Nanoclusters curve-fitting (How to use)

Before start

- Download the repository or the content of this folder.

Requirements and set up

Requirements:

- Tested in MATLAB v R2018b or higher
  - Required toolboxes:
    Image Processing Toolbox
    Statistics and Machine Learning Toolbox

Set up:

Copy the scripts provided (`./curve-fitting-scripts`) to the MATLAB folder (usually located in Documents directory).

This script performs the following analysis:

- Inputs: two comma-separated values (CSV) tables

  - **Calibration data or single molecule data** - table with intensity amplitudes of single mRNA molecules (i.e.: smFISH spots at low mRNA density).
    Should contain one column with the following data:
    - "**AMP**": intensity amplitudes of calibration data (detected smFISH spots).

Image	AMP
img1	569.148
img1	928.681
img1	1023.41
...	...
img(n)	925.117
img(n)	678.812
img(n)	1030.36
...	...

  - **Experimental data** - table with intensity amplitudes of experimental data (i.e.: smFISH spots in oocytes).
    Should contain three columns with the following data:
    - "**Image**": unique identifiers for each image in the dataset.
    - "**CELL_Oocyte**": unique identifiers for each cell/oocyte/ROI in the dataset.
    - "**AMP**": intensity amplitudes of experimental data (detected smFISH spots).

| Image	CELL_Oocyte	AMP
img1	1	1089.52
img1	1	997.423
img1	•••	•••
img1	2	949.656
img1	2	1209.73
img1	•••	•••
img1	3	1043.97
img1	3	1463.85
img1	•••	•••
img1	4	1150.72
img1	4	1119.68
•••	•••	•••
img2	1	1815.1
img2	1	1572.31
img2	•••	•••
img2	2	1037.62
img2	2	684.797
img2	•••	•••
img2	3	1703.06
img2	3	973.444
img2	•••	•••
img2	4	756.41
img2	4	1159.19
•••	•••	•••
img(n)	1	1256.4
img(n)	1	732.331
•••	•••	•••

- Analysis:

    1. The distribution of amplitudes in the calibration data will be fitted to an skewed gaussian distribution (curve-fitting: lsqcurvefit - Least Squares).
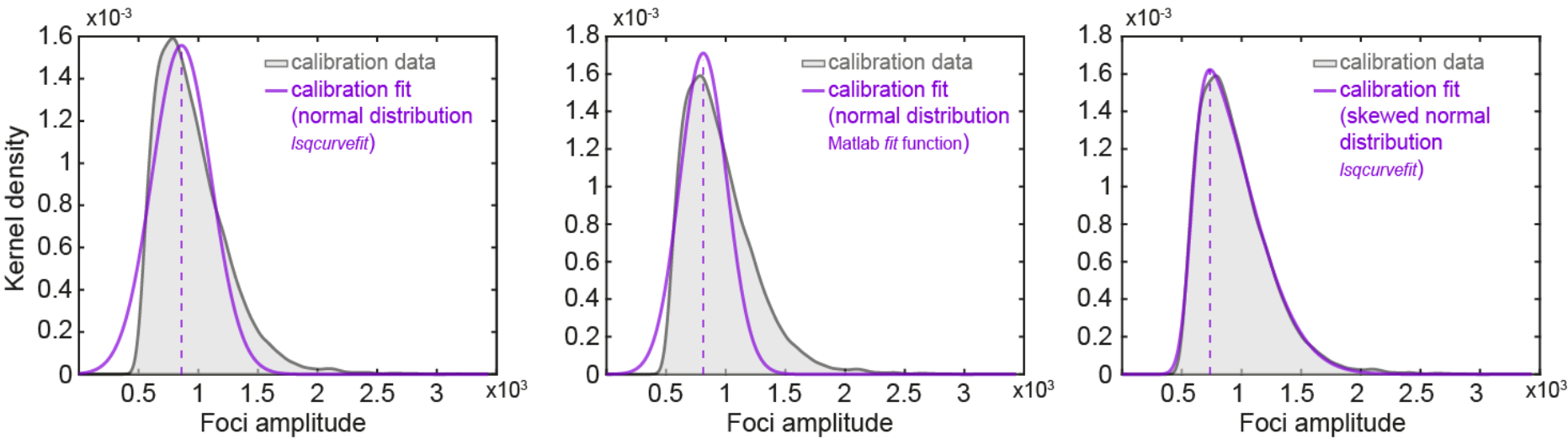
    - Calibration data fit examples. Skewed normal distribution yields a better fit.

    - The distribution of amplitudes in the experimental data will be fitted to an skewed gaussian model of two components, where the first curve estimates the single molecules (based on the calibration data) and the second curve estimates the nanoclusters.
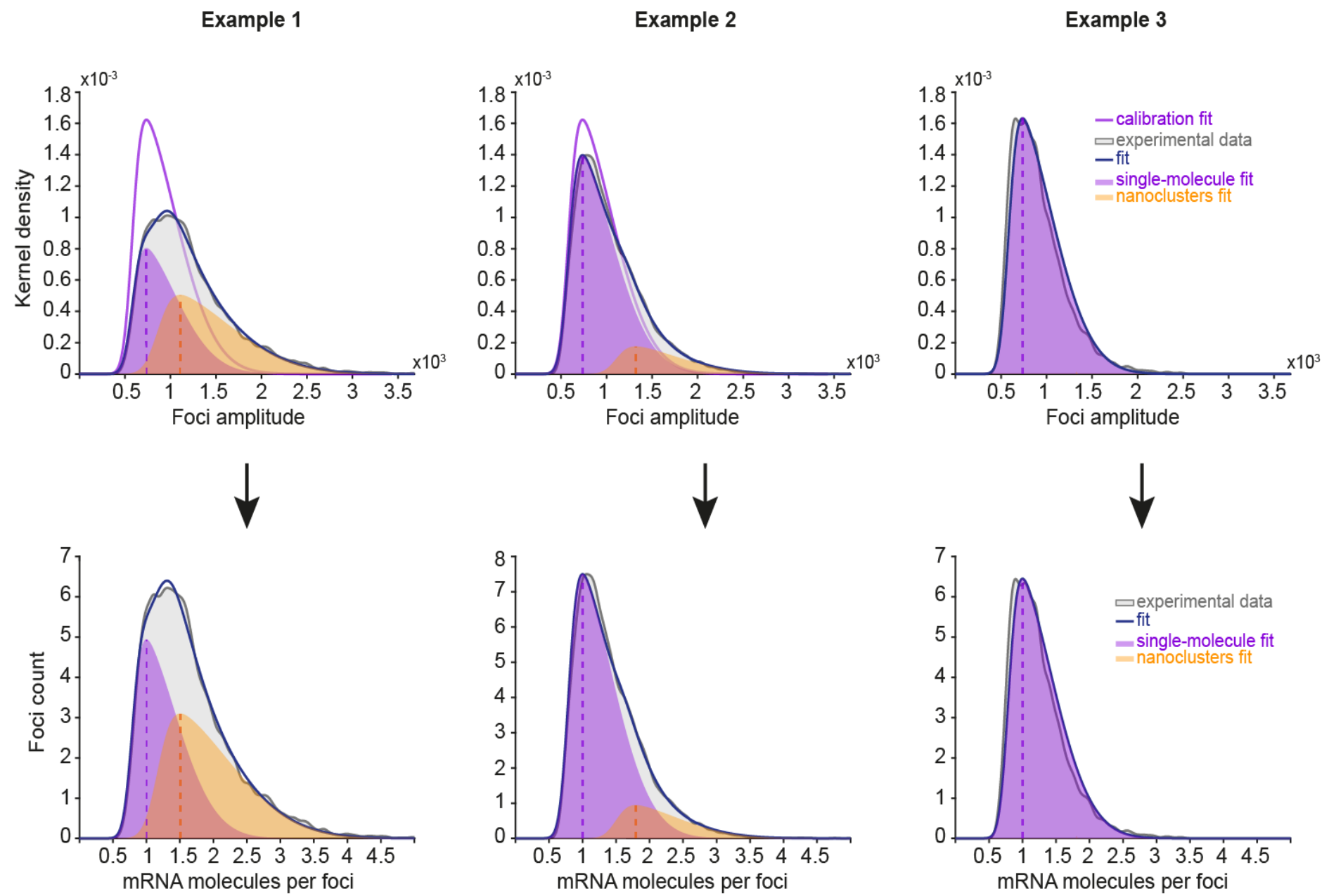
- Experimental data fit examples. Examples of single-molecule and nanocluster fitting.

- Outputs

  - Number of single molecules and number of molecules within nanoclusters.

image	oocyte	PDF_SM	foci_SM	molecules_SM	PDF_nano	foci_nano	molecules_nano	detected_foci
example1	oocyte_4	0.492620169	3028	3844	0.505399772	3106	6078	6146
example2	oocyte_4	0.856009444	4593	5831	0.150199915	806	1749	5366
example3	oocyte_4	1.000239346	3950	5014	0.000151529	1	1	3949

User guide

Data provided

- **Calibration data** (location: `./example-data/calibration_data.csv`):

  ```
  .csv file with column "AMP" among others % (i.e.: obtained using FISH-quant).
  ```

- **Experimental data** (location: `./example-data/experimental_data.csv`):

  ```
  .csv file with columns "Image", "CELL_Oocyte", "AMP", among others % (i.e.: obtained using FISH-quant).
  ```

Analysis (Procedure):

1. Open MATLAB.
2. In the command window type:
   `script_fit_nano_clusters`
   and then press enter.
3. Select the **calibration data** .csv file.

4. Select the **experimental data** .csv file.
5. Results are saved in the folder where **calibration data** is located.

The default parameters of the script include:

```
% ==== parameters for fitting
...

% sampling
parameter.pdf_step       = 0.5; % step for pdf evaluation from min : pdf_step : max(data)
parameter.pdf_Band_Width = 40;  % kernel band width
parameter.peak_fixed     = 1;   % fixed peak at boundary: force curved-fitted parameters of calibration data
to be fixed in the skewed gaussian model (Yes=1, No=0).
parameter.pdf_start      = 0;   % starting pdf from 0
parameter.separation     = 1.4; % controls the separation between dissolved and clustered curves

% show plots of fitted data
show.plots          = 0; % Yes= 1, No=0
show.display_range  = 5; % display range: number of molecules

...
```

pdf: probability density function.

To change this:

- Type in the command window:

  ```
  edit script_fit_nano_clusters
  ```
- Press enter.
- Then, default values can be manually modified (lines # 5 to # 14).

Outputs:

Provided output examples:

The fitting results, for the example data provided in this tutorial, are included (location: `./example-data-fitted`).