BIT - Bayesian Inference Team

Samudrajit Thapa,[1] Seongyu Park,[2] Yeong-Jin Kim,[2] Jae-Hyung Jeon,[2] and Michael A. Lomholt[3,*]

[1]*Institute for Physics and Astronomy, University of Potsdam, D-14476 Potsdam-Golm, Germany*
[2]*Department of Physics, Pohang University of Science and Technology, Pohang 37673, Korea*
[3]*PhyLife, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark*

(Dated: November 4, 2020)

Method: **Bayesian inference** using **annealed importance sampling** to (attempt to) sample from the posterior distribution.

1D TRAJECTORIES

For the 1D trajectories we use Bayesian inference with a uniform prior for the five different models. For the parameters we have chosen priors according to the provided information (e.g., for α we chose a prior uniform between 0.05 and 1 for ATTM and CTRW, and uniform between 0.05 and 2 for FBM, LW and SBM). For Task 3 we furthermore included a uniform prior for the changepoint along the full trajectory, and then independently inferred the model for each of the two segments (we ended up not including the option of no changepoint, since this gave us a lower score on the AnDi leaderboard, probably due to trajectories with a changepoint being wrongly inferred as without). For Task 1 we submitted the posteriors median α; for Task 2 we submitted the five model probabilities; for Task 3 we picked the most probable model for each segment instead and submitted together with the medians of the time of the changepoint and α-values. For all trajectories we discarded the data beyond 200 time points to save computational time.

LIKELIHOOD FUNCTIONS

All non-LW likelihood functions (also above 1D) were calculated as products of Gaussian distributions $P(\Delta\vec{x}_{i+1}|\Delta\vec{\boldsymbol{x}}_i)$ for each step $\Delta\vec{x}_{i+1}$ given the collection of previous step $\Delta\vec{\boldsymbol{x}}_i$ (and model parameters):

$$P(\Delta\vec{x}_{i+1}|\Delta\vec{\boldsymbol{x}}_i) \sim \text{Gauss}(\vec{\mu}_{i+1}, \vec{V}_{i+1}) \tag{1}$$

where $\vec{\mu}_{i+1}$ and $\vec{V}_{i+1}$ are the mean and variance of the steps as given by the individual models. For ATTM and CTRW the number and length of time intervals are included among the model parameters, and integrated over during the Monte Carlo sampling of the posterior. For ATTM, CTRW and SBM the variance for each step are then a priori as given by the organizers when defining the models, but to include measurement noise they are modified as explained in Appendix A of [1] (note that the formulas and derivations in that appendix generalizes straightforwardly to the situation where the variance depends on time (i.e., step number)). For FBM the method presented in [2] was used. The LW likelihood function in 1D was implemented as a hidden Markov model following [3] with a power law waiting time for each of the two directions generated using a trick presented at the beginning of Section II in [4]. For LW measurement noise was not included correctly, but instead included approximately by allowing diffusion during each step.

2D TRAJECTORIES

The 1D approach was generalized straightforwardly to 2D with one major exception: we gave up on LW since it now included a continuum of possible directions instead of just two. The annealed importance sampling was thus run with only the other four models. After the sampling a Bayesian goodness-of-fit test similar to [5] was performed: For a few parameter samples a Rosenblatt transformation [6] of the trajectory data was performed. If the fit was good, then this should give transformed data with values uniformly distributed in unit cubes. This was tested by comparing Bayesianly with densities perturbed on the unit cube by trigonometric functions. If any Bayes factor between the uniform distribution and an alternative perturbed model was larger than $10^{3.5}$ in favor of the alternative model the test was failed. In this case the LW model was selected (assigning a posterior probability of one to this model) and the value of α was estimated by fitting a straight line to the first nine points of the log-lagtime vs. log-estimated time averaged mean square displacement. No restriction was put on the value of α, in particular it was not restricted to $1 \leq \alpha \leq 2$.

3D TRAJECTORIES

The 3D likelihood functions was obtained by generalizing the 2D approach to one more dimension.

ANNEALED IMPORTANCE SAMPLING

Posterior sampling was done using Monte Carlo methods, with initial samples being obtained using annealed importance sampling [7]. Monte Carlo steps were per-

formed using either slice sampling or the Metropolis algorithm [8]. The latter were used for parameters that can affect the dimensionality of the parameter space, for instance by switching between the five models. Any new parameters in such an attempted step are assigned random values according to their prior. During annealing a walker with small weight was discarded at each temperature change, while the highest weight walker was duplicated, halving the corresponding two weights. For Task 3 a recombination step was included as a possible Monte Carlo move, where the model for individual segments can be switched between the current samples (also called a crossover in genetic algorithm terminology).

The lengths of the time intervals in ATTM and CTRW were also parameters that were included in the Monte Carlo sampling. Unfortunately, convergence was often slow for these parameters, so we were not able to sample them well. We think this is the main problem that limited our approach, at least in 1D , with the need to discard data beyond 200 time points as a second major issue.

IMPLEMENTATION

The code is implemented in Matlab. It can be found at [9]. Instructions for running the code is in the file `README.md`.

* `mlomholt@sdu.dk`

[1] J. Krog and M. A. Lomholt, Phys. Rev. E **96**, 062106 (2017).
[2] J. Krog, L. H. Jacobsen, F. W. Lund, D. Wüstner, and M. A. Lomholt, Journal of Statistical Mechanics: Theory and Experiment **2018**, 093501 (2018).
[3] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, 2007), 3rd ed.
[4] I. Goychuk, Phys. Rev. E **80**, 046125 (2009).
[5] I. Verdinelli and L. Wasserman, Ann. Statist. **26**, 1215 (1998).
[6] M. Rosenblatt, Ann. Math. Statist. **23**, 470 (1952).
[7] R. M. Neal, Statistics and Computing **11**, 125 (2001).
[8] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, 2003).
[9] `https://github.com/mlomholt/andi`.