Contribution ID: **121** Type: **Poster**

Preliminary Results of Vectorization of Density Functional Theory calculations in Geant4/V for amino acids

Tuesday, 25 October 2022 11:00 (30 minutes)

Density Functional Theory (DFT) is an extended ab initio method used for calculating the electronic properties of molecules. Considering Hartree Fock methods, the DFT offers appropriate approximations regarding the time calculations. Recently, the DFT method has been used for discovering and analyzing protein interactions by means of calculating the free energies of these macro-molecules from short to large scales. However, calculating the ground-state energy by DFT for many-body systems of molecules as proteins, in a reasonable time with enough accuracy, is still a very challenging and intensive task for the CPU's resources.
On the other hand, Geant4 is a toolkit for simulating the effects of energy through matter and the nature of materials with a wide range of specialized methods that include DNA and protein exploration. Unfortunately, the execution time to obtain an effective protein analysis is still a strong restriction for CPU processors. In this sense, the GeantV project searches to exploit the vectorization of CPUs, designed to tackle the problem of intensive charge of calculus at the cores of CPUs. In this work, we present the preliminary results of the partial implementation of the DFT in the Geant4 framework and the vectorized GeantV project. We show the advantages and the partial methods used for vectorizing several sub-routines in the calculus of ground-state energy for some amino acids and some molecules.

Significance

The novelty of this project consists of the addition of the DFT method to the workflow implemented in Geant4 with an extended analysis dedicated to proteins. Besides, with the incorporation of new strategies to vectorize several applications on modern CPU proposed by the GeantV project in HEP, we aim to include these advances to reduce the execution time of complex processes as the DFT calculations, following the philosophy of modern parallelism.

Finally, We would like to innovate with this work, expanding the aims and scopes of the GeantV project over the calculation of molecular structures at CPU with enough speedup and efficiency, contributing with part of libraries for fast simulation techniques. Therefore, we hope the final results will provide a straightforward approach to methods in software for parallelization.

References

• Collaboration at the project: GeantV. Results from the Prototype of Concurrent Vector Particle Transport Simulation in HEP: https://doi.org/10.1007/s41781-020-00048-6
• Vectorization techniques for probability distribution functions using VecCore: https://doi.org/10.1088/1742-6596/1525/1/012106
• Vectorization techniques for probability distribution function using Vec-

Core, the 19th-2019 ACAT conference: https://indico.cern.ch/event/708041/contributions/3272109/attachments/1810083/2955720/posterf.pdf
• Hot Spots & Hot Regions Detection Using Classification Algorithms in BMPs Complexes at the Protein-Protein Interface with the Ground-State Energy Feature: https://doi.org/10.1007/978-3-031-07750-0_1

Experiment context, if any

This work is inspired in Geant4-DNA project: http://geant4-dna.in2p3.fr/index.html Furthermore, this work is intended to be an expansion of the GeantV project: https://geant.web.cern.ch/geant/

Primary author: CHAPARRO AMARO, Oscar Roberto (Instituto Politécnico Nacional. Centro de Investigación en Computación)

Co-author: Prof. MARTÍNEZ CASTRO, Jesús Alberto (Instituto Politécnico Nacional. Centro de Investigación en Computación.)

Presenter: CHAPARRO AMARO, Oscar Roberto (Instituto Politécnico Nacional. Centro de Investigación en Computación)

Session Classification: Poster session with coffee break

Track Classification: Track 2: Data Analysis - Algorithms and Tools