**ManifoldEM 1.0.0**

ManifoldEM is a stand-alone MATLAB program to extract continuous conformational changes from single-particle cryo-electron microscopy data. It is developed in the Ourmazd research group at the University of Wisconsin Milwaukee (UWM) (https://sites.uwm.edu/ourmazd/). The underlying methodology with application to different experimental datasets is described in the following publications.

1. Dashti, A. et al. Trajectories of the ribosome as a Brownian nanomachine. *Proc Natl Acad Sci U S A* **111**, 17492-7 (2014).
2. Dashti, A. & Mashayekhi, G. et al. Retrieving Functional Pathways of Biomolecules

from Single-particle Snapshots. *Nature Communication,* (2020). (Under review)

If ManifoldEM is useful in your work, please cite the papers 1 & 2.

**Overview**

The software needs the Cryo-EM snapshots (aligned particle images) with the assigned orientations and defocus values (alignment file). For the preprocessing of the data and extracting the orientations and related parameters any software could be used. In this repository, a MATLAB function (getStarFile.m) is provided to extract the information from the .star file and put it in the appropriate format of ManifoldEM. As for the image stack the default setting is for .mrcs files. But such files/settings could be developed for any other data format to make the software compatible with.

Over the years, the MATLAB codes were developed in different MATLAB versions. The current distribution has been tested with MATLAB 2015b and 2019a 64 bit. Some functions used in ManifoldEM were part of standard MATLAB 2015, but moved to special toolboxes in the later versions.[[1]](#footnote-1)

**Known Issues**

Please note that there are some issues in the current ManifoldEM software, which are not yet fully resolved as of this release of MATLAB code (1.0.0) as follows:

* **Hyper parameter tuning**

There are multiple parameters being used in the ManifoldEM software. We have tried to come up with the set of parameters which works for most of the datasets. But further tuning of the parameters might give better results.

* **Automatic Outlier removal**

In some datasets, some outliers might appear in the manifolds which could affect the quality of conformational movies. Currently an automatic method is implemented which removes some of the outliers. But in some cases, we might need to remove the outliers manually (For more details refer to the tutorial).

* **Boundary Problem**

Overestimation of the occupancies near the boundaries of the occupancy map have been observed. So, the occupancies for up to 5% of the bins from the two ends of the conformational coordinate might not be reliable.

* **Propagation of 2D conformational coordinates across the whole S2**

In manifoldEM, first the 2D conformational changes are extracted in each viewing angle (projection direction), then they are mapped to 3D conformational changes. To do so, the conformational changes along the eigenfunctions of the manifold are extracted. Then the coordinates (CC) representing the similar motions from all projection directions are used to generate the 3D conformational movie. If there is one dominant and large conformational motion going on, this motion is most likely to be captured by the first eigenfunction of the manifold. But if there are subtle changes happening and/or there are more than one conformational coordinate or some other artifacts, the order of the eigenfunctions ranked according to the eigenvalue may be different in different projection directions. Also, the sense of the conformational changes may change due to sense-neutrality in eigen-decomposition. In these cases, we have following issues, which makes the propagation difficult:

* Duplicate CCs

Some of the generated 2D movies along different eigenfunctions might seem indistinguishable from one another. In these cases, we suggest selecting the eigenfunction with the higher corresponding eigenvalue. However, depending on the selected eigenfunction, the occupancy map and energy landscape might differ.

* Aberrant CCs

Sometimes the generated 2D movies along one eigenfunction looks like being a hybrid motion of two other eigenfunctions either in the same projection direction or neighboring one.

We have developed a couple of methods to propagate the conformational coordinates across the S2, but there is none which could be used for all the datasets in general and compete with the manual selection with acceptable accuracy.

Please note that, by wrong propagation of the conformational coordinate, the final 3D movie would be affected.

The software, code sample and their documentation made available on this website could include technical or other mistakes, inaccuracies or typographical errors. We may make changes to the software or documentation made available on its web site at any time without prior notice.

We assume no responsibility for errors or omissions in the software or documentation available from its web site.

1. You might encounter this issue when running great circle tessellation (selectGCs.m). This function uses kmeans which requires the Statistics and Machine Learning Toolbox in Matlab 2019a and higher. But also, this can be replaced by substitutes. [↑](#footnote-ref-1)