

IFM Analyser: version 2.0

IFM Analyser is written in Mathworks MATLAB for dealing with the quantification of sarcomeric structures using pre defined templates (single line, double line, gap) in drosophila flight muscles. The application can only be used on rainSTORM exports. (We plan to implement the support for open formats.)

Contributing

Pull requests are welcome. For major changes, please open an issue first to discuss what you would like to change. Please make sure to provide a demo dataset.

Requirements

Any PC (with Linux or Windows) that has Mathworks MATLAB R2018b installed on can run IFM Analyser. The software depends on rainSTORM v3.1.7 (or newer) to provide the exported localization data. See requirements for rainSTORM on [our software page](#).

The **Curve Fitting Toolbox** is required to quantify the epitope distribution for the **double lines** type structure. All other functions work without the Toolbox.

Tested on

The application has been tested on the following configurations:

Operating System	Hardware configuration	MATLAB version	Notes
Win 10 Pro 1809	Intel i7-6700K, 32GB DDR4, Intel HD Graphics 530	R2018b Update 4	all functions work perfectly
Ubuntu 18.04.2	Intel i5-8350U, 8GB DDR4, Intel HD Graphics 630	R2018b Update 4	all functions work perfectly
Win 10 Pro 1905	Intel E6500, 4GB DDR2, Radeon HD5570	R2019a (no Curve Fitting Toolbox)	quantification of double lines is not possible
Win 7 Ent	Intel i7-, 16GB DDR3, Asus GTX670	R2019a Update 1	all functions work perfectly

Installation

There are two ways to install IFM Analyser:

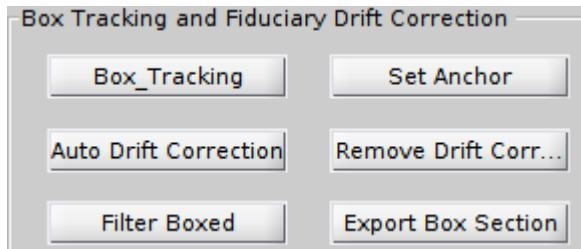
- Download and install the [packaged MATLAB App](#).
- Or clone the git repository:

```
git clone git@gitlab.com:adoptim/ifm-analyser.git
```

Initial notes

The application uses rainSTORM exported data files (Image Viewer -> Export Box Section). A rectangle shaped area around the studied structure ("double line", "band" or "gap") should be selected manually on the Super Resolved Image using the **Export box section** button. Localizations inside this selected area can now be analyzed further using IFM Analyser.

The stable [rainSTORM v3.1.7](#) has this feature included.



As the result, two files, a **csv** and a **mat** file are exported into the same folder where the raw measurement file is. Their names are created by attaching "ROI_Xxx_Yyy" to the original filename, where xx and yy are the selected area's center x and y coordinates (?). This naming convention has importance as when imported in the GUI application, the files are displayed with their cut-off, shortened names. Any deviation from this naming structure will cause an error when using the GUI application. If your files are named in a different way, use the launch scripts for the evaluation. The **csv** file contains only the event list (SectionParams) while the **mat** file is required by IFM Analyser. The latter has the following structure:

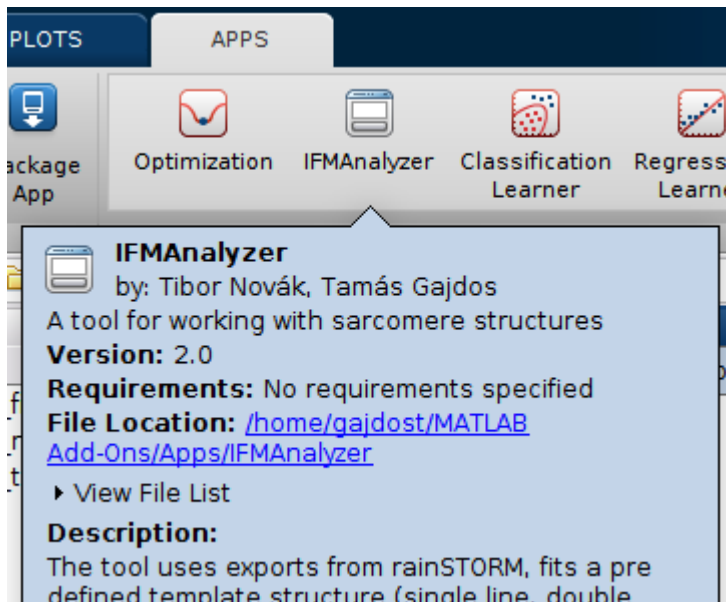
- export
 - SectionParams - full event list from the selected ROI
 - SectionPosits - x-y coordinate array in *nm* for each localization
 - SectionInfo - the four corners of the ROI
 - SectionRawDataFullFilename - the file from where the export was done
 - export_name - the name of the exported ROI: XxxYyy

Note: Before exporting we recommend using the following reconstruction filters to be applied in rainSTORM Image Viewer: **Thompson X-Y** 0-25 *nm*, **PSF size** 0.7-1.5 *px*. (Nem ezekkel értékeltek ki, és nem is ezek az alapértelmezten beállított paraméterek, valamint a pixelben mért PSF méret eléggé rendszerfüggő, másoknak valószínűleg más kell. Ettől még benne hagyjuk?)

Starting the application

Depending on the installation method the choices are:

Running the MATLAB app: Open MATLAB click on the Apps tab, and run the IFM Analyser.



Running the source code: Open MATLAB, change the current directory to the repository folder. Run:

```
startup; IFMAnalyser
```

When there is no Curve Fitting Toolbox detected the following message is displayed in the Command Window:

```
Mathworks MATLAB R2018b or newer is detected.  
There is no Curve Fitting Toolbox installed!  
Quantification of double line structures is not possible.
```

The GUI has three main sections:

- **Process Sections Data** - Fit the symmetry axis of one of the three structure types on multiple imported datasets and create the localization-symmetry axis distance histograms.
- **Process Histogram Data** - Fit a pre-defined template on the localization-symmetry axis distance histogram and quantify the structure (distance and width).
- **Merge** - The localization events of multiple exports can be aligned and merged to form a super-resolved image using the Simple Histogram method from rainSTORM.

Fitting a structure

There are two ways to import rainSTORM exported ROI files:

- **Add Files**, which opens a file browser window where one or several **mat** files can be selected.
- **Add Folder**, which will open all matching (**mat**) files within a folder.

Each imported file will appear in the **Sections** ListBox with their shortened names. To remove a single or multiple unwanted ROIs, select them and click **Remove Selected**.

Selecting a model:

- **Single Line** - fits a single line on a feature along the highest density. Then stores these localizations. (e.g. *obscurin*)
- **Double Line** - fits a single line between two high density parallel features, and calculates the distance of each localization from the line. (e.g. *kettin* or *p5*)
- **Gap** - fits a single line into a gap. (e.g. *actinRat*)

*Planned feature: if you are not satisfied with the quality of the localizations in the exported ROI, an import filter can be applied to remove low quality localizations based on **Precision** and **PSF Size**.*

To fit a template onto the selected Sections, click **Process Selected**. This will launch the background worker, which will import the selected files one-by-one and do the straightening. The worker can be put into an unattended mode, when **Auto Process** and **Auto Save** are both selected. To manually check each initial guess before accepting it, disable **Auto Process**.

After the structure fitting is done on the **Process Sections Data** tab, the selected exports can be moved to the **Process Histogram Data** tab or to the **Merge** tab by clicking the appropriate **Move to...** button. This will only move the elements in the ListBox and update the path connected to that export.

Structure fitting process

A straight line is coarsely fitted on the coordinates of the localizations inside the selected area. The centre coordinates of all the localizations, and the orientation of the straight line are used to define the position and the orientation of a smaller rectangle **X range * Y range** (typically set to 800*600 nm, here 800*400 nm) within which a further subset of localizations is cut out for quantitative evaluation. The size of this area is matched to the target feature: it is set approximately 3-4 times broader and 10% shorter than the actual structure to clearly separate it from its background and to eliminate artefacts at the myofibril edges.

Next, the symmetry axis of the structure is determined more precisely via a second-order polynomial fit (can be adjusted via **Polynomial degree of fitted line**), taking into consideration the curvature of the selected structure. In order to achieve this fit, a Gaussian kernel is applied on the localizations to get a smoothed localization density map. The size of the kernel (**Gaussian blur**) is chosen so that it would be large enough to effectively blur the localizations, but small enough to preserve the gap or the double line structures (~20 nm, here 30 nm). Then, a polynomial is fitted along the minima (in case of “double lines” and “gaps”) or along the maxima (in case of single “bands”) of the localization density, and is considered as the symmetry axis of the structure.

Fitting options to adjust

- **X, Y range** - the size of the search area and cropping rectangle
- **Gaussian blur** - the size of the gaussian kernel applied for smoothing the localization density map during the initial search
- **Line distance threshold** - the upper limit for the localization-symmetry axis distance (when applicable)
- **Number of bins** - used when generating the 1D histogram
- **Polynomial degree of fitted line** - 1-> straight; 2-> parabolic curve; etc...
- **PSF size maximum** - the upper limit for the PSF size during fit
- **Fitting power** - how to take the outliers on density map into account

The screenshot shows a software window with three tabs: "Process Sections Data", "Process Histogram Data", and "Merge". The "Process Sections Data" tab is active. It contains a list of sections on the left and a "Fitting Options" panel on the right.

Sections List:

- export_X25_Y128
- export_X44_Y139
- export_X59_Y12
- export_X8_Y116

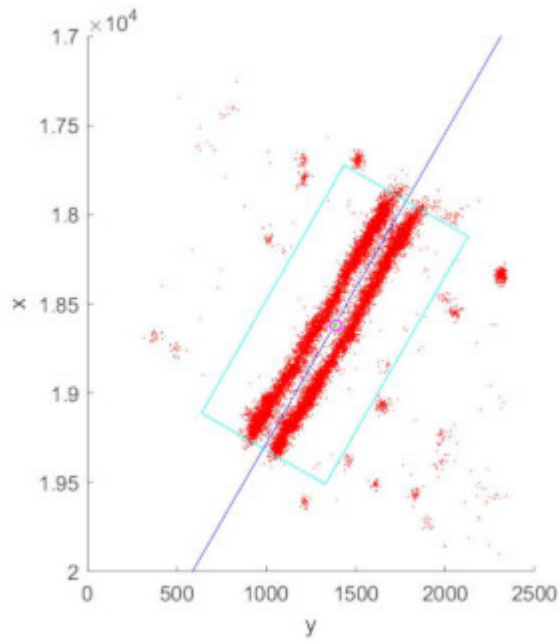
Fitting Options:

Parameter	Value
X range [nm]	800
Y range [nm]	400
Gaussian blur [nm]	30
line dist thr. [nm]	200
N of bins	50
poly. degree	2
PSF max [pixel]	1.1
fitting power	2

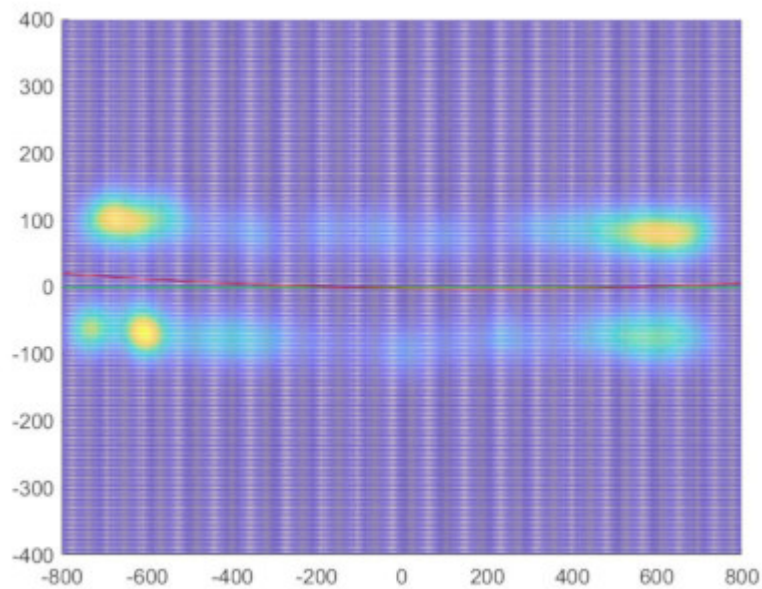
Buttons at the bottom: "Remove Selected", "Process Selected", "Move to 'Process Histogram'", "Move to 'Merge'", "Save List", "Add Files", "Add Folder", "Auto Process", "Auto Save", "Process All".

The fitting progress can be tracked in the Matlab Command window. The output files are generated using the following folder structure:

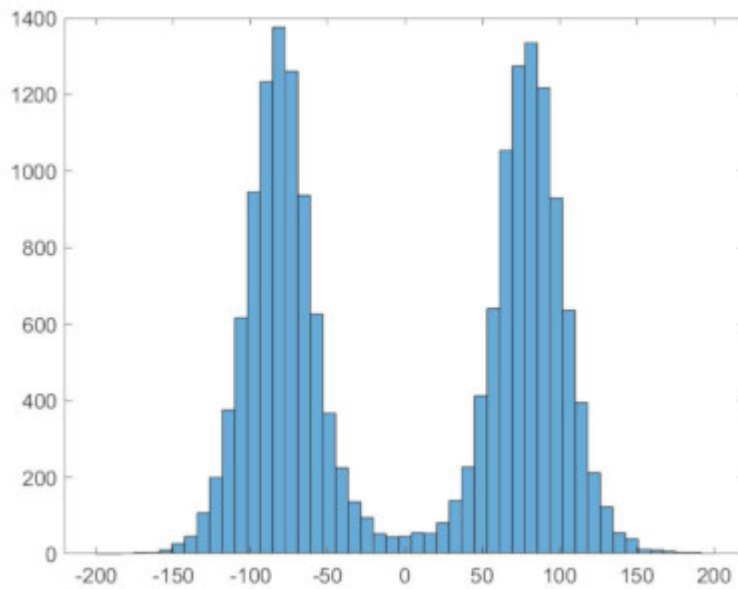
- [folder with exports]
 - **initLine** - the initial guess for the line with the localizations and the cropping box.



- **fittedCurve** - the actual line fits with the cropped localization density. These files should be checked for evaluating the fit quality.



- **histFigure** - the distance of each localized point from the fitted symmetry axis is determined numerically and depicted in a histogram.



- **histData** - the results are stored here (histogram, fitted curve, etc). These files are used for quantifying the epitope distribution.

Manual evaluating each fit

When **Auto Process** is not enabled the user will be prompted to fine tune each fit file-by-file.

?

Current angle of the initial line: 57.6122???

Current [x, y] displacement the rectangular region is: -2.4522 -2.9866 nm

Current length of the rotated rectangular region is: 800 nm

Current height of the rotated rectangular region is: 400 nm

Current standard deviation of the Gaussian blur is: 20 nm

Current maximum line-localization distance is: 200 nm

Current number of bins on the histogram is: 50

Current polynomial degree is: 2

Current max PSF size for filtering is: 2

Current power for curve fitting on surface is: 1

Accept parameters

Set parameters manually

Quit

The parameters used to guess for the symmetry axis and the **initLine** image are shown to the user to make a decision, whether to **Accept parameters**, **Set parameters manually**, or to **Quit**. When clicking the **Set parameters manually** button the user is prompted to adjust each parameter in a new dialog box.

Angle of the initial line:

[x, y] displacement the rectangular region [nm]:

Length of the rotated rectangular region [nm]:

Height of the rotated rectangular region [nm]:

Standard deviation of the Gaussian blur [nm]:

Maximum line-localization distance [nm]:

Number of bins on the histogram:

Degree of the polynomial to be fit:

maximal PSF size for filtering [pixel]:

Power minimizing the fitted curve along the surface:

Tip: In a typical case the angle and the displacement should be adjusted.

When **Auto Save** is not enabled the user will be prompted to make a decision about each fit file-by-file.

Choosing what to do:

Discard Current: Discards the results of the actual structure and proceeds to the next one. **Reevaluate Current:** Discards the current results and evaluate again the actual structure. The user will be prompted to modify the initial parameters to obtain a successful fit. **Save All:** Save the accepted results of all evaluated and yet unsaved structures in a single file. **Add Another or Leave:** Accepts the results of the actual structure and proceed to the next one. When saving, the accepted results of the different structures will be saved into a single file, although they will remain separate. When fitting the model function, the user will have the option to concatenate these results and fit on a single, averaged dataset.

Quantification of an epitope structure

The moved ROIs from the previous tab are shown in the **Histograms** ListBox. It is also possible to add previously fitted structures to this list by using the **Add Files** and **Add Folder** buttons similarly to the previous

case. The files to import should be selected from the *histData* folder.

It is possible to remove localizations based on their PSF Size to discard out of focus and overlapping localization events. The default acceptance range is between 0.8 (**PSF Min**) and 1.5 (**PSF Max**).

To quantify the epitope distribution histogram click **Process Selected** to start the process. To run in an unattended mode, the **Auto Process** and **Auto Save** options should be enabled.

The histogram data files are saved into a **mat** file with "_lineDiameterHistData" attached to the exported ROI file name.

There is an option to **concatenate** the localization-symmetry of all structures axis and calculate the epitope distribution from this averaged histogram. This option has only effect when the "The structure fitting process" was manually evaluated. (Ez lehet, hogy jobban illene a következő részhez.) (Változtassunk ezen a viselkedésen, hogy minden össze legyen fűzve, és konzisztensebb legyen?) **Tip**: enable this if the number of localizations are too low to rely on only one fit.

Quantification process

The distance of each localized point from the fitted symmetry axis is determined numerically and depicted in a histogram.

The acquired histogram shows the localization distribution rather than the epitope/binding site distribution which is closely related to the protein distribution. This discrepancy arises from two main reasons: **localization precision** and linker length (**Linker Radius**). Localization precision is determined by the localization algorithm based on the characteristics of the blinking event. Theoretically, the fluorophore distribution can be calculated by deconvolving the measured localization image with the localization precision distribution (which can typically be described by a Gaussian curve). The epitope distribution can be obtained from the fluorophore distribution (**Linker Type**) and the length of the linker (**Linker Radius**). We assume that the fluorophores are distributed evenly on a sphere around the epitope at a radius equivalent to the linker length (size of the primary and secondary antibodies, ~20 nm).

To quantitatively characterize the distribution of the epitopes, a theoretical curve, in which the aforementioned effects are considered, is calculated and fitted on the measured data. During this process, we approximate the epitope distribution (**Sample Type**) of the "band", the "double line" and the "gap" structures with a single Gaussian, with double Gaussians and with negative top-hat functions with additional peaks at the edges, respectively. Then the approximated epitope distribution is smeared with the linker and with the average Thompson precision using convolution. The resultant distribution is compared with the measured histogram and the fitting algorithm varies the epitope distribution parameters until good agreement is met. From the fitted parameters, the values quantitatively describing the "double line", "band" and "gap" structures can be extracted in a straightforward manner.

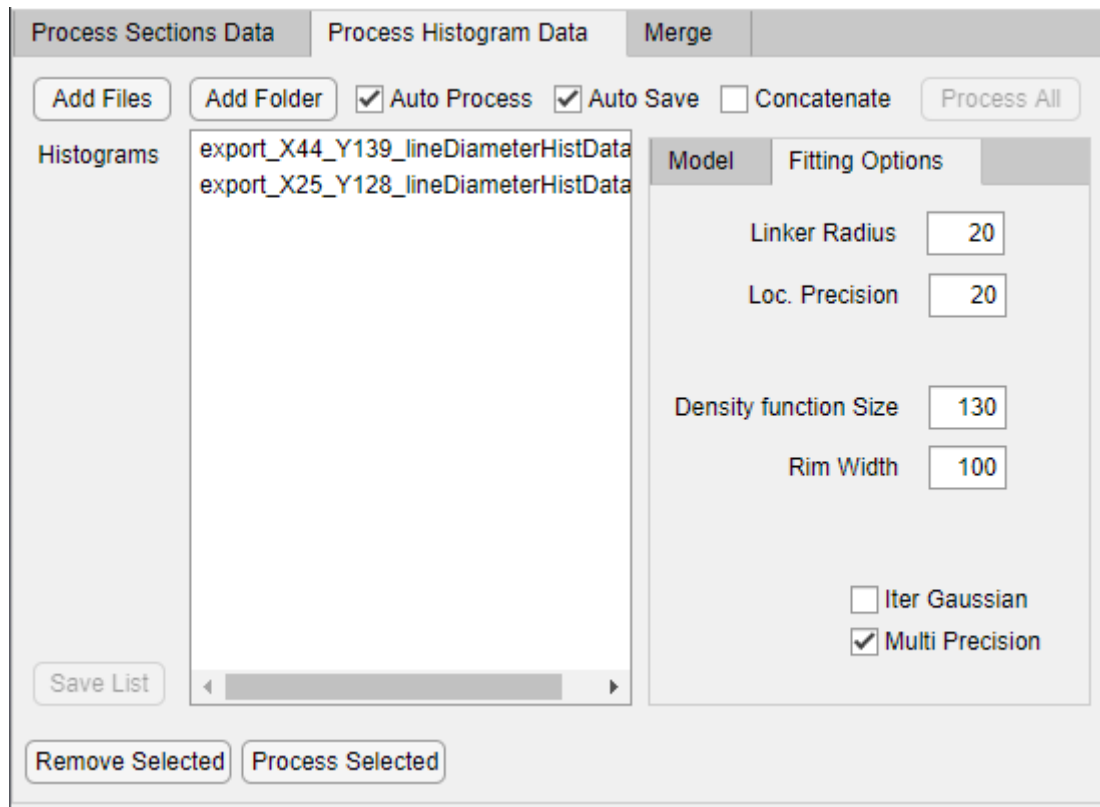
Model options to adjust

- **PSF Min, Max** - import filters to remove out of focus and overlapping localizations
- **Model Type** - select the type of the model function to fit
 - **Single or double band structures**
 - circle - dye density from projecting a circle, for filled tubes
 - disk - dye density from projecting a disk, for hollow tubes

- Gaussian - dye density following Gaussian distribution
- Lorentzian - dye density following Lorentzian distribution
- Rectangular - dye density following top-hat distribution
- **Gap Structure**
- Gap - dye density following negative top-hat distribution with additional rims at the edges. Should be used for the **gap** sample.
- **Sample Type** [single line, double lines, gap] - the same model should be used as on the previous tab.
- **Linker Type** [sphere, Gaussian] - fluorophore distribution around the epitopes

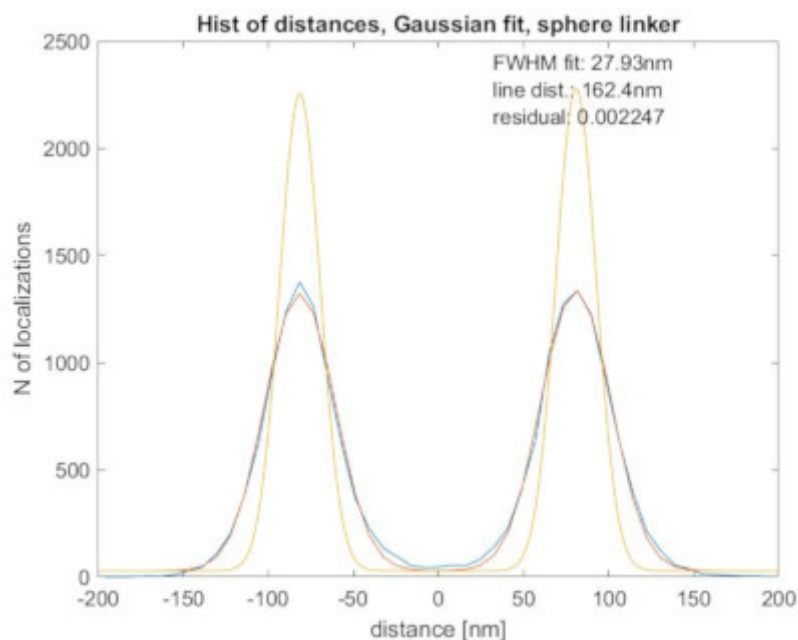
Quantifier fitting options to adjust

- **Linker Radius** - length of the linker, the size of the primary and secondary antibody combined
- **Localization Precision** - global localization precision to be taken into account when calculating the epitope distribution
- **Density function Size** - *only used for gap structure*
- **Rim Width** - *only used for gap structure*
- **Iter Gaussian** - *only used for gap structure*
- **Multi Precision** - the average localization precision is calculated for each histogram from the localization data instead of a manually set value. When it is enabled, the **Localization Precision** option is discarded.



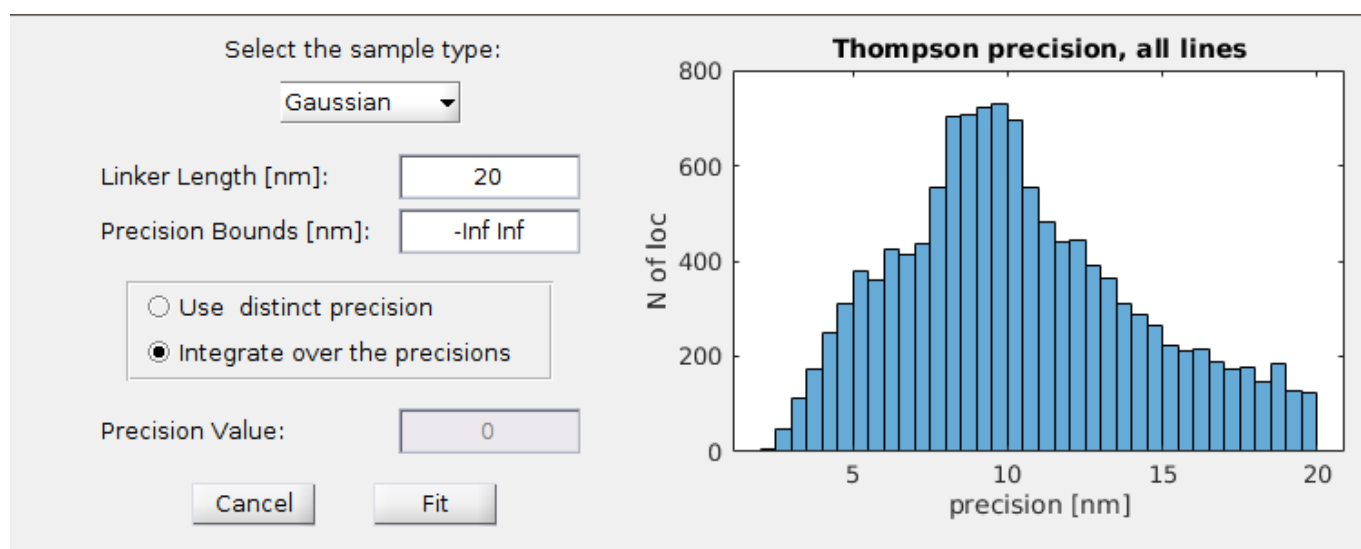
The histogram quantification progress can be tracked in the Matlab Command window. The output files are generated using the following folder structure:

- [folder with exports]
 - [histData]
 - **fittedParams** - contains the the fitted parameters histogram-by-histogram. Each model has its different parameters.
 - **allfittedparams.csv** - contains the parameters from all the files
 - **fittedWithFigure** - bug?
 - **fittedWithImage** - visual representation of the fit

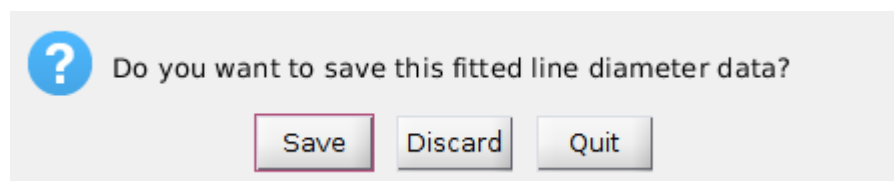


Manual evaluation of the quantifications

When **Auto Process** is not enabled the user will be prompted to fine tune each quantification histogram-by-histogram.



When **Auto Save** is not enabled the user will be prompted to make a decision about the quantification fit whether to **Save** it, **Discard** it, or to **Quit** the whole process.



Merge structures

The moved ROIs from the **Process Sections Data** tab are shown in the **Merge** ListBox. It is also possible to add previously fitted structures to this list by using the **Add Files** and **Add Folder** buttons similarly to the previous case. The files to import should be selected from the *histData* folder.

The default merge options should work for most cases. To create the final averaged image, select the list elements, and click **Process Selected**. The merge progress can be tracked in the Matlab Command window.

Merging process

During the initial structure fitting (on the **Process Sections Data**) each structure is aligned to have its center of mass at (0,0). Also the guessed curve (symmetry axis) is stored in the *histData* files.

The distance of each localization from the fitted symmetry axis, and the intersection of the perpendicular line is determined using the [distance2curve](#) function. Using these two coordinates the localizations can be straightened along the symmetry axis, and can be merged to create the Super-Resolved Image using the Simple Histogram method from rainSTORM.

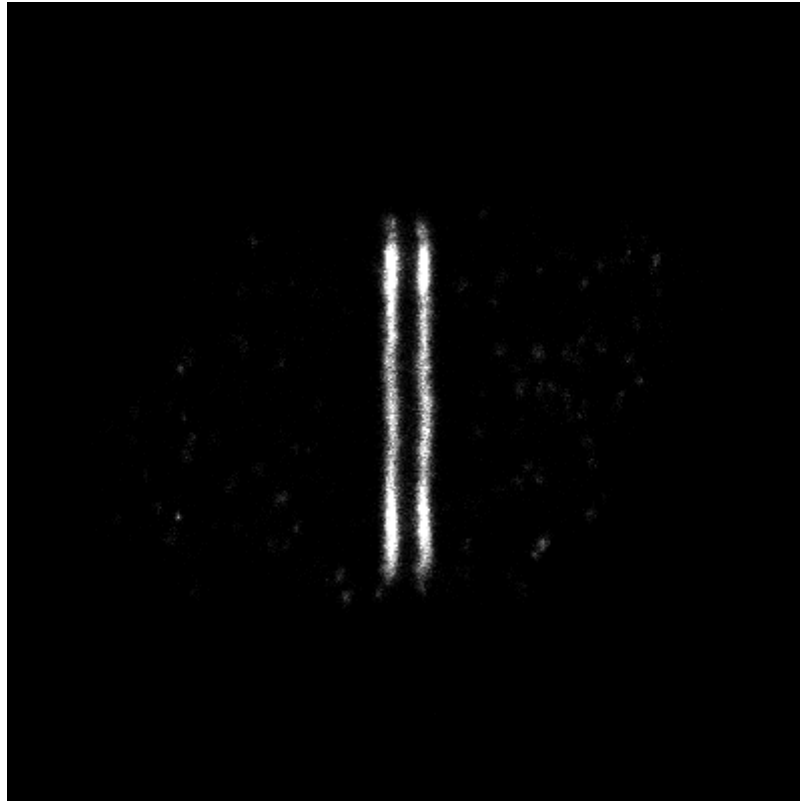
Merge options to adjust

- **Line Fit Model** - how to recreate the curve (symmetry axis), which was guessed during the structure fitting (**linear** is the fastest, and gives a good guess with a small **stepsize**)
- **Bounds** - the extra length to add to each end of the curve (helps with outliers)
- **Stepsize** - the iteration step-size used when recreating the guessed curve
- **Image Size X, Y** - the image size in *nm* to use when creating the Super-Resolved Image from the merged localization event list. The image size in pixels are calculated as **Image Size/Pixel Size**.
- **Pixel Size** - the pixel size to use when creating the Simple Histogram based Super-Resolved Image.

The output files are generated using the following folder structure:

- [folder with exports]
 - [histData]
 - **mergeData**

- **concatenated_xydata.mat** - contains x-y coordinates from all selected exports, but separated into cells
- **mergeOutput.mat** - a more comprehensive structure (localizations from exports in cells, merged localizations, image settings, image simple histogram array) which can be used to recreate the **output-MERGEImage.png**
- **output-MERGEImage.png** - the resulting Super-Resolved Image, which can be viewed in e.g. Fiji/ImageJ



Appendix, Examples, Known Bugs

License: [GPLv3](#)

Running without GUI

We created helper scripts to enable the usage without GUI. Each script replaces the functionality of a tab, and the options can be adjusted in the header. Here are the short extracts:

noGUI_launch_script_lineDiameter.m

```
% script for launching the line diameter determination tools without GUI

%% parameters for the line diameter determination
% parameters for automatic/manual evaluation
automaticSaveBool=false;
automaticProcessBool=false;

% parameters that appear in the pop-up window for setting cut-out region
initParams.xrange=600;           % "x" size of the ROI's rectangle
```

```

initParams.yrange=400;          % "y" size of the ROI's rectangle
initParams.sig=20;              % STD of the Gaussian smearing
initParams.sampLineDistThreshold=200; % maximum distance of
localizations considered from the central curve in nm
initParams.nBins=50;            % number of bins on the shown histogram
initParams.polyDegree=2;        % degree of the fitted central curve
initParams.PSFmax=2.0;          % maximal accepted fitted Gaussian STDs of
localizations in pixel size
initParams.surfFittingPower=1; % raise to this power the localization
density map

structureType='Single line';
% structureType='Double line';
% structureType='Double line (Fallback)';
% structureType='Gap';

```

noGUI_launch_script_lineDiameter_fit.m

```

%% Parameters to set manually
% parameters for automatic/manual evaluation
automaticSaveBool=false;
automaticProcessBool=false;

% parameters for the convolution to calculate the model function to fit
model.linkerRad=20; % radius of the linker, dyes can
only be positioned a sphere surfaces around the epitopes
model.multiPrecBool=true; % calculate the average of the
Gaussian distributions derived from the individual localizations precision
values
model.locPrecision=[]; % given standard deviation of the
localization precision, if empty or zero, it will take the average of the
STDs of the localizations
model.iterGaussianConvBool=false; % whether the the algorithm should
apply additional Gaussian smearing beside the localization precison, can be
used for the "gap" structure, the STD of this one is iterated

% select the type of the model function to fit
% model.modelFuncType='Circle'; % dye density from projecting a
circle, for filled tubes
% model.modelFuncType='Disk'; % dye density from projecting a
disk, for hollow tubes
model.modelFuncType='Gaussian'; % dye density following Gaussian
distribution
% model.modelFuncType='Lorentzian'; % dye density following
Lorentzian distribution
% model.modelFuncType='Rectangular'; % dye density following top-hat
distribution
% model.modelFuncType='Gap'; % dye density following negative
top-hat distribution with adittional rims at the edges

% string variable for whether the curve fitting should be executed on a
double line or on a single line or on a gap

```

```

model.sampleType='single line';
%model.sampleType='double lines';
%model.sampleType='gap';

% type of the dye distribution caused by the linker, "sphere" or "Gaussian"
model.linkerType='sphere';
%model.linkerType='Gaussian';

% some parameters for the gap fitting
model.densFuncSize_init=130;
model.rimWidth=50;

% might need some love, in the "rainSTORM_extras_lineDiameter" there is
"PSFmax"
% allowed fitted Gaussian STD range
clearing.PSFSizeMin=0.8;
clearing.PSFSizeMax=1.5;

averageDistBool=false; % boolean for whether the distances belonging to
different line should be catenated (false) or should be a handled
seperately, the "Concatenate" button in the GUI

```

noGUI_launch_script_lineDiameter_merge.m

```

% Launch script for aligning the individual single band (single line),
% double band (double line) or gap structures and merging them into a
single
% image.

%% input parameters to set:
mergeParams.sizeX=4000; % "x" size of the "meged" super-resolved
image in nm
mergeParams.sizeY=4000; % "y" size of the "meged" super-resolved
image in nm
mergeParams.pixelSize=10; % size of the pixels in the "meged" super-
resolved image
% interpolation methods for the straightening:
mergeParams.interp='linear';
%mergeParams.interp='pchip';
%mergeParams.interp='spline';
mergeParams.bound=100; % boundary region size for the curve for
the straightening
mergeParams.stepsize=0.5; % sampling step for the curve for the
straightening

```

Data structure and files

Running IFM Analyser without the Curve Fitting Toolbox

Every function will work as intended, except the quantification of **double lines**. Matlab will throw the following error:

```
Undefined function 'fit' for input arguments of type 'double'.

Error in lineDiameter_fit (line 358)
    f=fit(hist_X, histCounts,'gauss2');

Error in rainSTORM_extras_lineDiameter_fit (line 110)
    [fittedParamsTable, fittedHistFigure]=lineDiameter_fit(
lineDiam_histData_input(idxLine), ...

Error in IFMAalyzer/ProcessSelectedButton_HPushed (line 343)
    csvPath =
rainSTORM_extras_lineDiameter_fit(app.HistogramsListBox.Value, clearing,
model,...

Error using
matlab.ui.control.internal.controller.ComponentController/executeUserCallba
ck (line 335)
Error while evaluating Button PrivateButtonPushedFcn.
```