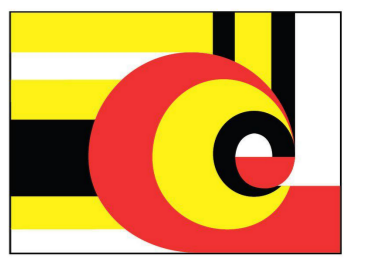# Preliminary Vectorization Results for Ground State Energy calculations on amino acids using the Density Functional Theory in Geant4/V

Oscar R. Chaparro Amaro · J. Martínez-Castro

Instituto Politécnico Nacional, Centro de Investigación en Computación, México.
Email-contact: ochaparroa2019@cic.ipn.mx

Introduction to the DFT Method

Kohn-Sham energy system approximation for crystal structures of amino acids [1]:

$$\mathbf{F}_i C_i = (\mathbf{T}_i + \mathbf{V}_i + 2\mathbf{J}_i + \mathbf{K}_i)C_i = \epsilon_i \mathbf{S}_i \qquad (1)$$

Linear Combination of atomic orbitals and STO-NG, centered at Nuclei coordinates [2]. Using the Hartree-Fock approximation (one-electron orbitals), the matrices dimensions are: $H_2O = 7$, $C_6H_6 = 36$ and Lysine $= 60$.

$$|\psi_{k,m,n}(\mathbf{r_A})\rangle \approx \sum_{\eta=1}^{\infty} C_\eta M_\eta AO_\eta = \sum_{\eta=1}^{\infty} C_\eta M_\eta (r - A_r)^a e^{-\alpha(r-Pr)^2} \qquad (2)$$

Overlap Int. S_{ij}:

$$S_{ij} = \langle \psi_i(\mathbf{r_A})|1|\psi_j(\mathbf{r_B})\rangle =$$
$$\int_{-\infty}^{\infty} \psi_i(\mathbf{r_A}) \cdot \psi_j(\mathbf{r_B}) dr \qquad (3)$$

One-electron expected kinetic energy Int. T_{ij} [4]:

$$T_{ij} = \left\langle \psi_i(\mathbf{r_A}) \left| -\frac{1}{2}\nabla^2 \right| \psi_j(\mathbf{r_B}) \right\rangle =$$
$$-\frac{1}{2}\int_{-\infty}^{\infty} \psi_i(\mathbf{r_A}) \nabla^2 \psi_j(\mathbf{r_B}) dr = \rightarrow \qquad (5)$$

The product Basis funct. $E = e^{-\lambda_{\eta\varsigma}(r-Pr)^2}$ and, summing over N basis [3]:

$$O_{r(i,a),(j,b)} =$$
$$\int_{-\infty}^{\infty} (r - A_r)^a (r - B_r)^b E dr \qquad (4)$$

$$T_{ij} \approx T1(O_r) + T2(O_r) + T3(O_r) \qquad (6)$$

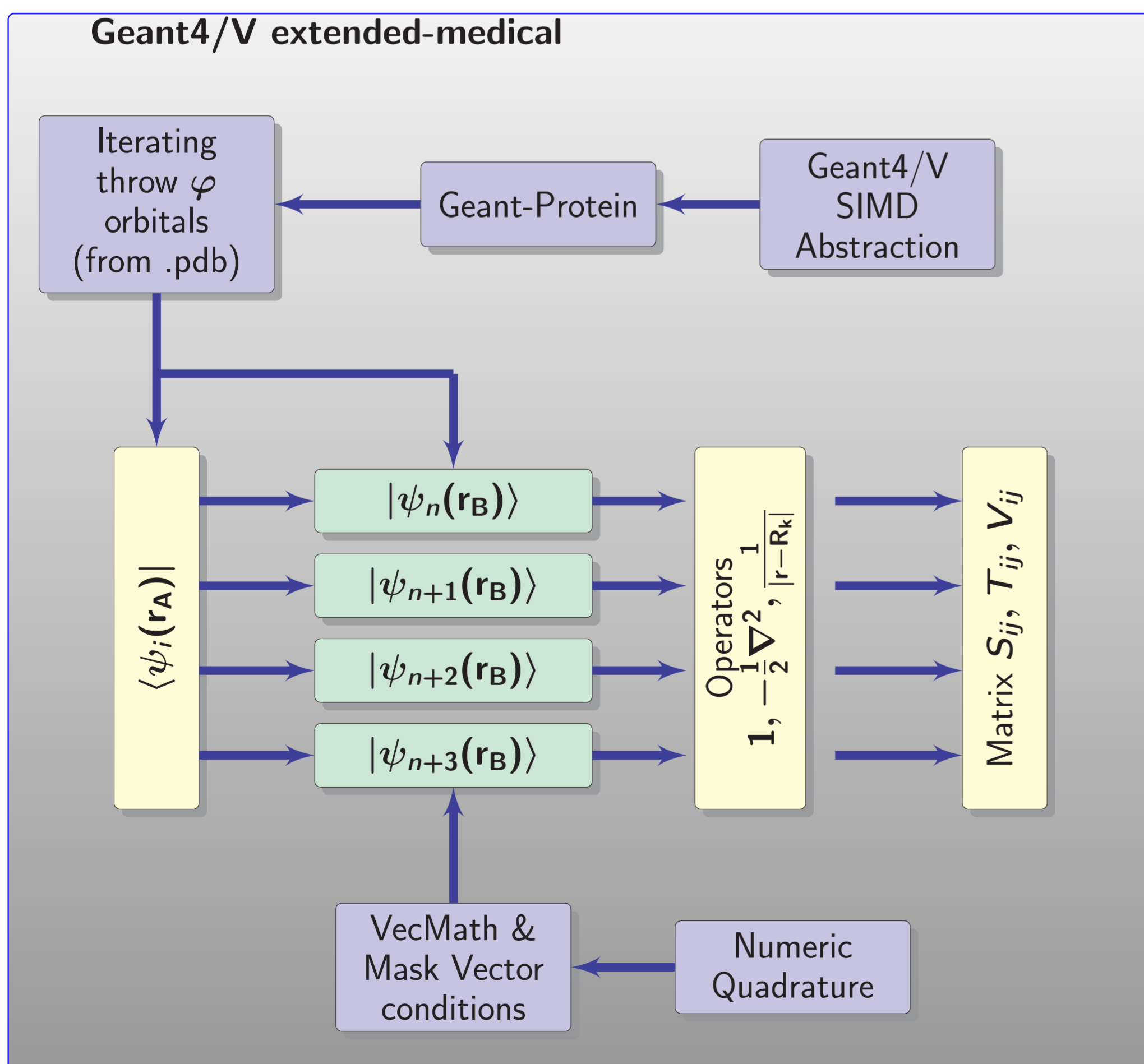One-electron expected Nuclei attraction energy Int: V_{ij}:

$$V_{ij} = \left\langle \psi_i(\mathbf{r_A}) \left| -\sum_{k=1}^{N_n} \frac{Z_k}{|\mathbf{r} - \mathbf{R_k}|} \right| \psi_j(\mathbf{r_B}) \right\rangle =$$
$$\sum_{k=1}^{N_n} Z_k \int_{-\infty}^{\infty} \psi_i(\mathbf{r_A}) \left(\frac{1}{|\mathbf{r} - \mathbf{R_k}|} \right) \psi_j(\mathbf{r_B}) d\mathbf{r} \rightarrow \qquad (7)$$

Gauss Chebyshev quadrature [5]:

$$F_m(a) = \frac{1}{2}\int_0^1 t^{2m}e^{-at^2}dt \qquad (8)$$

Vc/VecCore on Geant4/V

Geant-Protein is derived from the Geant4-dna project [6,7]. The vectorization applies Vc library from GeantV for vector operations in SIMD algorithms [8].

General Vectorization strategy. Mask Vector classes divide the input vector according to specific conditions (Maskfull for example). The run-time performance gain is mainly dependent on the code struct (numeric quadrature).

Partial Results
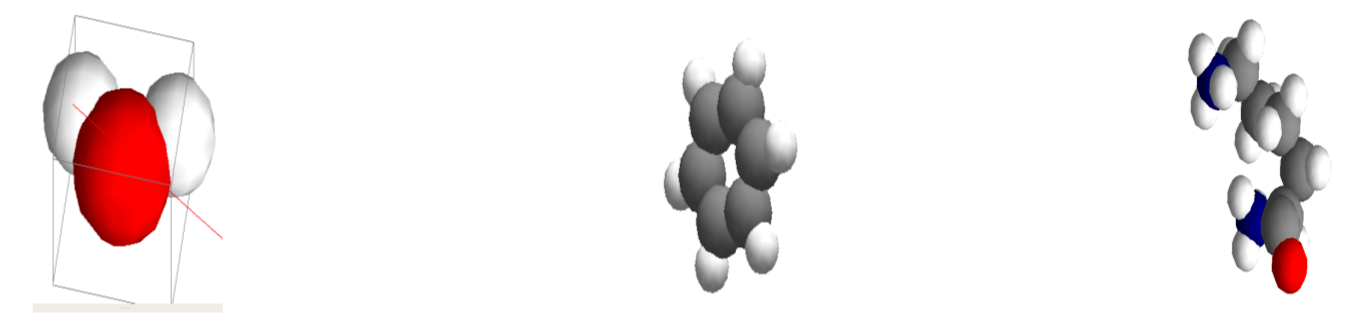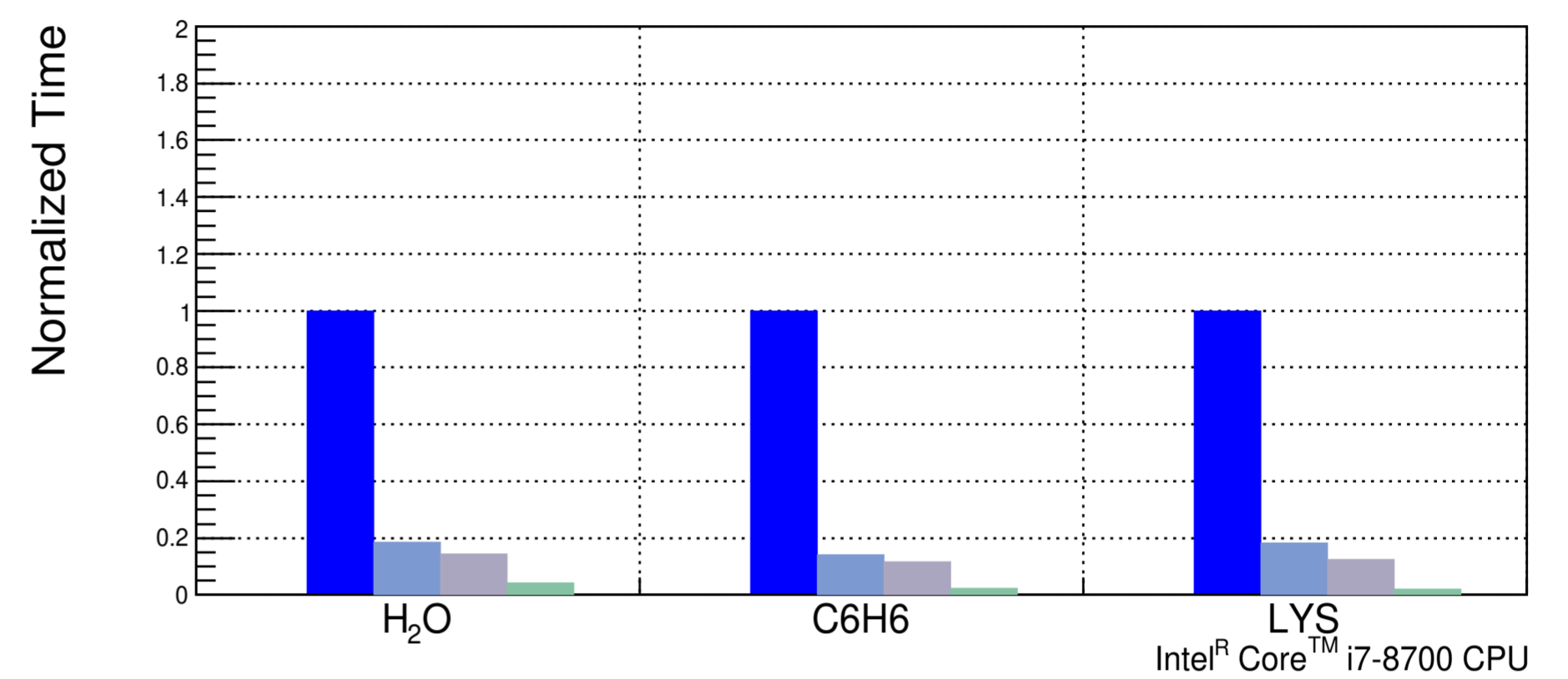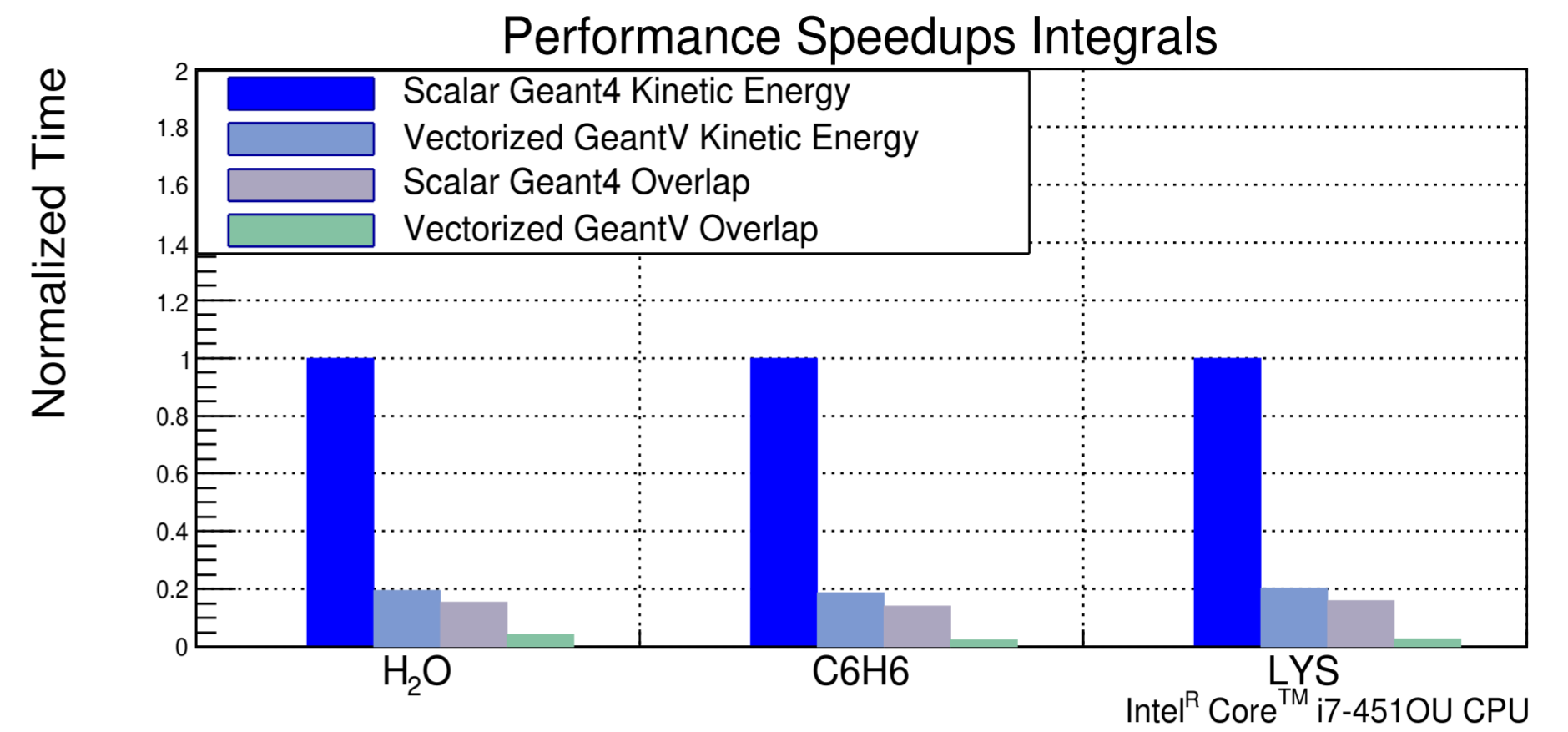
One hundred average performance between Geant4 and GeantV implementations. Speedups are normalized over the execution times of the scalar Geant4 kinetic energy matrices. The average normalized time for the scalar Geant4 electron-nuclei attraction energy matrices is **3913** for H_2O, **19729** for C_6H_6 and **43266** for the Lysine.

Future Work

-Double-electron Repulsion integral J_{abcd} of the form:

$$[ab, cd] = \left\langle \psi_a(\mathbf{r_A})\psi_b(\mathbf{r_A}) \left| \frac{1}{|\mathbf{r} - \mathbf{R}|} \right| \psi_c(\mathbf{r_B})\psi_d(\mathbf{r_B}) \right\rangle =$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_a(\mathbf{r_A})\psi_b(\mathbf{r_A}) \left(\frac{1}{|\mathbf{r} - \mathbf{R}|} \right) \psi_c(\mathbf{r_B})\psi_d(\mathbf{r_B}) d\mathbf{r_A} d\mathbf{r_B} \qquad (9)$$

-The Quadrature will use the Obara-Saika recurrence formula [9]:

$$[a(r)(b(r) + 1_r), c(r)d(r)] =$$
$$[(a(r) + 1_r)b(r), c(r)d(r)] + (\mathbf{A}_r - \mathbf{B}_r)[a(r)b(r), c(r)d(r)] \qquad (10)$$

-Applying Exchange correlation and local density approximations.
-Fully and partial Vectorization (Attraction, repulsion and exchange terms) using the SIMD acceleration (pseudo-vectorized algorithms).

References

[1] Stephen Fox, Jacek D. et al. Density functional theory calculations on entire proteins for free energies of binding: Application to a model polar binding site. Proteins-structure function and bioinformatics, 82:3335-3346, 2014.
[2] A. Beiser, S. Mahajan and S.R. Choudhury. Concepts of Modern Physics. Tata McGraw-Hill, Higher Education, 2003. ISBN 9780070151550.
[3] Toshikatsu K. and Tatsuya U. Approximate interaction energy in terms of overlap integral. The Journal of Chemical Physics, 85(3):14331437, 1986.
[4] Peter M.W. Gill. Molecular integrals over gaussian basis functions. Advances in Quantum Chemistry, 25(25):141205, 1994. S0065-3276(08)60019-2.
[5] José M., Pérez J. A simple, efficient and more reliable scheme for automatic numerical integration. Computer Physics Communications 77 (1993) 46-56.
[6] S. Agostinelli et al. GEANT4: A simulation toolkit. Nucl. Instrum. Meth., A506: 250303, 2003.
[7] S. Incerti, G. Baldacchino, M. Bernal et al. The Geant4-DNA Project. International Journal of Modeling, Simulation, and Scientific Computing Vol. 01, No. 02, pp. 157-178 (2010).
[8] Amadio, G., Ananya, A., Apostolakis, J. et al. GeantV. Comput. Softw. Big. Sci. 5, 3 (2021).
[9] Martin H. G. and John A. Pople. J. A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. Chem. Phys. 89, 5777 1988.