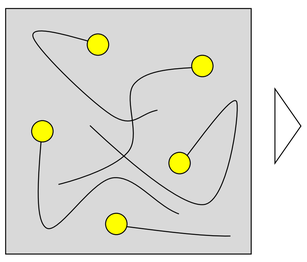
Estimating Errors in Diffusion Coefficients From Molecular Dynamics Simulations Without Replicates

Benjamin Crysup
Advised by Dr. Sachin Shanbhag
Department of Scientific Computing, Florida State University, Tallahassee, Florida

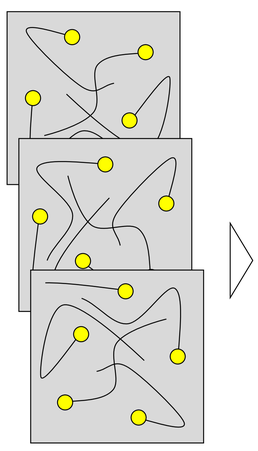
Motivation

Molecular dynamics simulations trace out the paths of atoms over time, and from these traces statistical mechanics can be used to glean various bulk properties.



Diffusion
Temperature
Modulus

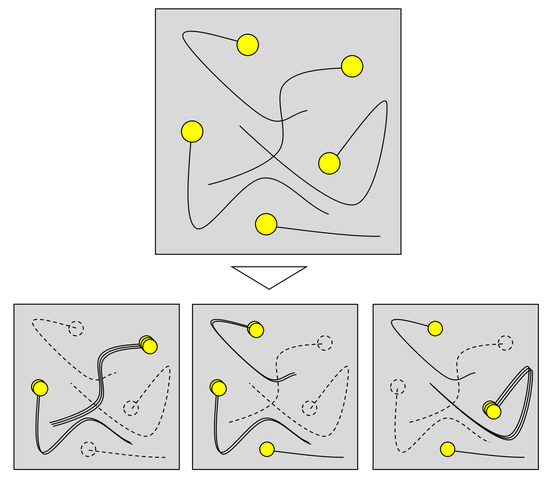
However, a single trace only provides a value, and leaves no indication of how much error is present. The proper way to estimate error is to rerun the simulation with randomized initial states, and take the standard deviation of the results. But, when each individual trace could take more than a day to run, then this process becomes painful very quickly.



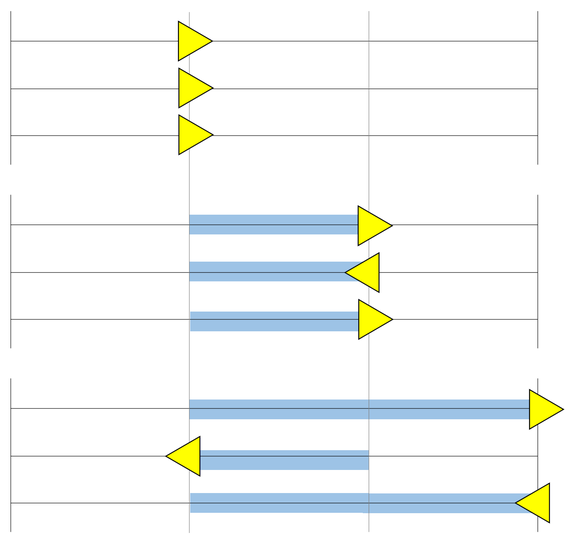
Diffusion ± Error

Method

Enter bootstrapping. The idea here is to invent new simulation traces by taking random samples with replacement of the individual particle trajectories, and then treat those new traces as independent samples in order to estimate the error.



In order to test this, a simplified model has been proposed. All of the particles in the simulation start at position 0 and start with a velocity of 1. Every time unit, each particle either reverses or maintains its velocity.



This system was run 1000 times each for different particle counts and time scales, and the diffusion coefficient for each run was extracted by performing a linear regression on mean square displacement:

$$MSD(lag) = \left\langle \left\langle \left(\overrightarrow{x_p(t_0 + lag)} - \overrightarrow{x_p(t_0)} \right)^2 \right\rangle_{t_0} \right\rangle_p$$

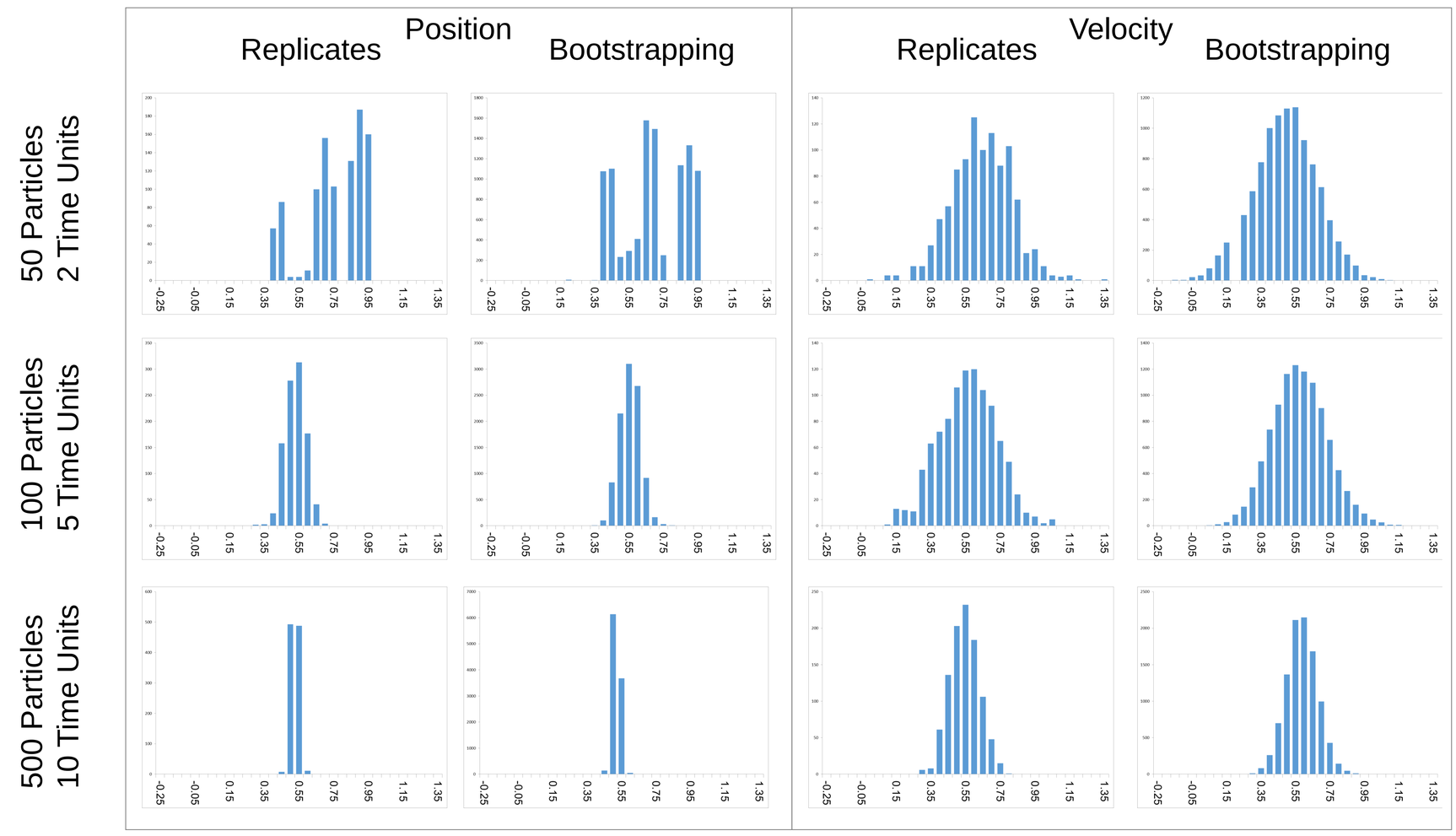
and also by integrating the velocity autocorrelation:

$$VAC(lag) = \left\langle \left\langle \overrightarrow{v_p(t_0 + lag)} \cdot \overrightarrow{v_p(t_0)} \right\rangle_{t_0} \right\rangle_p$$

The distributions of the extracted coefficients are displayed to the right under "Replicates".

Then, for each particle count and time scale, bootstrapping was performed on one of the runs. The distributions of the diffusion coefficients extracted from the bootstrap runs are displayed to the right under "Bootstrapping".

Results (Diffusion Coefficient Distributions)



Discussion

For each distribution calculated by running independent simulations, the corresponding distribution created using bootstrapping looks very similar. This suggests that, at least for this system, bootstrapping can be a very reliable method of estimating error. However, this simple system is slightly unrealistic, as each particle's trajectory was completely independent of the other particles. Future research will examine how bootstrapping handles a dense, interacting system.