Open Access | https://doi.org/10.25080/dvta2583

SciPy 2025
July 7 - July 13, 2025

Proceedings of the 24th
Python in Science Conference
ISSN: 2575-9752

Quantum Chemistry Acceleration: Comparative Performance Analysis of Modern DFT Implementations

Kyohei Sahara[1]

[1] Kyoto University

Abstract

This article examines the acceleration of quantum chemistry calculations through modern implementations of Density Functional Theory (DFT). We provide a comparative performance analysis between traditional frameworks and advanced implementations, demonstrating computational efficiency gains. Applications in electrolyte membrane structure analysis illustrate the practical benefits of these performance improvements. Our benchmarking results show how transitioning from traditional implementations to optimized frameworks enables more extensive quantum chemistry simulations within reduced timeframes, with substantial speedups obtained through modern code optimization techniques in Python-based quantum chemistry environments.

Keywords quantum chemistry, density functional theory, performance analysis, computational chemistry

Published Jul 10, 2025

Correspondence to
Kyohei Sahara
schwalbe10@gmail.com

1. Introduction

Quantum chemistry calculations have become indispensable tools for understanding molecular properties, reaction mechanisms, and material behavior at the atomic scale. Density Functional Theory (DFT) has emerged as one of the most widely used quantum mechanical methods due to its balance between computational efficiency and chemical accuracy. However, as research problems become increasingly complex and systems grow larger, the computational demands of traditional DFT implementations often become prohibitive. Modern scientific computing has witnessed advances in optimization techniques, parallel processing, and algorithmic improvements that directly benefit quantum chemistry calculations. The Python ecosystem, in particular, has seen growth in high-performance scientific libraries that enable researchers to leverage these advances more effectively.

This article presents a performance analysis comparing traditional quantum chemistry frameworks with modern optimized implementations. The development of environmentally-friendly materials for electrochemical applications is crucial for sustainable energy storage and electrolysis technologies. This study focuses on chlorinated analogues of phosphoric acid compounds ($\mathrm{CCl_3CCl_2PO_3H}$) as alternatives to fluorinated materials for use in electrochemical systems, aiming to reduce environmental impact while maintaining proton conductivity properties for electrolysis processes. Each test system contains 38 atoms with varying water molecule distributions to simulate realistic chemical environments. Our analysis demonstrates substantial computational speedups obtained through modern code optimization strategies while maintaining chemical accuracy.

2. Methods

2.1. Target System

Understanding proton exchange membrane structures represents a particularly challenging application for quantum chemistry calculations. The structural analysis involves complex multi-component systems including polymer chains, ionic species, and interfacial interactions. Traditional DFT implementations require substantial computational resources to adequately characterize these structures.

For example, studying environmentally-friendly phosphoric acid structures based on chlorinated analogues of $CF_3CF_2PO_3H$ [1], specifically $CCl_3CCl_2PO_3H$ compounds with water molecules, requires calculations on systems containing 38 atoms total with multiple conformational states. The fluorine-to-chlorine substitution aims to develop non-fluorinated materials for environmental sustainability. To investigate proton conductivity properties, several H_2O molecules are included in the computational model to simulate realistic proton-conducting environments. Using conventional frameworks like SIESTA [2], geometry optimization requires substantial computational time on modern computational clusters, making comprehensive structural studies computationally expensive.

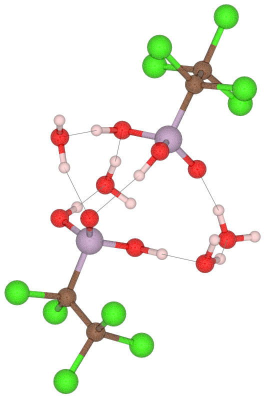

Figure 1. *Chemical structure of the* $CCl_3CCl_2PO_3H$ *compound with* H_2O *molecules studied in this work, showing the chlorinated analogue design for environmental sustainability and proton conduction pathway investigation.*

2.2. Computational Frameworks

Traditional quantum chemistry frameworks have provided the foundation for decades of molecular modeling research. Packages such as SIESTA have been widely adopted due to their robust implementations and extensive feature sets. However, traditional implementations like SIESTA often present computational bottlenecks when applied to complex systems requiring extensive parameter space exploration. SIESTA remains valuable for specific applications requiring its specialized basis sets, pseudopotentials, or particular exchange-correlation functionals, but may be less optimal for high-throughput studies.

Modern Python-based quantum chemistry environments offer computational advantages through optimized algorithms, efficient memory management, and leveraging of high-performance numerical libraries. Frameworks such as PySCF [3] and its GPU-accelerated extension GPU4PySCF [4] have emerged as powerful alternatives that can dramatically reduce computation times while maintaining chemical accuracy. PySCF achieves performance improvements through efficient integral computation algorithms and optimized parallel processing using OpenMP and MPI implementations, while GPU4PySCF further accelerates calculations by leveraging GPU parallelization for computationally intensive operations like two-electron integrals and matrix operations.

3. Experimental Conditions

Our benchmarking study compares computational performance across SIESTA, PySCF, and GPU4PySCF implementations using identical test systems. The benchmark suite consists of 85 individual input files, each being a 38-atom system containing a chlorinated phosphoric acid structure ($CCl_3CCl_2PO_3H_2$) with water molecules. These configurations feature varying water molecule distributions around the phosphoric acid structure to represent different water-membrane configurations, providing realistic system sizes for environmentally-friendly materials design with statistical reliability.

All calculations were performed using identical computational setups to ensure fair comparison:

- **System**: 38-atom systems containing 2 $CCl_3CCl_2PO_3H_2$ structures with 4 H_2O molecules
- **Exchange-correlation functional**: RPBE
- **Basis set**: DZP (Double-Zeta Polarized)
- **Convergence criteria**: SCF tolerance of 10^{-6} for all calculations
- **Computational environment**:
  - SIESTA (CPU): 16 parallel jobs (2 cores per job)
  - PySCF (CPU): 16 parallel jobs (2 cores per job)
  - GPU4PySCF (GPU): 16 parallel jobs
- **Hardware**:
  - CPU: Intel Xeon processors, 32 vCPUs, 256 GiB memory (AWS r7i.8xlarge, $2.117/hr)
  - GPU: 8x NVIDIA A100 80GB SXM4, 240 vCPUs, 1800 GiB RAM (Lambda, $14.32/hr)
- **Performance metrics**: Execution times and costs calculated based on cumulative processing time for all 85 configurations divided by parallelization factor (16) to reflect actual wall-clock time and resource usage

4. Results and Discussion

4.1. Performance Comparison

Figure 2 shows the performance differences between the three implementations. SIESTA represents the traditional approach with longer execution times, while PySCF provides acceleration through modern optimization techniques. GPU4PySCF shows the best performance by leveraging GPU acceleration, showing improvement in execution time while maintaining lower computational costs. These performance differences enable researchers to conduct more extensive conformational sampling and statistical analysis within available computational resources.

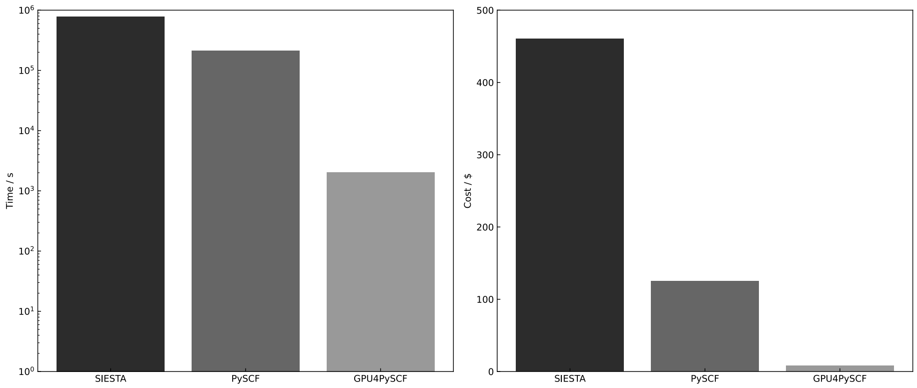

Figure 2. *Computational performance comparison between SIESTA, PySCF, and GPU4PySCF for 38-atom systems containing phosphoric acid structures, showing both execution time (left) and computational cost (right).*

4.2. Impact on Research Methodology

The performance improvements demonstrated in this study have broader implications beyond phosphoric acid structure analysis. The 3.7× speedup observed with PySCF and 390× speedup with GPU4PySCF can be expected to apply to similar molecule systems with comparable atom counts and electronic structure characteristics.

These computational advantages enable:

- **Expanded parameter space exploration**: Systematic investigation of molecular modifications, conformational changes, and environmental conditions across diverse chemical systems through sampling of structural configurations
- **Enhanced statistical reliability**: Larger sample sizes enabling more robust conclusions and confident predictions in materials design studies through increased sampling of molecular configurations and property distributions
- **Integration with modern workflows**: Seamless compatibility with machine learning approaches and automated high-throughput computational screening pipelines, facilitating development of predictive models for materials discovery

The methodology established here provides a foundation for accelerating quantum chemistry calculations across various molecular systems, particularly those involving compounds with heteroatoms (atoms other than carbon and hydrogen) and solvation effects (interactions between solutes and surrounding solvent molecules).

5. Conclusions

This study demonstrates computational advantages of modern PySCF implementations over traditional SIESTA for DFT calculations.

Key findings include:

1. **Performance improvement**: PySCF shows 3.7× speedup and GPU4PySCF demonstrates 390× speedup compared to SIESTA while maintaining chemical accuracy
2. **Cost efficiency**: Computational costs reduced from $460.50 (SIESTA) to $125.35 (PySCF) and $7.99 (GPU4PySCF), representing 73% and 98% cost reductions respectively

3. **Broader applicability**: The methodology extends to similar molecular systems, enabling systematic parameter exploration and integration with modern computational workflows

These results provide a foundation for accelerating quantum chemistry calculations across various molecular systems. The demonstrated efficiency gains illustrate the potential for making quantum chemistry calculations more accessible for complex materials science applications.

6. Code Availability

The computational scripts, input files, and data processing code used in this study are available at:

https://github.com/schwalbe10/quantum-chemistry-acceleration

The repository includes:

- SIESTA, PySCF and GPU4PySCF calculation scripts
- 85 input configurations for phosphoric acid structures
- Data processing and visualization code
- Detailed instructions for reproducing the benchmark results

References

[1] T. Ogawa, H. Ohashi, G. M. Anilkumar, T. Tamaki, and T. Yamaguchi, "Suitable Acid Groups and Density in Electrolytes to Facilitate Proton Conduction," *Physical Chemistry Chemical Physics*, vol. 23, no. 41, pp. 23778–23786, 2021, doi: 10.1039/D1CP00718A.

[2] J. M. Soler *et al.*, "The SIESTA Method for Ab Initio Order-N Materials Simulation," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, pp. 2745–2779, 2002, doi: 10.1088/0953-8984/14/11/302.

[3] Q. Sun *et al.*, "PySCF: The Pythonbased Simulations of Chemistry Framework," *WIREs Computational Molecular Science*, vol. 8, no. 1, p. e1340, 2018, doi: 10.1002/wcms.1340.

[4] X. Wu *et al.*, "Enhancing GPU-Acceleration in the Python-Based Simulations of Chemistry Framework," *WIREs Computational Molecular Science*, 2024, doi: 10.1002/wcms.70008.