Single Molecule Clustering for Super-resolution Optical Fluorescence Microscopy

Authors : Prakash Joshi, and Partha P. Mondal

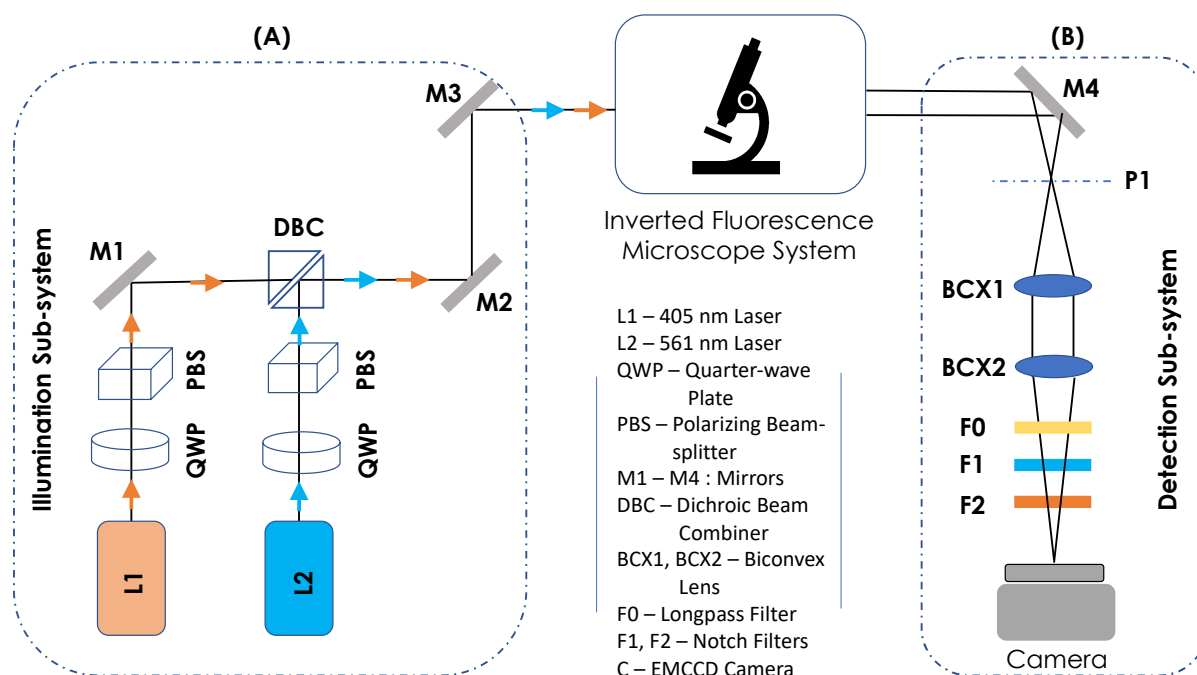# Supplementary S1: Experimental Optical Setup

Fig. S1: The schematic diagram of the actual experimental setup (excluding the optical arm for visualizing transfected cells). (A) Illumination subsystem, the inverted fluorescence microscope, and (B) the detection subsystem.

The optical setup for super-resolution system used for the clustering study is shown in Fig. S1. It consists of three major sub-systems : (A) illumination, and (B) Widefield detection sub-systems integrated with a high resolution inverted fluorescence microscope.

The illumination sub-system consists of two light sources. The first light source of wavelength 405 nm used for activation of single molecules, while the second light source of wavelength 561 nm for excitation. The linearly-polarized light from the lasers are allowed to pass through a combination of quarter-wave plate (QWP) and polarizing beam-splitter (PBS). This combination facilitates intensity control: QWP rotates the polarization of incident light, and the PBS splits the beam into s- and p- polarized light, of which p-polarized beam pass through and s-polarized light reflects in the perpendicular direction. The proportion of transmitted light passing through the combination can be altered by appropriately rotating the QWP. We have used 6.12mW for excitation and 112 μW for activation. The light from both the lasers are combined using a dichroic beam-combiner (DBC), and subsequently directed towards the back-port of inverted microscope by the mirrors M2 and M3.

We have used an Olympus IX81 inverted fluorescence microscope that is equipped with high NA objective (100X, 1.3 NA). The microscope is equipped with filter-cube containing dichroic mirror (of cutoff wavelength, 570 nm) and long-pass filter (of cutoff wavelength, 575 nm). This allows the both the lights (405 nm and 561 nm) to get reflected to the back-aperture of objective lens. Subsequently, the fluorescence emitted (at a maximum wavelength of 573 nm) is collected by the objective lens and pass through the dichroic mirror. The dichroic mirror eliminates the incident light and transmits the fluorescence light. The light is then focused by an internal lens (of focal length = 180 mm) to the point P1 and directed to the detection sub-system.

The detection consists of 4X magnifier system (consists of 2 lenses, BCX1 (focal length =75 mm) and BCX2 (focal length = 300 mm)). This 4X magnification along with the 100X magnification results in a total magnification of 400X. The light emitted from the specimen (Dendra-2HA with a peak emission at 573 nm) is then focused to the EMCCD camera (Andor 897 iXon Ultra camera, Andor Technology, UK) by a tube-lens BCX2 (focal length = 300 mm). On its way, the fluorescence light is then filtered by a combination of long-pass and notch filters (F0, F1 and F2) to remove the incident light. The data collection procedure is detailed in Supplementary S2. A complete detail of the optical elements used is given in Table S1.

Table S1. Optical elements used in the setup.

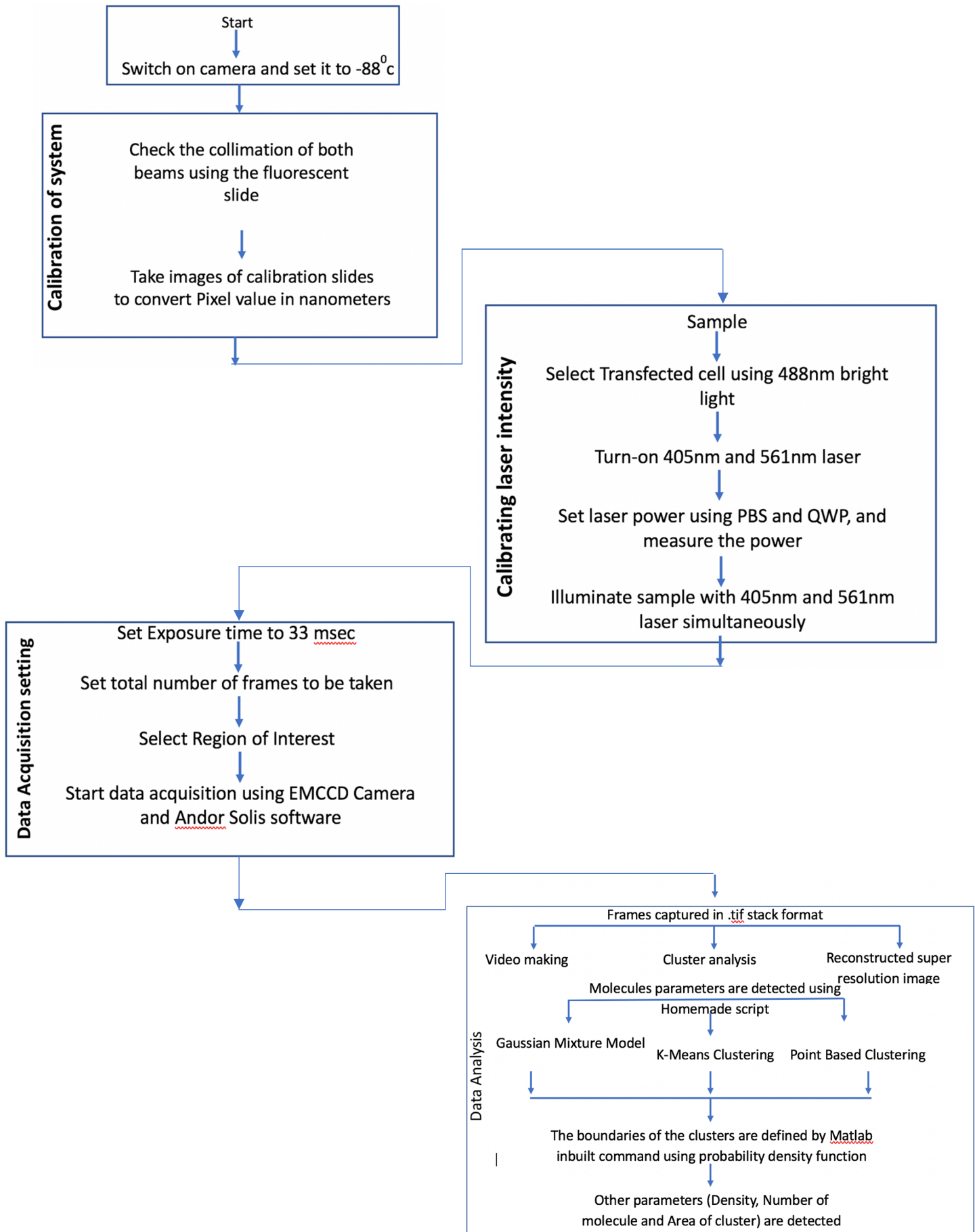| Optical Items	Company name	Item number
Dichroic mirror (425 nm)	Thorlabs	DMLP425R
Dichroic mirror (561nm)	Semrock	Di03-R561-t1-25*36
Lens (L1 = 75mm and L2 = 200mm)	Thorlabs	
Mirrors	Thorlabs	
Lasers (561nm and 405nm)	Oxxius	405 LBX and 561 LCX
Camera (EMCCD)	Andor Technology	Ixon 897 Ultra
Long Pass Filter (593/46-25 nm)	Semrock	FF01-593/46-25
Notch Filter (405 nm)	Semrock	NF03-405E-25
Notch Filter (561nm)	Semrock	NF03-561E-25
Auto-tunable lens	Edmund optics	88939
Objective Lens	Olympus	100x,NA 1.3 oil

Supplementary S2: Data Collection and Analysis

During Experimentation, a large number of single molecule images were recorded at an exposure time of 30 ms. The raw data was then subjected to spot-detection for identifying single molecules. This is achieved by subtracting the background followed by thresholding. An appropriate threshold is chosen to extract the spots indicating single molecules. To determine the threshold, an average of a large number of pixels in several frames is considered and the threshold value is chosen 4 times larger than the average value. All the recorded data (approximately, 5000 images) are subjected to the Einzel Matlab script [1-4], and important parameters are extracted. The images comprising single molecules and real-time reconstruction of super-resolved image for both sparse and dense clustered data are shown in Supplementary Videos S1 and S2, respectively. The parameters include, detected photons per molecule, centroid, molecule position, variance among others. Once the single molecules are extracted and characterized, it is subjected to cluster analysis by three different methods (K-means, GMM, Point-Clustering). The entire process from data collection to analysis is summarized the flow-chart.
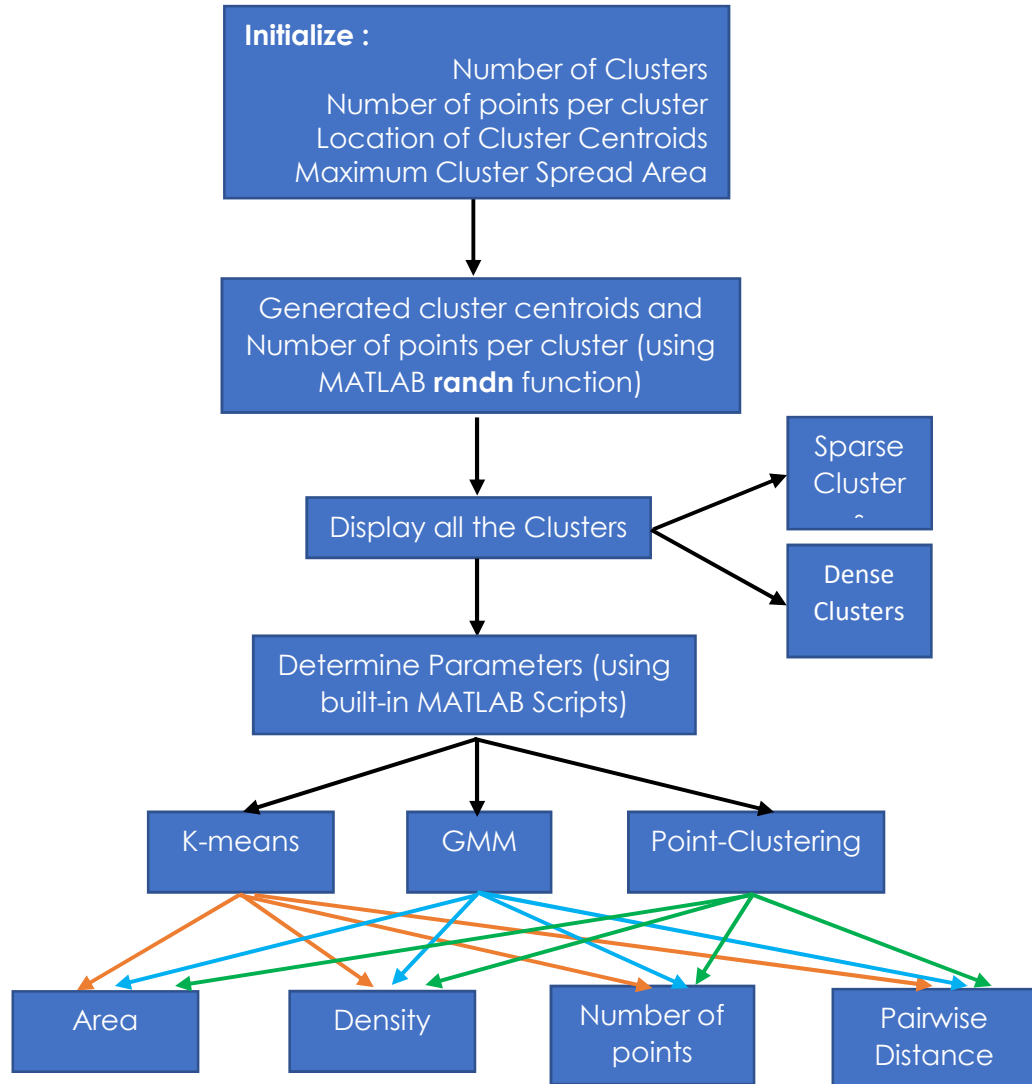
References:

1. S. T. Hess, T. P. K. Girirajan, and M. D. Mason, Biophys.J. 91, 4258 (2006).
2. N. M. Curthoys , Biophys J . 116, 893-909 (2019).
3. Partha P. Mondal and S. T. Hess, App. Phys. Lett. 110, 211102 (2017).
4. Partha P. Mondal, Micros. Res. Tech. 80, 333 (2017)..

Flow Chart

Start

Switch on camera and set it to -88^{0}c

Calibration of system

Check the collimation of both beams using the fluorescent slide

Take images of calibration slides to convert Pixel value in nanometers

Calibrating laser intensity

Sample

Select Transfected cell using 488nm bright light

Turn-on 405nm and 561nm laser

Set laser power using PBS and QWP, and measure the power

Illuminate sample with 405nm and 561nm laser simultaneously

Data Acquisition setting

Set Exposure time to 33 msec

Set total number of frames to be taken

Select Region of Interest

Start data acquisition using EMCCD Camera and Andor Solis software

Data Analysis

Frames captured in .tif stack format

Video making   Cluster analysis   Reconstructed super resolution image

Molecules parameters are detected using Homemade script

Gaussian Mixture Model   K-Means Clustering   Point Based Clustering

The boundaries of the clusters are defined by Matlab inbuilt command using probability density function

Other parameters (Density, Number of molecule and Area of cluster) are detected

Supplementary S3. Test Sample Generation and Processing

Flow Chart

The test cluster sample is computationally generation by first initializing the number of clusters and related parameters. Subsequently, in-built MATLAB " randn " function is employed to distribute the clusters and also for fixing the number of points in clusters. The sampling of pixels is kept same as that of real super-resolved image and subsequently the pixels are converted to micrometers. The area to each cluster is assigned using MATLAB " polyarea" function that use Green's theorem to calculate area. Finally, all the methods (K-means, GMM and point-

clustering) are used to determine the parameters (Area, Density, number of points and pairwise distance) [1-5].

References:

1. G. McLachlan, and D. Peel. Book: Finite Mixture Models. Hoboken, NJ: John Wiley and Sons, Inc., 2000.
2. M. Ankerst, M. M. Breunig, H. P. Kriegel, J. Sander, Proc. ACM SIGMOD'99 Int. Conf. on Manage. Data, Philadelphia PA, 1999.
3. Y. Marcon, MATLAB Central File Exchange. 2021.
4. D. Arthur and S. Vassilvitskii. Proc. 18th Annual ACM-SIAM Sym. Disc. Algo. 1027 (2007).
5. S. P. Lloyd, IEEE Trans. Info. Theory 28, 129 (1982).