![System requirements](https://img.shields.io/badge/python-3.8-red.svg)

![System requirements](https://img.shields.io/badge/platform-win%2064,%20linux%2064-green.svg)

[![License: GPL v3](https://img.shields.io/badge/license-GPLv3-blue.svg)](https://www.gnu.org/licenses/gpl-3.0)

# Analysis code for "the role of protein shape in multiphasic separation within condensates"

The GitHub repository contains analysis programs for the study "the role of protein shape in multiphasic separation within condensates" by V Pandey et al. [1].

All programs were written in Python3.8 (Windows) and designed for the analyses of output from LASSI simulation engine [2].

[1] https://www.biorxiv.org/content/10.1101/2024.08.26.606306v1

[2] https://github.com/Pappulab/LASSI

| directory | contents |

| -------- | -------- |

| \*\*`bin`\*\* |Executable programs. |

| \*\*`lib`\*\*| Shared libraries. |

| \*\*`workspace`\*\*| Accessary programs. |

### specification\_datasets.py

Analyses are conducted through two steps: the conversion of lammpstrj files into intermediate data files and the visualization based on the intermediate data. Their directories and filenames are specified as instance variables of a superclass "SpecDatasets" in "specification\_datasets.py." In the methods of "SpecDatasets", the instance variables below must be defined. They are referred by executable programs (bin/).

| instance variable | variable type | content |

| -------- | -------- | -------- |

| \*\*`self.dir\_lammpstrj`\*\* | str | Directory for target lammpstrj files. |

| \*\*`self.dir\_edited\_data`\*\*| str |Directory for intermediate data files. |

| \*\*`self.dir\_imgs\_root`\*\*| str | Directory for image files. |

| \*\*`self.filenames\_lammpstrj`\*\*| list/tuple | Filenames of target lammpstrj files. |

| \*\*`self.filenames\_edited`\*\*| list/tuple | Filenames of intermediate data files. |

| \*\*` self.filename\_lammpstrj\_matrix `\*\*| func(v, l) | Filename of a target lammpstrj file specified by valency/GluN2B conc (v) and length/STG conc (l). It is used to draw matrix graphs and phase planes. |

| \*\*`self.filename\_edited\_matrix`\*\*| func(v, l) | Filename of the intermediate data file specified by valency/GluN2B conc (v) and length/STG conc (l). It is used to draw matrix graphs and phase planes. |

### Some .py files

| .py file | functions |

| -------- | -------- |

| \*\*`lib/paramters.py`\*\* | It defines basic parameters such as “ID versus molecular name“, “the size of lattice space” and so on. The defined variables (e.g., space) are referred such as "p.space" in executable programs (bin/). |

| \*\*`lib/colormap.py`\*\*| It defines colors. The color universal design was utilized for color blindness. |

| \*\*` control\_board\_\*.py `\*\*| Workspaces. They would be edited depending on further simulation and analyses. |

### Example

| diretory / file | contents |

| -------- | -------- |

| \*\*`control\_board\_example.py`\*\* | Executable programs. |

| \*\*`example\_lammpstrj`\*\*| CG.lammpstrj, PS.lammpstrj, CGPS.lammpstrj |

| \*\*`example\_edited`\*\*| Empty. |

| \*\*`example\_imgs`\*\*| Empty. |

It creates objects of the classes below. Each object ‘obj’ calls examples() to specify the target lammpstrj files (obj.exmaples()), and executes the method for processing (obj.run(), obj.repeat\_run()).

It calls the following programs.

* EditData

It converts the lammpstrj files (‘CG.lammpstrj’, ‘PS.lammpstrj’, and ‘CGPS.lammpstrj’; example\_data/lammpstrj) into intermediate data files (‘CG\_.pickle’, ‘PS\_.pickle’, and ‘CGPS\_.pickle’; example\_data/edited). Each pickle file has a dict variable with the following keys:

| key | value type | description |

| -------- | -------- | -------- |

| \*\*`mc\_step`\*\* | int | Sampled MC step. Final MC step in general. |

| \*\*`sampling\_frame`\*\* | int | Sampled time frame. Final time frame in general. |

| \*\*`region\_condensate\_in\_grid\_mesh`\*\* | dict[X], 'dilute' | Condensate region of X. Each condensate region is defined by the region over the half maximal levels of blurred Y. |

| \*\*`conc\_condensate`\*\* | dict[Y][X] | Conc of X in the condensate region Y. |

| \*\*`conc\_periphery`\*\* | dict[X] | Conc of X in the periphery region. The periphery region is defined by the region over d/2 distant from the center in the lattice space [d, d, d]. |

| \*\*` concs\_in\_grid\_mesh`\*\* | dict[X] | Conc of X in the 3d space (3d np.array, float). Grid locations of beads were blurred by the gaussian (sd: p.sigma in lattice unit). |

| \*\*` locs\_in\_grid\_mesh`\*\* | dict[X] | Conc of X in the 3d space (3d np.array, int). Grid locations of beads. |

| \*\*`rdf\_bins`\*\*| 1d np.array, int | Bins for radial distribution profile (RDP) (in lattice unit). |

| \*\*`rdf\_sampling\_frames`\*\*| list, int | Sampled frames for RDP |

| \*\*`rdf `\*\*| dict[X] | Radial distribution profile X (2d np.array, float) |

| \*\*`dir\_lammpstrj `\*\*| dir\_lammpstrj. |

| \*\*`filename\_lammpstrj`\*\*| filename\_lammpstrj. |

| \*\*`rotated\_region\_condensate\_in\_grid\_mesh`\*\* | dict[X] | Same as \*\*` region\_condensate\_in\_grid\_mesh `\*\* but rotated. |

| \*\*`rotated\_concs\_in\_grid\_mesh`\*\* | dict[X] | Same as \*\*`concs\_in\_grid\_mesh`\*\* but rotated. |

Here, X, Y ∈ [‘All’, ’CaMKII’, ‘GluN2B’, ‘STG’, ‘PSD95’]