

CMSC6950 Project Report: Tidynamics

Karina Barcelos

Spring - June 2021

1 Introduction

Tidynamics [1] is a tiny and simple package that calculates mean-square displacements (MSD) of a trajectory and correlation functions of a input data using the Fast Correlation Algorithm. [2] This package is especially beneficial for quantitatively evaluating dynamics of stochastic and molecular simulations from numerical trajectories. The MSD is a deviation of the position of a certain particle regarding a reference position at time intervals. Whereas the other useful computational tools of this software are the autocorrelation function (ACF) and correlation function, in which both are conceptually similar mathematical representations of similarity degree. ACF uses a variable with itself between two successive at two points in time (time lags) and correlation function is the relationship between two different time series at successive time lags.

The data format input and output of tidynamics [1] is array, simply depending only on Python and NumPy. The NumPy arrays where the first index indicates the timestep and the second index, in case of correlation function between two time-series, shows the spatial coordinates of the trajectory. Both ACF and MSD are defined by the respective equations below in tidynamics algorithm [1], where the angle brackets denote an average over time. [2]

$$C_{AB}(\tau) = \langle A(0)B(\tau) \rangle, \quad (1)$$

where $A(0)$ and $B(\tau)$ represent the input data quantity A and the quantity at some later time B, respectively.

$$MSD(\tau) = \langle (x(\tau) - x(0))^2 \rangle, \quad (2)$$

where $x(\tau)$ and $x(0)$ refer to current and reference positions, respectively.

The aim of this project was to implement two computational tasks of dynamical systems from the chosen open source package tidynamics[1]. For the first task, the ACF was computed for the bond length of C-H and C=O from a certain Tyrosine over 700 ns of my Molecular Dynamics (MD) simulation of a certain protein. It has been chosen to work with a more applicable data from a MD simulation to test the function in a real system. For the second task, the random walk values and its mean-square displacement were generated. Several modules were utilized during the project, including numpy, pandas, math, sys, and matplotlib.

2 Results

The results will be reported for use of ACF and MSD from tidynamics package[1].

2.1 Task 1: CF of C=O and C-H bond length over MD time

The objective of this task was to analyze the C=O and C-H bond stretching motion using its time ACF. The input data were obtained from a long MD simulation trajectory with total duration of 1.5

ms of a Tyrosine amino acid of a certain protein. A MD simulation is a computer simulation method that allows to predict and visualize at atomistic level the motion of a system (often a protein) of time evolution. The time evolution herein was in nanosecond (ns) from a period of 800-1500 ns to ensure data analysis in the protein equilibrium. The final input files had the C=O and C-H bond lengths in nanometer (nm) over 700 ns time series and an autocorrelation function calculated from todynamics [1] in a time-series. After load the data, time and length were converted into a dataframe. The result after applying the ACF in C-H and C=O lengths by saving into the same dataframe. The resulting txt file was a intermediate file for further plotting the Figure 1. The ACF showed near zero positive autocorrelation for those length data of those types of bonds, which do not vary drastically over long period of simulation.

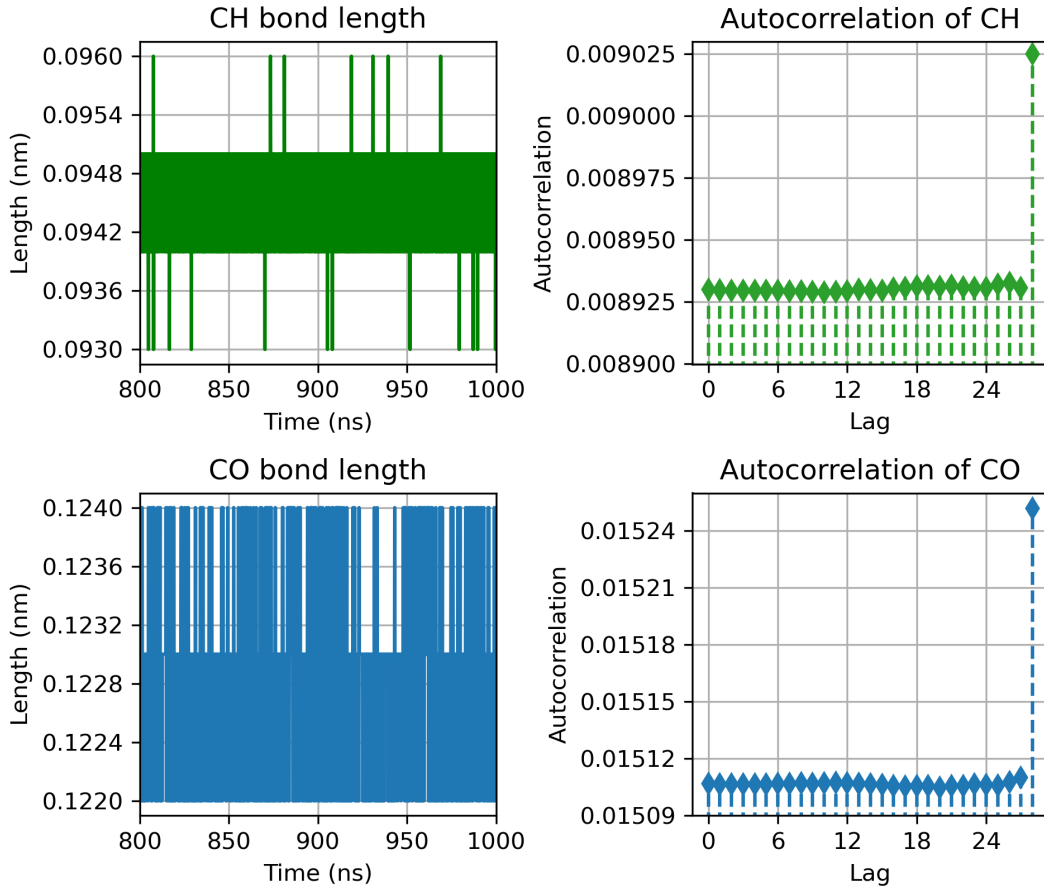


Figure 1: C-H and C=O bond length from a Tyrosine amino acid over MD simulation time and their ACF after 800 ns to ensure the system is in equilibrium.

2.2 Task 2: MSD of a random walk

For the second computational task, it was generated a random walk of 1000 steps. By starting with an example with a certain walker at the origin, meaning $(x,y) = (0,0)$, it then chooses a random direction between $0 \leq \theta \leq$ for the first step of the walk. For that, it is necessary to determine a random angle

value number using a random-number generator in order to the sequence of random numbers do not probably repeat in each further random walk step. In case of this task script, a random value was obtained after a random angle was applied in x and y coordinates, such as $x = \cos \theta$ and $y = \sin \theta$. For each of N step, those prior values were calculated to the cumulative values for x and y being updated for each step. Then the (x,y) values were saved into a list and after a dataframe. The MSD were applied in x and y as it is a 2D position dataset. The Figure 2 was then plotted from another script, containing the 2D random walk coordinates with start and end markers and its numerical and theoretical MSD curves. The theoretical value is the time (N=1000) multiplied by 2 according to 2 since it is a 2D random walk.

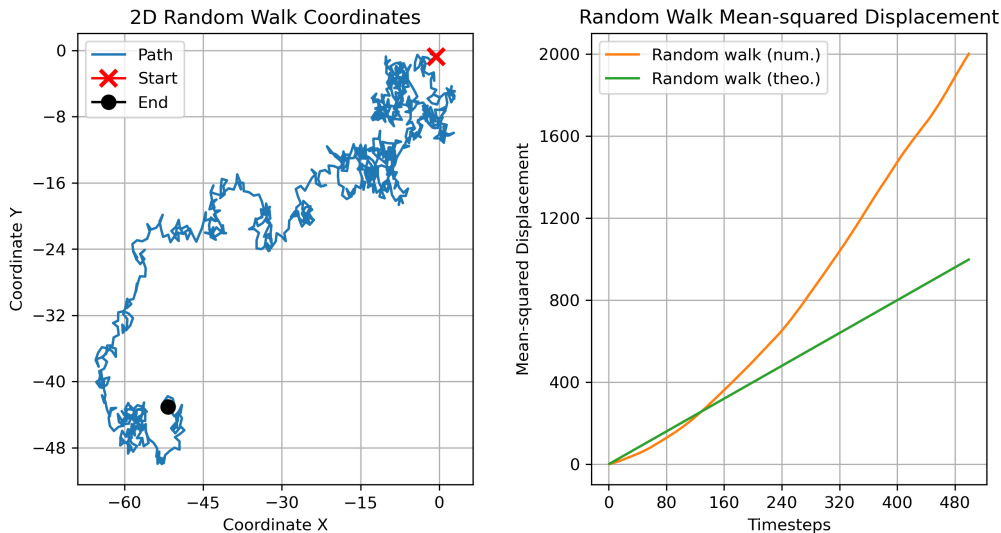


Figure 2: 2D random walk in a (x,y) coordinates and its mean-squared displacement.

3 Conclusions

Therefore, two computational tasks were performed using tidynamics package to calculate its common tools, such as ACF and MSD, and plot their results in different visualizations. The first task provided a visualization of the bond length of C-H and C=O and its ACF over 700 ns a MD dynamics simulation, data obtained from a trajectory of a Tyrosine. It required tidynamics, matplotlib, and pandas python libraries. Whereas the second task, MSD of a random walk was plotted, in which x and y coordinates were applied cosine (theta) and sine (theta) while theta was randomly sorted and multiplied by 2π . It was utilized numpy, math, tidynamics, and pandas libraries.

References

- [1] Pierre de Buyl. tidynamics: A tiny package to compute the dynamics of stochastic and molecular simulations. *Journal of Open Source Software*, 3(28):877, 2018.
- [2] Gerald R Kneller, Volker Keiner, Meinhard Kneller, and Matthias Schiller. nmoldyn: a program package for a neutron scattering oriented analysis of molecular dynamics simulations. *Computer physics communications*, 91(1-3):191–214, 1995.