NaR Data Analysis: Mean First Passage Time

Giuseppe Sacco

(version June 2021)

1 Scientific background

The aim of the project is to get a better understanding of the properties of ion transport in nanotubes. In particular, in our simulation box, aqueous solutions of $NaCl$ are separated by a lipidic membrane, at whose center there is a carbon nanotube, permeable by ions and water (periodic boundary conditions are applied). In order to facilitate and get better statistics of the permeation events, the movement of one of the ions Na is constrained along the axis of the nanotube (z axis). Here we focus on the kinetics of the permeation of this ion of reference, called hereafter NaR.

2 Purpose of the program

The aim of the program *MFPT.py* is to perform data analysis of the trajectory of the NaR in order to get insights about the Mean First Passage Time (MFPT) of the transition of the ion between two local minima of the potental of mean force, situated just before and just after the entrance of the nanotube. Here the MFPT is operationally defined as the mean residence time in the starting minimum (outside the nanotube) before the transition to the other minimum.

3 Input and output

The program needs *Python 3.9* or any later version.

3.1 Modules required

The following modules are required:

- Built-in modules: *os, re, sys, typing, argparse, configparser, timeit, warnings, pickle*;
- Modules for scientific application: *numpy, scipy, matplotlib.pyplot*;
- Custom modules (the respective .py files must be included in the directory 'Libraries'):
  - *custom_plt.py* containing custom plotting functions for each plot;
  - *custom_errors.py* containing custom exceptions;
  - *custom_class.py* containing the definition of the class used to initialise the input parameters.

3.2 Input

The programs takes in input from the command line an INI file (option aliases are '-i','- -input','--input_file'; if not given the defalut is *input_file.ini*). The INI file must contain the following required parameters in the correct sections (time must be expressed in ps and distances in $\mathring{A}$):

[simulation params]

- *delta_z*: half-width of boxes corresponding to local minima of the potental of mean force;
- *lag_time*: time interval analyzed in order to determine if transition is valid;
- *acceptance_rate*: Transitions are considered valid only if, during the lag time, $n_{frameinbox}/n_{frame} \geq$ acceptance rate.

[files and directories]

- *pickle_directory*: name of the directory containing the .pickle files of the NaR trajectory;
- *output_file*: name of the file with the results of the data analysis.

[plotting params]

- *plot_graphs* (bool): choose if graphs are to be created;
- *save_figures* (bool): choose if graphs are to be saved;
- *nbins_t*: number of bins along the time axis;
- *n_bins*: number of bins along the z axis;
- *savefig_directory*: name of the directory where figures are saved (if not existing it is created);
- *add_savefig_name*: Figures are saved as: graphtype + filename + add_savefig_name + '.png'.

Booleans can be: yes/no, on/off, true/false, 1/0 (case insensitive).

- ☐ Working directory
  - *Libraries*
    - *custom_plt.py*
    - *custom_errors.py*
    - *custom_class.py*
  - pickle_directory (u)
    - Pickle files (u)
  - input INI file (u)
  - *MFPT.py*

Elements given by the user are marked with '(u)' above. The program looks in the given *pickle_directory* for files that match the pattern *00-00-00-00_0M_NAR.pickle* where 0 indicates any digit (one or more digit can precede 'M').

3.3 Output

After the correct execution of the program, the file *output_file* is produced, containing the results of the data analysis for each .pickle file in the *pickle_directory*.

Optional: The following graphs are created:

- NaR trajectory;
- NaR histogram z counts;
- Potential of mean force;
- Transition times distribution.

Output file is saved in the working directory, graphs in the *savefig_directory*. If files with the same name are present, they are subscribed.

From the obtained distribution of the transition times, it is also computed the fitted value and standard deviation of τ through the function $Ae^{t/\tau}$, where t is the transition time.

Results printed are: number of transition events, total residence time in the starting box, MFPT as (total residence time/ number of transition events), MFPT as mean of the transition times, tau fitted, standard deviation of tau, greatest and lowest transition time.

4 Data analysis

The following process is executed for each .pickle file given.

The data of the trajectory of the NaR is extracted from the .pickle file. Two boxes are defined whose center corresponds to the local minima of the potental of mean force right before and after the nanotube entrance. The half-widht of these boxes is *delta_z*. Following the ion trajectory, each transition time is the computed as the residence time in the starting box before the transition to the second box. Transitions are considered valid only if, during the lag time, $n_{frameinbox}/n_{frame} \geq$ acceptance rate.