Single vs. Double Precision in MD Simulations: Correlation Depends on System Length-Scale
Anqi Zou¹, Tyson J. Lipscomb¹, and Samuel S. Cho¹,²
¹Departments of Computer Science and ²Physics
Wake Forest University, Winston-Salem, NC, USA

Introduction to Molecular Dynamics (MD) Simulations
Understanding how biomolecules transition from an unstructured unfolded state to a well-defined structured folded state is one of the grand challenges in biology because biomolecules sometimes unfold to cause disease such as Alzheimer’s and Parkinson’s disease. Other times, the biomolecules themselves cause a disease effect, such as HIV’s protease, a protein that is integral in HIV replication. The structure of biomolecules can be viewed at a microscopic scale using molecular dynamics simulations (MD) of the underlying conformation. To accurately calculate the inherent calculations, the representation of numbers is essential. By current standards, researchers can directly observe only the unfolded and folded states, and MD simulations can help fill in the critical blanks.

Length- and Time-Scales of MD Simulations

Single vs. Double Precision Overview
It is well known that computers can only represent numbers using a limited number of significant figures. That limited number, such as it cannot be represented as an exact mathematical value is computer memory. The discrepancy caused by this representation limitation is called “round-off error.” Other computers have the capacity to represent numbers in the type of precision, single (32-bit) and double (64-bit) precision.

Single precision (32 bits) to represent the significant digits and can represent at about 7 decimal digits, while double precision used 64 bits to represent about 15 decimal digits. The limit of available bits limits the precision of a numerical representation. For example, it is represented as 3.141592653589793 (double precision). In general, using double precision reduces the consequence of roundoff error. Memory and computational cost associated with double precision could be very expensive, so single precision calculations are preferred when the precision of the calculation is not required to be an error.

Motivation
To evaluate how single or double precision, and to determine which one is more suitable for various computer architectures, and the accuracy of single precision can be compared to the accuracy of the final result.

Single vs. Double Precision MD Simulations Give Different Trajectories for Small Systems

We performed coarse-grained MD simulations of the ribosome, a biomolecular macromolecular machine, using three different precisions: single (32-bit) and double (64-bit) precision. Single precision (32 bits) to represent the significant digits and can represent about 7 decimal digits, while double precision used 64 bits to represent about 15 decimal digits. The limit of available bits limits the precision of a numerical representation. For example, it is represented as 3.141592653589793 (double precision). In general, using double precision reduces the consequence of roundoff error. Memory and computational cost associated with double precision could be very expensive, so single precision calculations are preferred when the precision of the calculation is not required to be an error.

Coarse-Grained Simulations
We performed coarse-grained simulations of the ribosome, a biomolecular macromolecular machine, using three different precision: single (32-bit) and double (64-bit) precision. Single precision (32 bits) to represent the significant digits and can represent about 7 decimal digits, while double precision used 64 bits to represent about 15 decimal digits. The limit of available bits limits the precision of a numerical representation. For example, it is represented as 3.141592653589793 (double precision). In general, using double precision reduces the consequence of roundoff error. Memory and computational cost associated with double precision could be very expensive, so single precision calculations are preferred when the precision of the calculation is not required to be an error.

Structural Analysis of Trajectories

We calculated the Pearson correlation coefficient (R) of the coordinates from our MD simulations. When R = 1, the correlation is perfect, meaning that any data point can be accurately represented using a line. When R = 0, there is no correlation, meaning that no data point can be accurately represented using a line. For each residue, R values range from 0.0 to 1.0. For different systems, there is a clear relationship between the structure similarity and system size.

Conclusions

We performed MD simulations using single vs. double precision computer, and we quantify the difference between single vs. double precision frames using three different structural measures: difference in root-mean-square (RMSD), difference in root-mean-square deviation (RMSD), and structural overlap (g). For each measure, R values range from 0.0 to 1.0. For different systems, there is a clear relationship between the structure similarity and system size.

Future Directions

• Round-off error correction for small systems
• Evaluate other MD simulation types to determine the size-dependent accuracy and precision limitation

Acknowledgements

• (TIG) Wake Forest University Computer Science Graduate Fellowship for Anqi Zou
• (SISC) NVIDIA Academic Partnership
• (WFU) Wake Forest University Science Research Fund