second (or third) hit usually works. Once the procedure will be aborted, StatSTEM will show a message asking whether the user wants to select a local minimum in the current ICL graph.

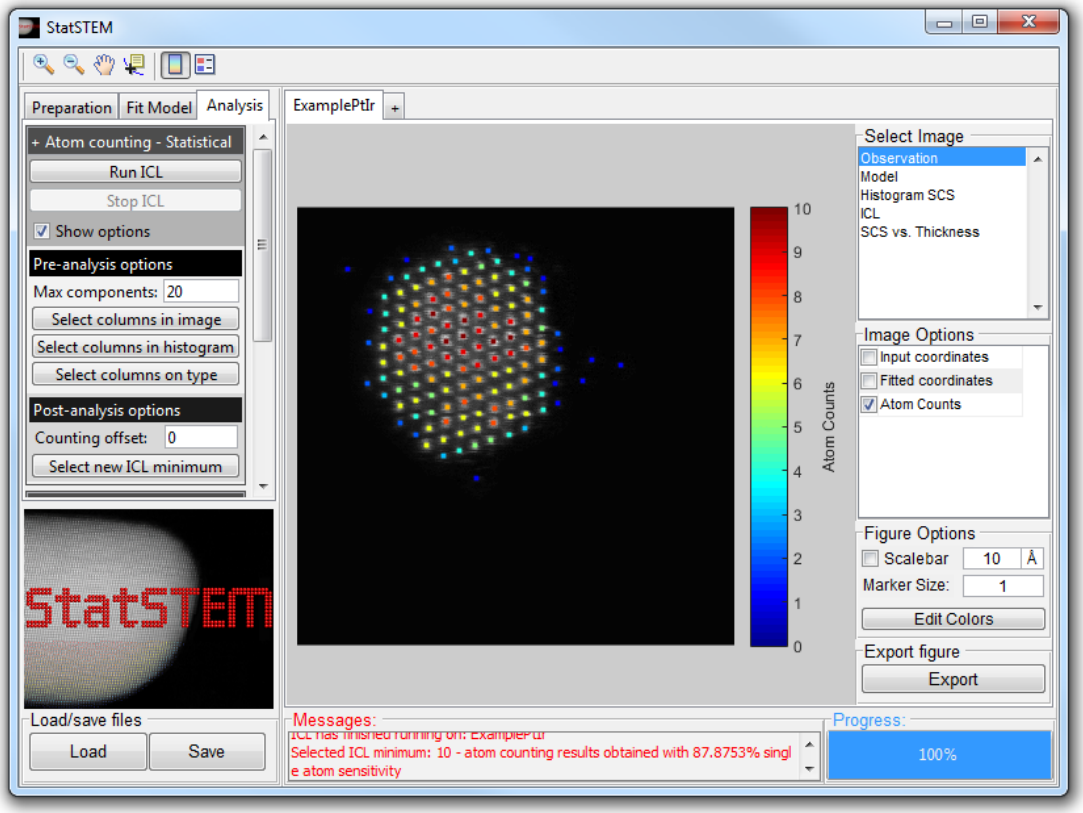

Figure 7.2: The atom counts of the different atomic columns in the experimental HAADF STEM image of a Pt/Ir particle.

Counting offset

Once a local minimum in the ICL curve is selected, counting results may be rescaled by giving the counts an offset. This is particularly useful for thick particles in which no thin columns are present.

Select new ICL minimum

With this option a new local minimum in the ICL curve can be selected.

7.2 Atom counting - Simulation based

Another method for doing atom counting is by comparing simulated library's of scattering cross-sections with the experimentally measured ones.

7.2.1 Load library values

A MATLAB (*.mat*) or text (*.txt*) file can be loaded with simulated values of the scattering cross-sections. The simulated values in function of column thickness must be stored in a (n $\times$ 1) vector. In the figure *SCS vs. Thickness*, the scattering cross-sections obtained by the statistics based method can be compared to the simulated values (see Fig. 7.4).

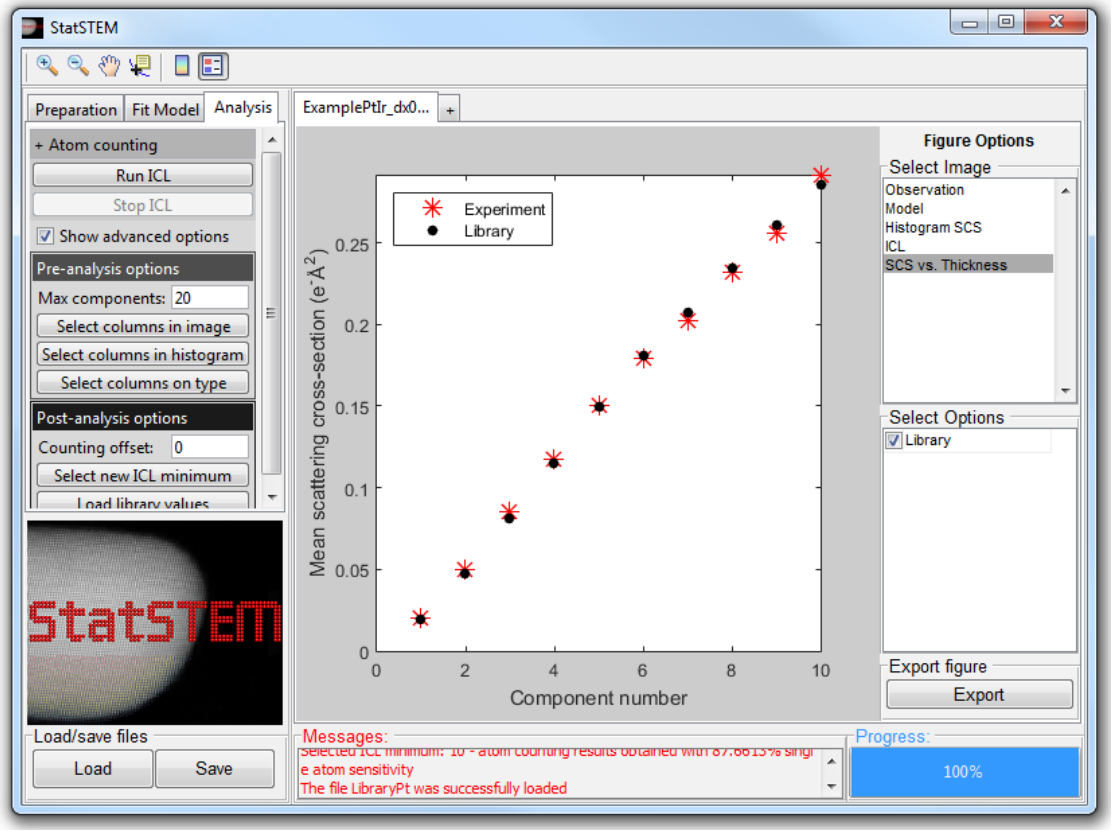

Figure 7.3: A comparison between the scattering cross-sections obtained by the statistics-based method (Experiment) and image simulations (Library).

7.2.2 Match with simulations

By clicking on this button, the measured scattering cross-sections from the fitted model will be compared to the loaded library of sim-

ulated scattering cross-sections. This direct comparison will give the counting results.

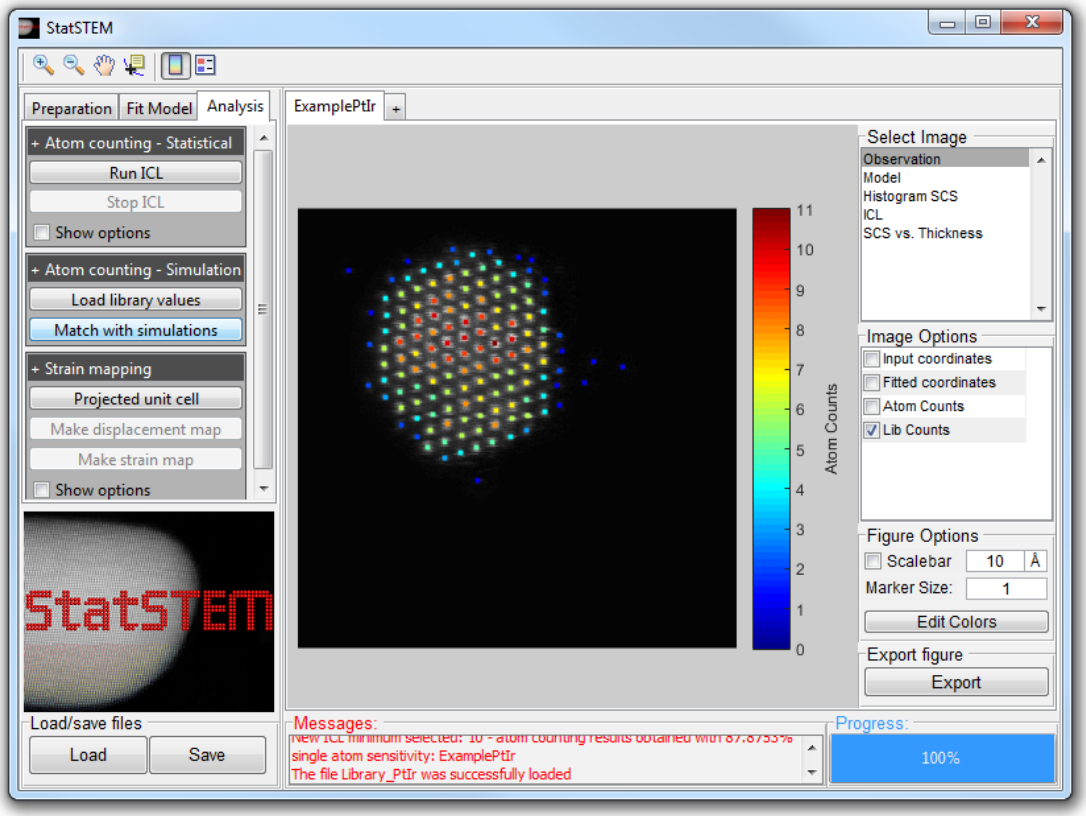

Figure 7.4: The atom counts of the different atomic columns in the experimental HAADF STEM image of a Pt/Ir particle obtained by a comparison with simulated library values of SCS.

7.3 Strain mapping

A new feature from version 2.0 is strain mapping. In this method, a displacement map of the atomic column with respect to their relaxed positions can be made. From the derivative of the displacement map, a strain map with atomic resolution can be made.

7.3.1 Projected unit cell

A necessary input for strain mapping is the location of the different column positions in a projected unit cell, shown in the Fig. 7.5. A database is added to StatSTEM for some common materials and viewing directions. Make sure that you fill in the correct pixel size of the image as the lattice parameters in the projected unit cell should be close to the experimental values.

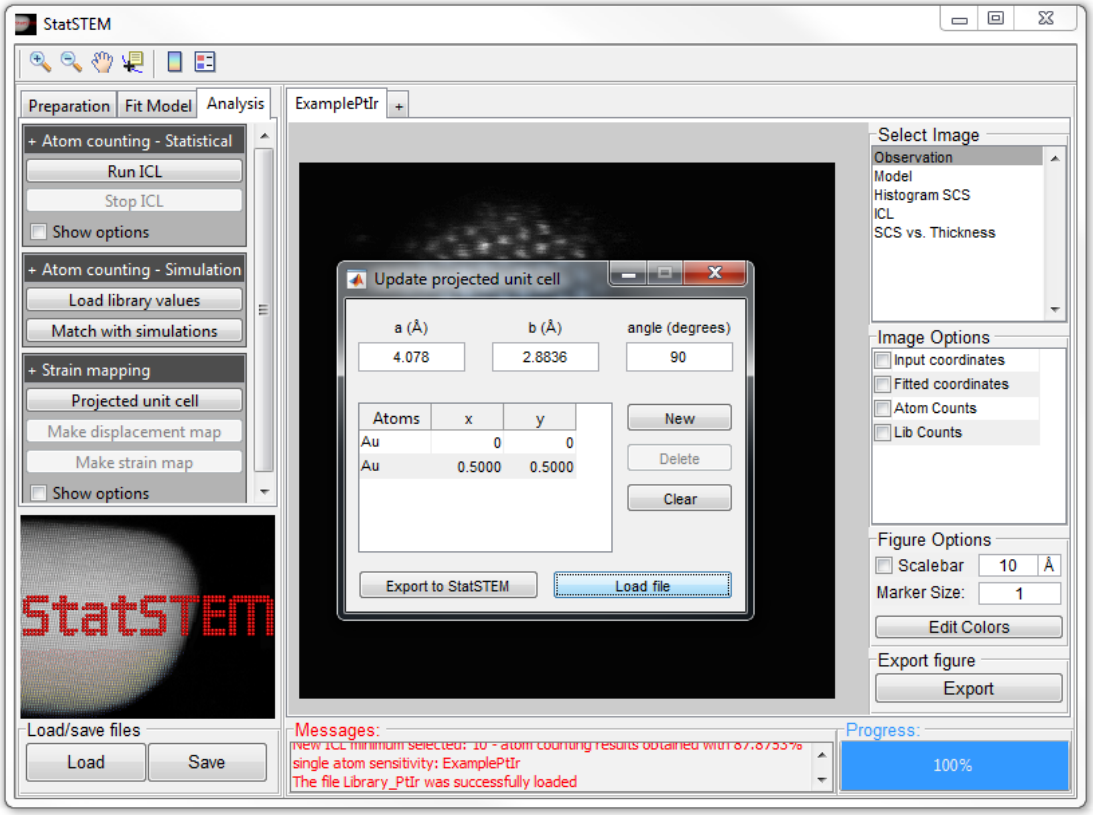

Figure 7.5: A project unit cell can be made by defining the projected lattice constants (a and b) and the relative positions of the atoms.

7.3.2 Make displacement map

By clicking on this button a displacement map will be made based on the projected unit cell that is given. For determining the direction

of lattice vectors and other variables, parameter can be specified in the advanced options part discussed in section 7.3.4. During the creation of the displacement map, indices of each column with respect to the a- and b-direction of the lattice will be generated. Each index indicates the distance with respect to a selected reference coordinates in number of unit cells in a- and b-direction. For each column type (for example one at (0,0) and (0.5,0.5)), indices will be generated in the output file.

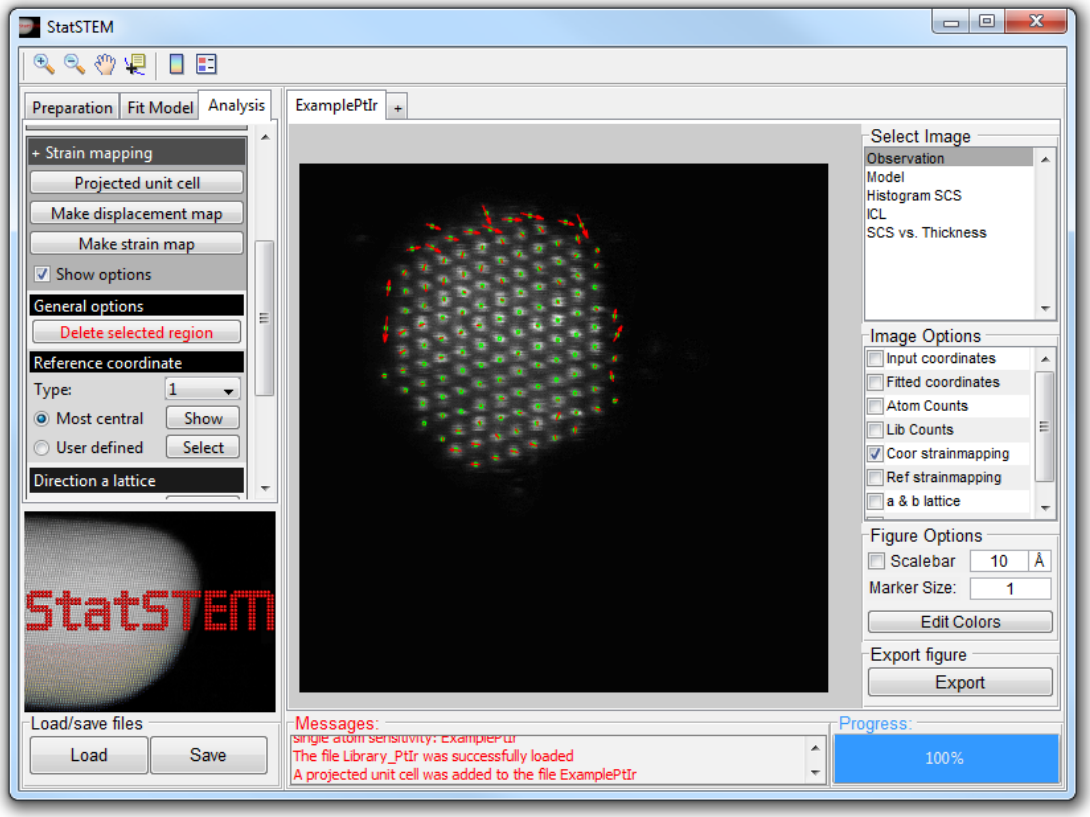

Figure 7.6: A displacement map of a Pt/Ir particle.

7.3.3 Make strain map

By clicking on this button, the ϵ_{xx}, ϵ_{xy}, ϵ_y and ω_{xy} strain maps will be generated. Here, the peak pair method discussed in [6] is used. An example is shown in Fig. 7.7.

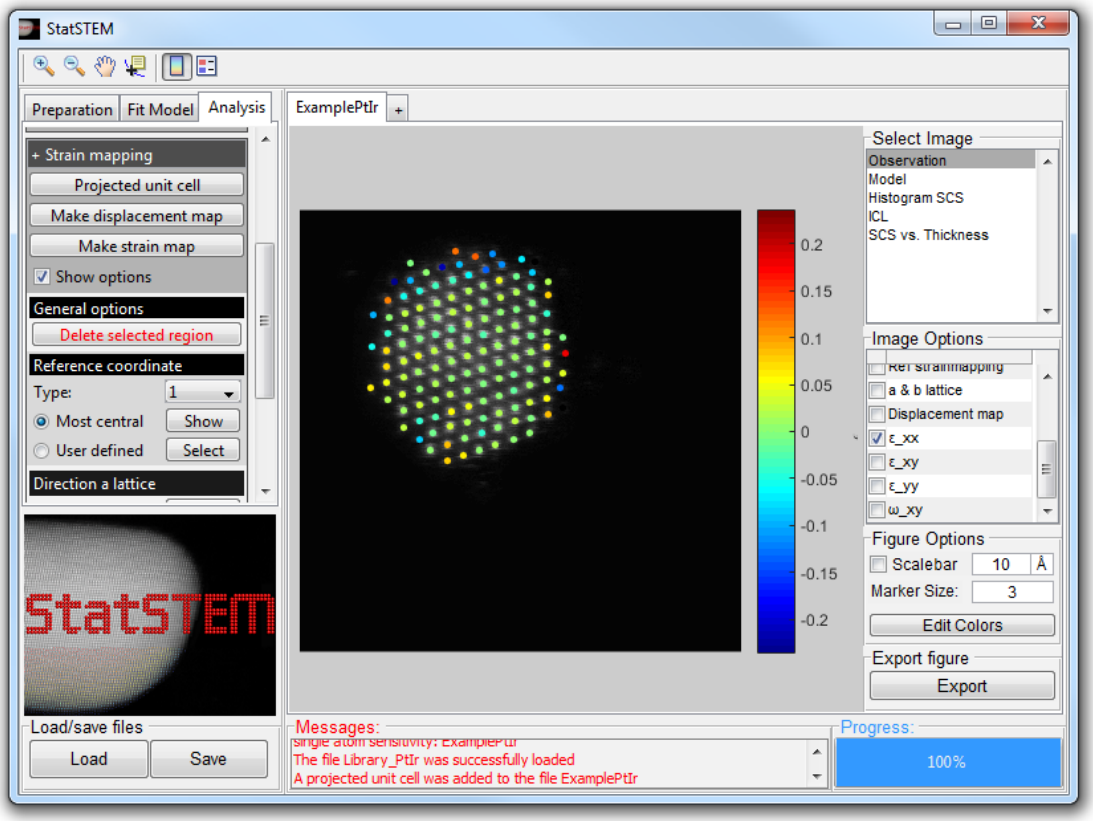

Figure 7.7: An ϵ_{xx} strain map of a Pt/Ir particle.

7.3.4 Advanced options

Select coordinates

With this options a selection of coordinates used for a displacement map and strain map can be made.

Reference coordinate

In this section a reference coordinate can be chosen. A displacement map is made from this coordinate. Furthermore, this coordinate gets the index (0,0) during the creation of the strain map. One can choose between different column types for selecting a reference coordinate. Automatically StatSTEM will use the most central coordinate as a reference, this can be change by manually selecting another coordinate.

Direction a lattice

For finding the most optimal direction of the a (and b) lattice a fit can be made. In this procedure a box of N unit cells (standard 3 UC to each side) around the reference coordinate is used for finding the direction of the a (and b) lattice parameter. Furthermore, one can select the a-direction to improve the values for the a and b lattice parameters by fitting. Automatically StatSTEM will use this fitting procedure. This option is advised to be used, as the pixel size recorded by an electron microscope is never that accurate. Be however, aware that when you are 100% sure about the pixel size, this option changes the values of a and b.

8 Closing StatSTEM

StatSTEM can be closed by hitting the red cross in the top-right corner. When closing StatSTEM, make sure that all files are saved as no warnings of unsaved files are given.

9 Remarks and suggestions

When downloading MATLAB, a folder with examples is included. In this folder MATLAB scripts are present to use StatSTEM without the graphical user interface. Our StatSTEM website offers a forum in case you have any questions, remarks or suggestions. New releases will be announced on this website.

10 References

[1] Annick De Backer, Karel H. W. van den Bos, Wouter Van den Broek, Jan Sijbers, and Sandra Van Aert. StatSTEM: An efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images. *Ultramicroscopy*, 171:104–116, 2016.

[2] S. Van Aert, K. J. Batenburg, M. D. Rossell, R. Erni, and G. Van Tendeloo. Three-dimensional atomic imaging of crystalline nanoparticles. *Nature*, 470:374–377, 2011.

[3] S. Van Aert, A. De Backer, G. T. Martinez, B. Goris, S. Bals, and G. Van Tendeloo. Procedure to count atoms with trustworthy single-atom sensitivity. *Physical Review B*, 87(064107), 2013.

[4] A. De Backer, G. T. Martinez, A. Rosenauer, and S. Van Aert. Atom counting in HAADF STEM using a statistical model-based approach: methodology, possibilities, and inherent limitations. *Ultramicroscopy*, 134:23–33, 2013.

[5] G. T. Martinez, L. Jones, A. De Backer, A. Bch, J. Verbeeck, S. Van Aert, and P. D. Nellist. Quantitative STEM normalisation: The importance of the electron flux. *Ultramicroscopy*, 159:46–58, 2015.

[6] Pedro L Galindo, Sławomir Kret, Ana M Sanchez, Jean-Yves Laval, Andrés Yáñez, Joaquín Pizarro, Elisa Guerrero, Teresa Ben, and Sergio I Molina. The peak pairs algorithm for strain mapping from hrtem images. *Ultramicroscopy*, 107(12):1186–1193, 2007.

Appendix A Structure of saved files

StatSTEM files may be saved in a MATLAB (*.mat*) format. In this file different structures are created in which all variables are stored. For each step or analysis, a structure is created. A list of all the variable and structure names is shown below in Table A.1.

Variable name	**Description**
input	Structure of preparation step
obs	image
dx	pixel size (Å)
coordinates	(n × 3) vector with the x- and y-coordinates and the column types (numbered)
types	names of the column types (cell format)
output	Structure of the fitted model
coordinates	x- and y-coordinates ((n × 2) vector)
rho	width of the Gaussian peaks ((n × 1) vector)
eta	height of the Gaussian peaks ((n × 1) vector)
zeta	background value
model	Gaussian model of the image
volumes	scattering cross-sections (volumes) of the Gaussian peaks ((n × 1) vector)
lsq	least squares sum
atomcounting	Structure of the statistical based atom counts
coordinates	x- and y-coordinates ((n × 2) vector)
volumes	scattering cross-sections (volumes) of the Gaussian peaks ((n × 1) vector)
Counts	number of atoms per atomic column
TotalNumber... Atoms	total number of atoms in the image
selType	selected column type
N	number of columns
offset	offset in atom counts
selMin	selected minimum
mLogLik	the negative log likelihood
ICL	ICL values per number of components

	((1 × n) vector)
estimated... Distributions	Gaussian mixture model information per number of components
estimated... Locations	Locations of the components of the Gaussian mixture model
estimated... Width	Width of the components of the Gaussian mixture model
estimated... Proportions	Proportions of the components of the Gaussian mixture model
libcounting	Structure of the simulating based atom counts
coordinates	x- and y-coordinates ((n × 2) vector)
volumes	scattering cross-sections (volumes) of the Gaussian peaks ((n × 1) vector)
Counts	number of atoms per atomic column
TotalNumber... Atoms	total number of atoms in the image
library	The scattering cross-section in function of column thickness ((n × 1) vector)
strainmapping	Structure of the strain mapping results
coor_sel	x- and y-coordinates ((n × 2) vector)
refCoor	(x,y)-coordinate of the reference column
teta	angle of the a lattice (mrad) with the 95% error in the second column
a	a lattice (Å) with the 95% error in the second column
b	b lattice (Å) with the 95% error in the second column
dir_teta_ab	boolean indicating whether the b lattice is rotated in the positive direction from the a lattice
coor_relaxed	(x,y)-coordinates of the relaxed positions of the atomic columns ((n × 2) vector)
types	the column type ((n × 1) vector)
indices	indices of the columns with respect to the reference coordinate. First column is displacement in a, second in b ((n × 2) vector)
eps_xx	The ϵ_{xx} strain values of the atomic columns ((n × 1) vector)

eps_xy	The ϵ_{xy} strain values of the atomic columns ((n $\times$ 1) vector)
eps_yy	The ϵ_{yy} strain values of the atomic columns ((n $\times$ 1) vector)
omg_xy	The ω_{xy} strain values of the atomic columns ((n $\times$ 1) vector)
errEps_xx	The error on the calculated ϵ_{xx} strain values of the atomic columns ((n $\times$ 1) vector)
errEps_xy	The error on the calculated ϵ_{xy} strain values of the atomic columns ((n $\times$ 1) vector)
errEps_yy	The error on the calculated ϵ_{xy} strain values of the atomic columns ((n $\times$ 1) vector)
errOmg_xy	The error on the calculated ϵ_{xy} strain values of the atomic columns ((n $\times$ 1) vector)

Table A.1: List of all the variable names in the saved MATLAB file by StatSTEM