Pushing the Accuracy Limit of Foundation Neural Network Models with Quantum Monte Carlo Forces and Path Integrals

Anouar Benali
Qubit Pharmaceuticals Inc, Boston, MA 02116, USA

Thomas Plé
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 75005 Paris, France

Olivier Adjoua
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 75005 Paris, France

Valay Agarawal
Department of Chemistry, University of Chicago, Chicago, IL 60637, USA

Thomas Applencourt
Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439, USA

Marharyta Blazhynska
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 75005 Paris, France

Raymond Clay III
HEDP Theory Department, Sandia National Laboratories, Albuquerque, NM 87185, USA

Kevin Gasperich
Qubit Pharmaceuticals Inc, Boston, MA 02116, USA

Khalid Hossain
Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439, USA

Jeongnim Kim
Intel Corporation, Hillsboro, OR 987124, USA

Christopher Knight
Computational Science Division, Argonne National Laboratory, Argonne, IL 60439, USA

Jaron T. Krogel
Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Yvon Maday
Sorbonne Université, Université Paris Cité, CNRS, INRIA, LJLL UMR 7598, Paris 75005, France

Maxime Maria
Université de Limoges, XLIM UMR CNRS, 87060 Limoges, France

Mathieu Montes
Conservatoire National des Arts et Métiers, GBCM, 75003 Paris, France

Ye Luo
Computational Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

Evgeny Posenitskiy
Qubit Pharmaceuticals, 75014 Paris, France

Corentin Villot
Qubit Pharmaceuticals, 75014 Paris, France

Venkat Vishwanath
Argonne Leadeship Computing Facility, Argonne National Laboratory, Argonne, IL 60439, USA

Louis Lagardère
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS and Qubit Pharmaceuticals, Paris, France

Jean-Philip Piquemal
Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS and Qubit Pharmaceuticals, Paris, France
Email: jean-philip.piquemal@sorbonne-universite.fr

***Abstract*—We propose an end-to-end integrated strategy for the production of highly accurate quantum chemistry (QC) synthetic datasets aimed at deriving atomistic Foundation Machine Learning (ML) Models. We first present a GPU-accelerated QC database generation Exascale protocol able to produce the required energies and forces. A "Jacob's Ladder" approach leverages computationally-optimized layers of massively parallel high performance software with increasing accuracy to compute: i) Density Functional Theory (DFT); ii) Quantum Monte Carlo (QMC); iii) Selected Configuration Interaction**

(sCI), within large volumes and optimized time-to-solution performances. Handling this ambitious computational pipeline would be impossible without exascale computing resources, particularly for the notoriously difficult and computationally intensive calculation of QMC forces and for the combination of multi-determinant QMC energies and forces using selected CI wavefunctions methodologies. To our knowledge, this is the first time that such quantities are computed at such scale. We combine these data with the FeNNix-Bio-1 foundation ML model to bridge the gap between highly accurate QC calculations and condensed-phase Molecular Dynamics (MD). We demonstrate stable multi-ns simulations using the resulting beyond DFT accuracy fully reactive model coupled to full path integrals adaptive sampling quantum dynamics. A complete 1 million-atom plant virus solvated structure, including its full genetic material, is simulated using Ring-Polymer MD quantum dynamics along as its response to acidification under physiological NaCl concentrations. These new capabilities open the door to the possibility to monitor bond breaking/creation and proton transfers chemical interactions taking place in biosystems allowing us to reach a deeper understanding of their complex internal machinery.

1. Justification for Prize

- Record Quantum Monte Carlo and Selected Configuration Interaction (energies and forces) computations in terms of number of calculations: 20,251 QMC (energies+forces), 20,338 QMC (energies-only), 2000 s-CI (energies+forces); and achieved accuracy. Double precision performance of 124/62/52 PFLOPS on 2048 nodes of Aurora (12,288 GPUs), representing approximately 60%/30% and 25% of theoretical peak performance for energy calculations, force calculations, and multideterminant force calculations, respectively.
- For a 32-million-particule (1-million atom X 32 beads) Satellite Tobacco Virus system (STMV) Ring Polymer quantum molecular dynamics simulation with a foundation neural network model (using project quantum chemistry data) while using multi-replica via an unsupervised adaptive sampling strategy, time-to-solution of 0.44ns/day (128 nodes, 1024 GPUs, mixte precision) on Jean Zay. Establish new state-of-the-art for foundation model quantum dynamics simulation of biological system. Simulations captured the STMV response to acidification under physiological NaCl concentrations for the first time.

2. Performance Attributes

Performance attribute	Our submission
Category of achievement	Time-to-solution, scalability
Type of method used (1)	Quantum Chemistry
Type of method used (2)	Neural Network potential Molecular Dynamics
Results reported on basis of	Whole application including I/O
Precision reported	Double precision, mixed precision
System scale	Measured on full systems / extrapolated times
Measurements	Timers, FLOP count

3. Overview of the problem

Simulating the structural dynamics of biological systems at full quantum accuracy is a daunting task requiring precise modeling of the Born-Oppenheimer electronic potential energy surface while including nuclear quantum effects in the dynamics to accurately reproduce condensed phase properties. Up to now, achieving such high-fidelity simulations has been prevented by the computational cost of first-principles quantum approaches and remains a Graal for the molecular simulation community. In this context, neural networks potentials have started to gain speed and precision offering a clear alternative to ab initio molecular dynamics (AIMD) and recently even to force fields simulations thanks to high performance computing. Thus, the field evolves very fast, and larger foundation models dedicated to atomistic Molecular Simulation have been recently introduced. These single general-purpose ML potentials have various capabilities and can be applied to various types of applications, efficiently replacing quantum mechanical approaches. However, in practice, such ML approaches cannot perform computations at an accuracy better than the models' reference data. Therefore, foundation models rely heavily on the quality and the quantity of their reference synthetic data. The choice of the quantum chemistry approaches and their capability to approximate the Schrodinger equation therefore have a considerable impact on the performance of the model. Additionally, on the molecular dynamics side, since foundation models can be seen as Born-Oppenheimer potential surfaces generators, it is critical to go beyond classical dynamics. Indeed, it is paramount to be able to include the effects of nuclei to perform quantum dynamics to access real condensed-phase properties. Overall, the rise of exascale computing offers the possibility of designing a new, multi-level, pipeline for the generation of accurate quantum chemistry data that would strongly impact future fondation machine learning model developments.

4. Current State of the art

4.0.1. Quantum Chemistry Datasets. In that context, the present consensus for the computation of reference quantum chemistry datasets is to use Density Functional Theory (DFT) potential energy surfaces which provide a reasonable accuracy at an affordable cost. However DFT's precision strongly depends on the choice of the functional, which is by essence non-exact and therefore not fully transferable across the chemical space, especially if one wants to simultaneously model structures, weak interactions and reactivity of complex biological systems. Therefore, there

is a need for more exact ab initio methods. for several decades CCSD(T) [72], [80] – often termed the "gold standard" – has been widely employed. Nonetheless, CCSD(T) exhibits steep $O(N^7)$ scaling, and its local approximations (e.g., DLPNO-CCSD(T)) rely on domain truncations that are typically benchmarked against known data but can fail if the chosen domains omit critical electron correlation [80]. As a result, such local-correlation variants, while computationally attractive for large systems, can yield unpredictable errors outside the familiar chemical space. Similarly, Density Matrix Renormalization Group (DMRG) offers high accuracy for strongly correlated systems but remains challenging for broad chemical data generation due to the need for carefully chosen active spaces. [106]
By contrast, Quantum Monte Carlo (QMC) methods [73], and particularly Diffusion Monte Carlo (DMC) [27], [74], have proven capable of near-exact ground-state energies across wide chemical spaces [10], [11], [23], [28], [47], [48], [61], [83], [84], [85]. Large-scale QMC studies focused on benchmark energies [12], [44], but until recently lacked an efficient pathway to obtain forces, rendering them less suitable for force field generation. Computing forces in Diffusion Monte Carlo (DMC) remains inherently challenging due to the fixed-node approximation and the intrinsic stochastic nature of the algorithm. Fortunately, over the past several years, a variety of technical [26], mathematical [96], and performance improvements [25] to the original Zero-Variance Zero-Bias force estimator framework of Assaraf and Caffarel [4] have made QMC forces usable and trustworthy for large scale applications.

Generating multideterminant expansions to reduce nodal-surface errors has proven challenging. Methods such as CASSCF have been recognized for decades as effective but require extensive prior knowledge of the target system's chemical properties, which limits their suitability for broad machine-learning applications. More recently, Caffarel and co-workers [32] revived the concept of selected Configuration Interaction (sCI) [38], automating the construction of multideterminant expansions in a largely black-box manner and enabling various sCI flavors [37], [93]. As a multideterminant trial wavefunction, sCI systematically improves nodal accuracy. Coupled with an optimized Jastrow factor, it resolves both static and dynamic correlations by addressing strong-correlation effects through the sCI expansion while capturing dynamical correlations in the Jastrow [12], [29], [51], [52], [53], [54], [63], [65], [75], [76]. In turn, QMC forces can validate or directly compute sCI forces, establishing a synergistic approach with high reliability [87].

Nevertheless, the computational cost of these methods remains substantial, restricting routine usage for many years. Over the past two decades, the community has extensively developed QMCPACK [43], [45], [58], [59] to exploit leadership-class supercomputers—from IBM Blue Gene architectures to exascale systems such as Aurora and Frontier at Argonne and Oak Ridge National Laboratories. Early efforts prioritized a GPU-based energy estimator, particularly via efficient B-spline orbital evaluations on massively parallel platforms, targeting material sciences application. Building on this foundation, the present work introduces new optimizations and data layouts that optimize the LCAO branch of QMCPACK for GPU acceleration for molecular applications.
In parallel, a modern sCI program has been developed with GPU-friendly data structures and offloaded core operations in preparation for large-scale parallelism. Taken together, these advances enable QMC–sCI workflows on exascale platforms, providing the high-fidelity datasets needed for force-field generation and facilitating the study of previously intractable systems in chemical and materials research.

4.1. Foundation Machine Learning Models for atomistic molecular dynamics Simulations

The application of Neural Networks to molecular dynamics is recent and started with the availability of suitable neural architecture [6], [9]. Since then, numerous neural network potentials (NNPs) have been proposed [8], [9], [14], [20], [21], [30], [31], [34], [35], [46], [55], [62], [66], [71], [77], [78], [81], [88], [94], [100], [101], [104], [105] while their high performance implementations reached a mature stage [39], [41], [86] and as specialized libraries are now available to NNPs developers [22], [67], [79], [102]. Very recently, in link with the breakthroughs of the AlphaFold [42] and RoseTTAFold [5] protein structure foundation models, a major shift in the developments started to reshape molecular simulation with introduction of the MACE-MP-0 foundation model. [7]. As explained by its authors, MACE is a single general-purpose ML potential, designed for material science atomistic simulations. Its capabilities range from single molecule to full materials simulations, and enables modeling of chemical reactivity. Beyond solids, its extended transferability allowed it to be tested even on small proteins where stable molecular dynamics could be obtained. However, since MACE and its associated neural architecture are primarily designed for materials, its applicability to condensed phase simulations of large complex biological molecules remains limited both in term of accuracy and computational performances. For these reasons, some of us recently introduced the FeNNix-Bio1 Foundation Model [68]. Note that its associated paper can be seen as the machine learning counterpart of the present Exascale computing report as both projects were launched together and as the FeNNix-Bio1 development used preliminary data extracted from the present quantum chemistry pipeline. Thus, it is able to perform atomistic condensed phase molecular dynamics simulations of biological systems while allowing for reactivity modeling. FeNNix-Bio1 is fast and accurate, being able to address simultaneously the issues of the computational requirements generated by the biological timescales and the difficult transferability of charged-species, that are numerous in Chemistry and Biology. Finally, FeNNix-Bio1 is also able to compute meaningful condensed-phase properties since it explicitly takes into account nuclear quantum effects in its

(quantum) dynamics. [18], [60], [69]. This allows the foundation model to predict accurate condensed phase properties such as free energies, solvent properties, etc. .

5. Innovations

5.1. Dataset Generation: High Performance Quantum Chemistry Pipeline

We first introduce our GPU-accelerated quantum chemistry database generation Exascale protocol which has been optimized to able to produce the required energies and forces. Organized as a "Jacob's Ladder" approach (see Figure 1) with methods of increasing complexity, it leverages 3 different layers of computationally-optimized massively parallel high performance softwares dealing with increasing levels of accuracy. Note that exact, i.e. Full-Configuration interaction (or Full-CI) computations in the complete basis set limit, numerically reaching the quality of the electronic time-independent, non-relativistic, Schrödinger's equation quality computations are mostly out of reach of classical computers due to the existence of a complexity exponential wall (see [95] and references therein). Such a quality may be reached in a near future using quantum computers and fault tolerant Quantum Processing Units (QPU) [103]. For now, and since the goal of our work is to apply the quantum data to chemistry and biology simulations, the multi-level GPU-accelerated computational strategy proposed here aims at reaching the chemical accuracy which is the required level of precision needed to perform realistic chemical, i.e. comparable to experiment, predictions. Chemical accuracy implies errors below 1 kcal/mol (1.59 mH/particle) with respect to the exact Schrödinger total energy. Since forces are the cornerstone of the design of machine learning potentials, each proposed level includes the efficient evaluation of the energy gradients.

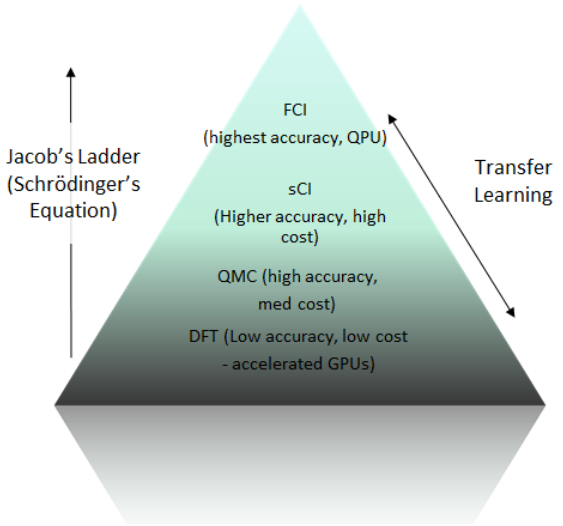

Figure 1. Jacob's Ladder for quantum chemistry methods towards the full Schrödinger's Equation. Acronyms: FCI= Full Configuration Interaction, sCI= Selected Configuration Interaction; QMC= Quantum Monte-Carlo; DFT=Density Functional Theory; QPU= Quantum Processor Unit.

5.1.1. The SPICE2(+)-ccECP dataset. In order to test the computational pipeline, we computed a first training set denoted SPICE2(+)-ccECP. It used the molecular configurations of the SPICE2 database [24], [89], which we extended with a carefully selected subset of 100k configurations from the ANI-2X dataset [21]. These latter were selected via a round of active learning using the uncertainty given by a preliminary version of the FeNNix-Bio1 model and recomputed using our selected level of theory.This combined dataset provided comprehensive coverage of the conformational space relevant to biomolecular systems, including diverse functional groups, a wide range of environmental conditions, and various non-equilibrium configurations critical for modeling reactive processes. Diverging from the original SPICE2, we combined in the present work DFT with the correlation-consistent effective core potentials (ccECP) [97] and their companion aug-cc-pVTZ basis sets [97], as implemented in the PySCF package [91]. The choice of the present effective core potentials conserves accuracy while increasing the computational efficiency compared to the explicit inclusion of core electrons. Moreover, this enables us to extend the SPICE2 dataset to post Hartree-Fock methods such as QMC where the use of valence-only calculations is crucial. Overall, the dataset encompasses 2,108,628 conformations that were all treated (energies and forces) at the DFT level. Among these latter, 20,251 conformations were computed using QMC (energies and forces), 20,338 QMC (energies-only) and 2000 (energies and forces) with multideterminant QMC with selected CI wavefunctions.

5.1.2. Density Functional Theory.

- Summary of Contributions

For the first layer of our quantum chemistry database generation protocol, we performed Density Functional Theory

(DFT) calculations using the PySCF code [50], [90], [99], specifically leveraging the GPU-accelerated implementation. We strategically allocated computational resources based on system size. For molecules up to 175 electrons (approximately 1.3 million configurations), calculations were executed on the Leonardo supercomputer at Cineca (Italy), larger molecular systems with up to 275 electrons (approximately 0.8 million configurations) were computed on the Polaris supercomputer at Argonne National Laboratory. Both supercomputers feature NVIDIA A100 GPUs-40Gb. The largest molecules with more than 275 electrons (65,521 conformations) were computed on Aurora Supercomputer which features Intel Xe GPUs. A significant challenge for large-scale DFT calculations is the computation and storage of electron repulsion integrals (ERIs), which cannot be generated on-the-fly on GPUs for larger molecules due to memory limitations of the A100-40Gb cards. The conventional CPU implementation would require over 6 hours per molecule on AMD Epyc 7303 processors and up to 300GB of temporary storage per calculation, making the complete dataset generation computationally infeasible. To overcome these limitations, we employed our gpu4mrh implementation, which optimizes memory usage and enables efficient offloading of integral calculations to GPUs even for large molecular systems.

We employed the ωB97M-D3BJ functional [33], [56], [57], which provides an excellent balance between accuracy and computational efficiency for organic and biological systems. This range-separated hybrid meta-GGA functional incorporates D3BJ dispersion corrections to accurately capture non-covalent interactions critical for biological systems. All calculations utilized correlation-consistent effective core potentials (ccECP) [3], [13] to reduce computational cost while maintaining high accuracy for valence electrons.

- Algorithmic Innovations

To optimize computational resources while ensuring sufficient accuracy, we employed an adaptive quadrature grid strategy. For energy calculations, we utilized a grid level of 3, while for gradient calculations (forces), we increased the grid density to level 4 to ensure accurate force evaluation. This strategy significantly reduced computational cost and memory usage while maintaining chemical accuracy in both energies and forces.

The DFT calculations were distributed across the Nvidia A100 nodes systems (Polaris and Leonardo) with one molecule assigned per GPU, using 4 GPUs per node. This distribution strategy optimized the trade-off between parallel efficiency and memory requirements. Specifically, the DFT calculations achieved approximately 16 TFLOPS per node (40% of theoretical peak) for smaller molecules (0–100 electrons), 20 TFLOPS per node (50% of theoretical peak) for intermediate-sized molecules (100–200 electrons), and 28 TFLOPS per node (70% of theoretical peak) for larger molecules (200–250+ electrons). Each DFT simulation was distributed across 128 nodes, with 4 Nvidia A100 GPUs per node (512 GPUs per simulation in total). This large-scale deployment was essential, as it was the only feasible way to generate the extensive dataset of 2.1 million molecular configurations within the allocated time frame. Specifically, the simulations achieved peak performances of approximately:

- 1.98 PFLOPS) for smaller molecules (0–100 electrons),
- 2.48 PFLOPS for intermediate-sized molecules (100–200 electrons),
- 3.48 PFLOPS for larger molecules (200–250+ electrons).

These performances represent approximately 40–70% of the peak theoretical capability of the utilized hardware nodes.
On Aurora, we used PySCF coupled with the gpu4mrh code that leverages the Intel Xe architecture through efficient data layout patterns and kernel optimizations to maximize use of matrix-multiply operations central to quantum chemistry calculations. The gpu4mrh code, initially developed to accelerate multi-configurational electronic structure methods, uses a lightweight Python/C++ abstraction layer to seamlessly offload fundamental operations from PySCF (namely, formation of the Coulomb and exchange matrices) to GPUs via CUDA, HIP, or SYCL and respective vendor-optimized math libraries (e.g. MKL on Aurora). With an efficient hashing strategy to minimize data-transfer GPU-host data transfers, the majority of the GPU runtime is spent computing single and batched matrix-matrix products concurrently across multiple GPUs in a compute node. The gpu4mrh code was used as a drop-in replacement enabling SCF energy calculations to leverage the full 768 GB of HBM from the 6 GPUs on each Aurora node. The cost of this implementation is dominated by several dense gemm operations, so it is able to make use of vendor-optimized BLAS routines and achieves high GPU utilization for the systems considered here. In total, we performed calculations on 2,1 million molecular configurations, requiring approximately 20k nodes-hours on Leonardo, 35k node-hours on Polaris and 25k nodes hours on Aurora. The resulting database of energies and forces serves as the foundation layer for our multi-level quantum chemistry approach.
In addition to providing the first level of our Jacob's Ladder approach, these DFT calculations served as essential inputs for our higher-level quantum chemistry methods. The DFT wavefunctions were used as trial wavefunctions for the single-determinant Quantum Monte Carlo calculations. Additionally, the one- and two-electron integrals generated during the DFT calculations were utilized as inputs for the selected Configuration Interaction (sCI) calculations, ensuring consistency across all three levels of theory.

5.2. Quantum Monte Carlo

QMC techniques are stochastic methods that solve the many-body Schrödinger equation with high accuracy using a limited but controlled number of approximations. By explicitly including many-body electronic interactions, these methods achieve chemical accuracy while maintaining favorable scaling compared to traditional high-level quantum

chemistry approaches. In this work, we leverage the extensive refactoring of QMCPACK that has been specifically optimized for exascale computing platforms over the past two decades. Building on previous optimizations that targeted material science applications, we extended and enhanced the Linear Combination of Atomic Orbitals (LCAO) branch of QMCPACK with new optimizations and data layouts specifically designed for GPU acceleration in molecular applications. These enhancements enable unprecedented scaling on Aurora's architecture for complex molecular systems. Among the various QMC approaches, we employed Diffusion Monte Carlo (DMC) due to its ability to recover a significant portion of the correlation energy while maintaining favorable scaling with system size. DMC solves the Schrödinger equation in imaginary time $\tau = it$, ensuring that any initial state $|\psi\rangle$ not orthogonal to the ground state $|\phi_0\rangle$ will converge to the ground state in the long-time limit:

$$\lim_{\tau\to\infty} \Psi(\mathbf{R},\tau) = c_0 e^{-\epsilon_0\tau}\phi_0(\mathbf{R}) \quad (1)$$

To address the fermion sign problem, we employed the fixed-node (FN) approximation [2], which constrains the nodal surface of the many-body wavefunction to match that of a trial wavefunction. This introduces the only systematic error in DMC when the reference wavefunction is not exact. To minimize this error, we used a trial wavefunction of the Slater-Jastrow form:

$$\Psi_T(\vec{R}) = \exp\left[\sum_i J_i(\vec{R})\right] \sum_k^M C_k D_k^{\uparrow}(\phi) D_k^{\downarrow}(\phi) \quad (2)$$

where $D_k^{\uparrow}(\phi)$ and $D_k^{\downarrow}(\phi)$ are Slater determinants for up and down spin electrons expressed in terms of single-particle orbitals $\phi_i = \sum_l^{N_b} C_l^i \Phi_l$ The Jastrow factor $\exp[\sum_i J_i(\vec{R})]$ explicitly accounts for electron-electron correlations and significantly improves the trial wavefunction quality while maintaining the original nodal structure.

For our QMC calculations, we used the QMCPACK code [43], [45], [58], [59], which has been specifically optimized for exascale computing platforms. The single-particle orbitals in our trial wavefunctions were generated using the ωB97M-D3BJ functional as implemented in the PySCF package. This range-separated hybrid meta-GGA functional with D3BJ dispersion corrections was chosen for its excellent balance between accuracy and computational efficiency for organic and biological systems. All calculations employed correlation-consistent effective core potentials (ccECP) to reduce computational cost while maintaining high accuracy for valence electrons.

For each molecular system, we optimized up to 40 variational parameters, including one-body, two-body, and three-body Jastrow factors, using a variant of Umrigar's linear method. This optimization was performed using variational Monte Carlo (VMC) before the subsequent DMC calculations. The Jastrow parameters were optimized independently for each molecule to ensure optimal trial wavefunction quality. All DMC calculations used a time step of 0.01 a.u., which we verified through time-step extrapolation studies to be within error bars of the zero time-step limit. To ensure statistical independence and proper convergence, we used 4096 walkers per calculation, distributed across the exascale computing resources.

The stochastic nature of QMC offers exceptional parallel efficiency on modern supercomputers, with computational tasks being evaluated independently across thousands of processing units. This "embarrassingly parallel" characteristic positions QMC as an ideal candidate for exascale applications, allowing us to simulate systems containing thousands of electrons while maintaining quantum mechanical accuracy.

5.2.1. QMC Energy and Wavefunction Evaluation .

- Summary of Contributions

We developed a highly optimized implementation of quantum Monte Carlo calculations in QMCPACK that leverages the massive parallelism of exascale hardware. Building on previous optimizations for material science applications, we specifically enhanced the Linear Combination of Atomic Orbitals (LCAO) branch to accelerate molecular applications on GPU architectures. Our implementation enables unprecedented calculation throughput, processing over 40,000 molecular configurations with QMC (20,251 with forces) on the Aurora supercomputer, achieving 124 PFLOPS sustained performance on 2048 nodes, and while using over 4500 walkers per rank (see Fig3).

- Algorithmic Innovations

The core innovation in our QMC implementation is a multi-level batching approach that maximizes GPU utilization and minimizes kernel launch overhead:
Walker-Level Batching Multiple Monte Carlo walkers are evaluated simultaneously, transforming independent stochastic samples into batched operations that better utilize GPU parallelism.
Multi-Center Batching Rather than evaluating atomic orbitals one center at a time, we group centers by species and process them in batches. This approach:

- Flattens the combined dimension of {centers, electrons} into a single array dimension
- Groups centers by species to minimize kernel launches Maintains consistent indexing schemes for reading and writing data
- Accumulates all partial orbital contributions in a minimal number of passes.

Unified Data Layout: We implemented consistent flattening schemes that eliminate index mismatches between writing and reading orbital data, preventing out-of-bounds memory accesses and simplifying debugging.
Optimized Kernel Structure: After computing radial (Rnl) and angular (Ylm) expansions in separate kernels, we combine partial results in a single kernel that maximizes

GPU occupancy.

These implementations reduced kernel launch overhead by an order of magnitude and significantly improved memory throughput, particularly important for the LCAO approach used in molecular systems where the number of electrons is small. The architecture-independent optimizations for GPUs achieved 60% of theoretical peak performance for typical molecular configurations in our dataset. Additional batching strategies will allow us to reach

5.2.2. Single Determinant QMC Forces .

- Summary of Contributions

For the Quantum Monte Carlo (QMC) layer of our quantum chemistry database generation protocol, we build upon the existing force evaluation formalism in QMCPACK, developed according to Filippi et al.'s [25] approach. Rather than reimplementing the mathematical framework, we focused on optimizing the computational efficiency for exascale architectures by integrating the force evaluation with QMCPACK's batched structure. This optimization enabled us to compute accurate interatomic forces for over 20,000 molecular configurations, achieving 62 PFLOPS sustained performance on Aurora with scaling efficiency of 30% up to 2048 nodes.

By optimizing the force evaluation code to utilize the same batching structures we developed for energy calculations, we achieved an acceleration of 12× compared to the original implementation while maintaining the same statistical accuracy. This performance gain removes a critical bottleneck in using QMC for large-scale force calculations, enabling the largest database of QMC-Forcesfor machine learned potentials. Additional components of the force evaluation remain on the CPU and lead to significant data transfers hindering performance. Future developments will focus particularly on porting them to GPU.

- Algorithmic Innovations

Our primary innovation was adapting the existing force evaluation implementation to utilize the batched structure we developed for wavefunction evaluation:
Efficient Matrix Formalism: We implemented the complete mathematical framework from Filippi et al., which expresses derivatives and one-body operators through a unified approach. The logarithmic derivative of a Slater determinant with matrix $A_{ij} = \phi_j(r_i)$, the logarithmic derivative with respect to any parameter λ is:

$$\frac{d \ln D}{d\lambda} = \text{tr}(A^{-1}B) \quad \text{where} \quad B = \frac{dA}{d\lambda} \tag{3}$$

This formulation allows us to compute forces as:

$$\frac{\partial E_L}{\partial R_a} = \text{tr}\left(A^{-1} \frac{\partial B}{\partial R_a} - X \frac{\partial A}{\partial R_a} \right) \tag{4}$$

where $X = A^{-1}BA^{-1}$. The derivatives of the matrices A and B with respect to nuclear coordinates are fully implemented, including the complex expressions for the non-local pseudopotential contribution. This implementation verified that the computational cost scales as $O(N^3 * 3 * N_{atom})$ with minimal overhead compared to energy-only calculations.

Integration with Multi-Walker Batching: We restructured the force evaluation code to process multiple walkers simultaneously. Rather than independently computing forces for each walker, we leveraged the batched walker approach to transform the calculations into batched matrix operations. This significantly improved GPU utilization and reduced kernel launch overhead.
Extending Multi-Center Batching to Forces: We extended our multi-center batching approach from energy calculations to force evaluations. This required careful adaptation of the force calculation kernels to process multiple atomic centers in batches, maintaining the same flattened indexing scheme used in the wavefunction evaluation. Each atomic derivative contribution is now computed in batches grouped by atomic species, dramatically reducing the number of kernel launches.
Unified Memory Management: We implemented an optimized memory management scheme that reuses the data structures from the energy evaluation for force calculations. This approach minimizes memory transfers between CPU and GPU and ensures consistent data access patterns across different parts of the calculation.
Optimized Space-Warp Implementation: We adapted the space-warp coordinate transformation to our batched structure. The transformation, which reduces statistical variance in force estimators, was implemented with vectorized operations that process multiple walkers and centers simultaneously, improving computational efficiency.

By integrating force calculations with the batched structure of QMCPACK, we eliminated redundant computations and maximized hardware utilization, enabling force calculations at scales previously impossible with QMC methods. The performance improvements from our batched implementation make QMC force calculations practical for large-scale applications, opening new possibilities for highly accurate molecular dynamics simulations and foundation model training based on QMC-level accuracy.

5.2.3. Multideterminant QMC Forces.

- Summary of Contributions

An important advancement in our work is the full implementation and integration of multideterminant QMC forces in QMCPACK. While some elements of the mathematical framework existed in previous versions, these implementations were incomplete and non-functional for practical calculations. We significantly extended, completed, and optimized these components, then fully integrated them with our batched computation structure to enable exascale performance. With this implementation, we successfully computed forces for 2,000 molecular configurations using multideterminant QMC with selected Configuration Interaction (sCI) wavefunctions, achieving 52 PFLOPS on Aurora, using 200 molecules per run and 20 nodes per molecule, with

1.4M determinants per molecule. This represents a meaningful methodological advancement, as multideterminant approaches provide a systematically improvable path toward more accurate solutions of the many-body Schrödinger equation by reducing the fixed-node error that limits single-determinant QMC. Prior to our work, such calculations remained largely theoretical or limited to small proof-of-concept systems.

- Algorithmic Innovations

Our implementation required completing and integrating several key components of the multideterminant QMC force framework:

Completion of Mathematical Framework: Building on Filippi et al.'s theoretical approach, we completed the implementation of the excitation-based formalism for evaluating multideterminant forces. For a multideterminant wavefunction we finalized all necessary components to evaluate derivatives of ratios of sums of determinants, debugging and extending the existing code to handle excited determinants properly.

Integration with Batched Architecture:We restructured the multideterminant force evaluation to leverage our batched computation framework. This integration ensures optimal performance on exascale hardware by:

- Processing multiple walkers simultaneously
- Grouping determinants by excitation level for batch processing
- Utilizing the same flattened center indexing scheme used elsewhere in the code

Optimization of Table Method Implementation: We refined and optimized the table method algorithm for evaluating excited determinant quantities within our batched framework. This approach now efficiently handles the complex data structures required for multideterminant calculations. Our implementation of multideterminant force calculation re-uses many of the quantities that are already required for the single-determinant force calculation and adds a cost that scales as $O(N_{det} * 3 * N_{atom})$.

Improved Memory Management: We developed enhanced memory management schemes that:

- Minimize data transfers between CPU and GPU
- Ensure optimal data locality for GPU processing
- Reuse memory allocations across different parts of the calculation

Integration with sCI Workflow: We established a computational pipeline that connects our selected Configuration Interaction (sCI) implementation with the multideterminant QMC force evaluation. This integration enables the automatic generation, optimization, and force evaluation of multideterminant wavefunctions in a cohesive workflow.

This implementation makes multideterminant QMC forces practically usable at scale for the first time. By completing, optimizing, and integrating this capability with our batched computation structure, we have created a pathway for generating higher-accuracy reference data for machine learning potentials, particularly for systems where electronic correlation effects are significant.

5.2.4. Selected Configuration Interaction.

- Summary of Contributions

The third layer of our quantum chemistry database generation protocol employs selected Configuration Interaction (sCI) methods, which provide a systematically improvable path toward the exact solution of the many-body Schrödinger equation. Our sCI implementation is based on the CIPSI (Configuration Interaction using a Perturbative Selection done Iteratively) algorithm, which we implemented in C++ and ported to GPUs to address Exascale computing.

Our sCI code is only several months old and is still undergoing rapid development. There are many more algorithmic improvements underway and planned for the near future, but it is already capable of leveraging GPUs, and for this work it was able to generate 1.4M-determinant wavefunctions for 2000 molecules using approximately 1k node-hours, which represents less than 1% of the total computational cost of the full {DFT→sCI→VMC/DMC} workflow per molecule. So even at a very early stage of development, with many planned improvements that will increase the capability of the code, we are still able to greatly improve our nodal surface quality (by adding multideterminant expansions) without adding any significant cost in performing the sCI calculations.

- Algorithmic Innovations

The core of our implementation follows an iterative procedure in which we begin from an eigenvector of the Hamiltonian in a given set of determinants, and we repeat the following steps until some convergence criterion is met:

1) Generate all single and double excitations from the current space
2) Compute second-order perturbative energy contributions
3) Select determinants with the largest contributions
4) Diagonalize the Hamiltonian in the new variational space

The C++ implementation has been specifically optimized for exascale computing with a focus on efficient memory management and parallel execution. Key optimizations include:

- Bit-encoded determinant representation using native fixed-width integer types for compact storage and fast comparison operations
- Task-based parallelism for generation and evaluation of candidate determinants
- GPU-accelerated Hamiltonian diagonalization (Davidson-Liu)

For force evaluation with sCI wavefunctions, we leverage our QMC implementation rather than developing a separate analytical gradient procedure. This approach offers several advantages: it avoids the complication of computing

analytical derivatives of the sCI wavefunctions (which would require more input data than the standard FCIDUMP interface provides), it naturally incorporates the Jastrow factor which significantly improves the wavefunction quality, and it provides a unified computational framework for energy and force evaluation. The multideterminant wavefunctions generated by sCI are directly imported into QMCPACK, where they serve as a starting point for further optimization before force evaluation.

This integration of sCI and QMC enables an efficient workflow where high-quality nodal surfaces are generated through the sCI procedure, then refined and used for force evaluation in QMC, circumventing many of the challenges associated with analytical gradients for large multideterminant expansions while maintaining high accuracy.

5.3. Performance Analysis and Benchmarks

To demonstrate the computational efficiency and accuracy of our new Quantum Monte Carlo (QMC) implementation, we performed extensive performance analyses on several representative molecular systems. We primarily focused on water clusters (monomer, dimer, and trimer) due to their biological relevance and the well-known computational challenges they pose for accurate quantum chemical treatment. Additionally, we benchmarked a representative SPICE molecule with heavier atoms containing 16 electrons for force evaluation, as well as a larger, computationally demanding molecule with 176 electrons to examine scalability.

5.3.1. Single-Determinant QMC Performance. Our initial performance analysis revealed a counterintuitive finding: the original GPU implementation of the QMC algorithm significantly underperformed compared to the CPU-only implementation, particularly for molecular systems with electron counts common in biological applications (Figure 2). As shown, the CPU version consistently delivered higher throughput for molecular systems containing 32, 82, and 240 electrons. The GPU implementation only outperformed the CPU in extremely high walker-count scenarios, a regime practically limited by memory saturation and therefore infeasible for routine calculations.

This stark performance gap highlighted the need for fundamental algorithm redesign, specifically addressing GPU kernel overhead, parallel efficiency, and memory optimization as discussed in Sections 5.2.1 and 5.2.2.

Figure 3 illustrates the substantial performance gains achieved with our new algorithm. We observe clear improvements across water clusters with electron counts ranging from 8 to 152 electrons. Quantitative speedups measured against the CPU baseline are summarized in Table 1, highlighting consistent and significant GPU acceleration.

The significant GPU acceleration is due to our multilevel batching approach, optimized data layouts, and careful tuning of the number of walkers per GPU rank, maximizing occupancy and minimizing kernel overhead. Because QMC methods are embarrassingly parallel, increasing the number of walkers per rank enabled higher throughput with fewer

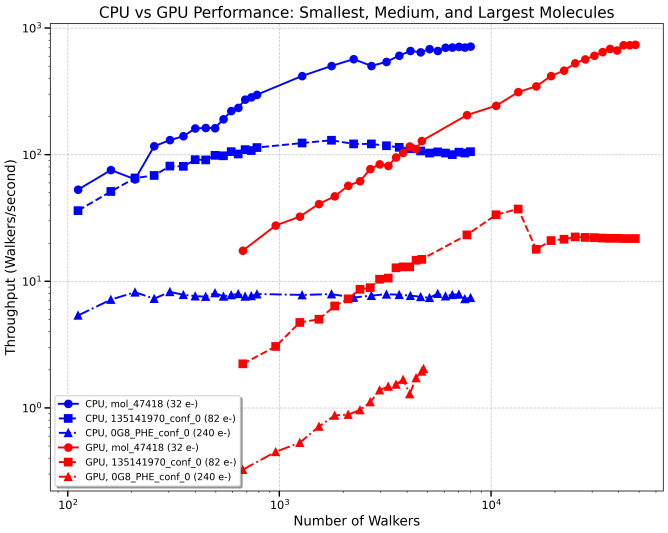

Figure 2. Performance comparison of the initial QMC algorithm on CPU versus GPU for three representative molecular systems with different electron counts (32, 82, and 240 electrons). The CPU implementation consistently outperformed the GPU version, highlighting limitations in data layout, memory overhead, and kernel-launch inefficiencies in the original GPU algorithm.

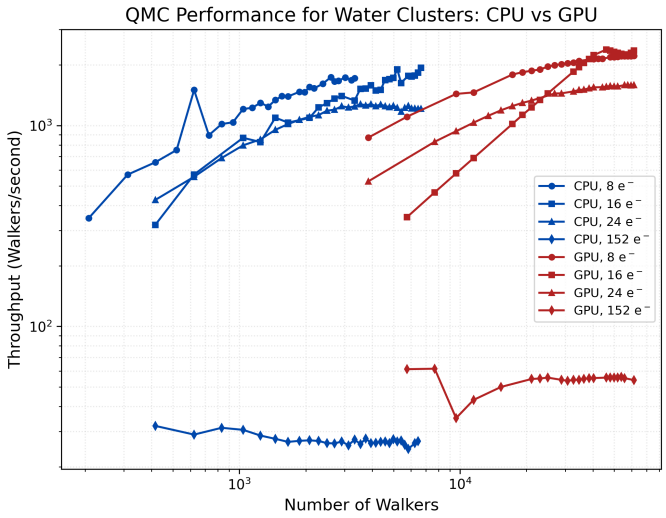

Figure 3. Performance comparison of CPU versus optimized GPU implementations for QMC calculations on water clusters with varying electron counts (8, 16, 24, and 152 electrons). Throughput is shown as walkers per second versus total walker count, demonstrating robust GPU acceleration and scalability.

TABLE 1. MEASURED SPEEDUPS FOR QMC ENERGY EVALUATION ON GPUS COMPARED TO CPUS, UTILIZING THE OPTIMIZED ALGORITHM.

Electron Count	Speedup (GPU vs. CPU)
8	1.29
16	1.22
24	1.31
152	2.01

overall simulation steps, thus effectively utilizing GPU resources.

5.3.2. Resource Allocation Strategy. Through detailed performance analyses, we established an optimal resource allocation strategy that balances computational efficiency with statistical accuracy requirements. As QMC is an embarrassingly parallel method, we calibrated our runs to achieve a fixed statistical error target of approximately 0.0005 mHa (below the chemical accuracy threshold of 1 kcal/mol). Following standard Monte Carlo statistics, the error bars scale as $1/\sqrt{N_{\text{samples}}}$, requiring us to carefully balance the number of walkers per rank with the total number of sampling steps.

To optimize throughput while maintaining accuracy, we determined that the ideal configuration was 4320 walkers per GPU rank. We then scaled the number of nodes based on system size to achieve the desired statistical convergence within reasonable wall-clock times compatible with supercomputer queue policies. This translated to allocating two Aurora nodes per small molecule (up to 16 electrons), six nodes for medium-sized systems like the water trimer (24 electrons), and correspondingly more for larger systems.

For force calculations, which are approximately 100× more computationally intensive than energy-only evaluations due to known bottlenecks in the current implementation (e.g. finite differences for space warp transformation), this scaling approach was particularly critical. Previous QMC implementations were simply incapable of executing large-scale force evaluations, making direct comparative benchmarks unavailable. Our batched implementation made these calculations feasible within our computational budget, enabling us to process approximately 20,000 molecular configurations with full force evaluation. Future optimization efforts will focus on further reducing GPU-to-host data transfers and increasing GPU kernel efficiency to improve both throughput and statistical precision.

5.3.3. Multideterminant QMC Performance. Multideterminant calculations drastically increase computational demands compared to single-determinant approaches. In principle, multideterminant wavefunction evaluations scale as $\sqrt{N_{det}}$, where N_{det} represents the number of determinants included and in this context, limited to 1.4M determinants leading to a PT2 of ∼2 mHa. This characteristic scaling implies that achieving the highest accuracy in multideterminant calculations rapidly becomes computationally prohibitive as the number of determinants increases.

We successfully executed multideterminant QMC calculations for the water monomer; however, performing analogous calculations on the water dimer exceeded our computational allocation limits, underscoring the extreme intensity of such computations. To achieve optimal performance with multideterminant wavefunctions, we allocated approximately 20 Aurora nodes per molecule and processed batches of 100 molecules per job (utilizing 2000 nodes per job submission). Our implementation reached a sustained computational throughput of 24 TFLOPS per node configuration on Aurora, corresponding to 25% of the peak theoretical performance.

These substantial accuracy gains justify the computational costs incurred, demonstrating the critical importance of multideterminant QMC methods in achieving chemically accurate reference datasets for training next-generation machine learning models.

5.4. Large Scale Quantum Dynamics Simulations with a "beyond DFT" Foundation Neural Network Model

- Summary of Contributions:

Using the QMC and sCI data produced in the project, we improved an early version of the FeNNix-Bio1 model [68] using transfer machine learning via the FeNNol library [67] to obtain a "beyond DFT" improved approach. We then coupled the improved model to the GPU-accelerated massively parallel Tinker-HP molecular dynamics package [1], [49] and leverage simultaneously its Deep-HP (neural networks), Quantum-HP (quantum nuclear effects) modules while using enhanced sampling via non-supervised adaptive sampling molecular dynamics. We tested the new approach on both Jean-Zay (H100) and Aurora(Xe) machines on the full structure of the Satellite Tobacco Virus, a 1-million atom virus. The simulation encompassed 32-million particles (1-million atom X 32 beads) since it leverages Ring Polymer quantum molecular dynamics simulation. We extended the conformational space sampling with adaptive sampling replica to decompose the global exploration problem into a set of separate MD trajectories. These latter can be restarted thanks to a selective process in order to achieve sufficient phase-space sampling.

Algorithmic Innovations

- **Bridging a Foundation Model to Ring Polymer Quantum Molecular Dynamics and to Adaptive Sampling**
  We coupled our newly developed JAX-based [16] FeNNoL library [67] that allows for the construction and inference of atomistic machine learning potentials such as FeNNix-Bio1 [68] to the Tinker-HP molecular dynamics package [1], [49] via the Deep-HP interface [39]. By doing so, we leverage the Ring Polymer Molecular Dynamics (RPMD) [36] capabilities of Tinker-HP (Quantum-HP module [70]) in combination with FeNNix-Bio1 in order to recover quantum dynamics and statistics of the Satellite Tobacco Virus (STMV) that we previously studied [15] using polarizable molecular dynamics. Accounting for these has been shown to be mandatory to recover correct properties in that context. Such simulations leverage two levels of parallelism: one at the beads level and a second one within each bead. The choice of the STMV was motivated as it is a model system

for studying viral assembly. Moreover, since it is includes a single-stranded RNA genome located within an icosahedral capsid, it is also a good model to study the structural dynamics stability of viruses [15]. Finally, to push further the sampling of the conformational space, we added on top of the discussed framework an additional layer of parallelism using automatized (unsupervised) adaptive sampling following described previously [40]. Here multiple iterations of independent molecular dynamics simulations replica are launched. The framework is an iterative process and the initial conditions at each iteration are based on the results of previous iteration steps. The selection of the initial structures at each iteration follows an adaptive, fully automated, procedure enhancing the exploration of a low-dimensional space of slow variables. This unsupervised selection step has the advantage of suppressing the problem of the choice of the initial collective variable at the beginning of the simulation. In practice the procedure maximize the conformational space exploration by penalizing areas that have already been extensively visited. The choice of the number of replica is only limited by the number of available GPUs, making adaptive sampling particularly suited for exascale computing (see section of reference [40] for a detailed discussion. On Jean-Zay, each of the bead was parallelized on 8 H100 GPUs for a total of 8x32=256 H100s. Exploiting parallelism within each bead is mandatory to fit in the memory of each GPU thanks to memory distribution. All in all, we obtain 0.11 ns/day (mixte precision) per single adaptive sampling replica, making production of tens of nanoseconds per day of such quantum dynamics simulation accessible through adaptive sampling given enough resources. To our knowledge, achieving such level of quantum dynamics is unprecedented. The Jean Zay machine having 1280 GPUs, we were able to launch 4 simultaneous replica leveraging 1024 GPUs, producing therefore 0.44 ns/day of adaptive quantum dynamics. This motivated the portage to Exascale on Aurora to increase time to solution thanks to the large availability of computing nodes.

- **Aurora Portage of Tinker-HP/Deep-HP**
  In order to port Tinker-HP/Deep-HP (which is mainly written in Fortran 90) on Aurora, we resorted to OpenMP directives. This could be naturally done as we could start from a Fortran90/OpenACC version designed to run on Nvidia GPUs. This strategy allows for efficient data transfer between CPUs and Aurora GPUs as well as parallelization of the computationally intensive kernels. Additionnally, distributed parallelism through the native 3D domain decomposition of Tinker-HP [1], [49] can be directly exploited through MPI calls between Intel Xe GPUs. Contrary to our simulations on Nvidia GPUs, we resorted to an Aurora-compatible version of Pytorch [64] and not JAX for the computation of the neural network potential energy and forces. Following the same protocol described on Jean Zay (see previous section), we obtain 0.03 ns/day (mixte precision) per single adaptive sampling replica (32 nodes, 384 GPUs).As expected since this is a new portage of the code that will benefit of further optimization over the next months, the present production per replica on Aurora is lower than the numbers per replica obtained on Jean Zay which used the reference Tinker-HP implementation. However, Aurora benefits from its Exascale capabilities and from its associated extremely large availability of computing nodes.

Large scale simulations on STMV: analysis of results
Using the STMV simulations, we open a new window on the impact of pH fluctuations on biological systems, particularly viruses, which is a critical factor influencing their structural dynamics and function. Traditional molecular dynamics (MD) approaches often simulate pH effects by manually adjusting the protonation states of ionizable residues to match the desired pH conditions or rely on specific constant-pH MD implementations [17] that are strongly depending on the force field quality. For STMV, as for any other viral particle, pH-dependent changes play a key role in the infection process, likely affecting capsid stability and RNA-capsid interactions. Simulating the pH transition from physiological to more acidic reflects conditions likely to be encountered during the early stages of viral cell entry and intracellular transport [82]. Implying the "beyond DFT" Foundation Neural Network, we simulated the transition from pH=7 to pH=5 by explicitly introducing the calculated number of free protons into the system, based on pKa estimations. This approach allows proton exchange to occur dynamically and interactively with the molecular environment. As a proof of concept, we extracted a viral STMV capsomer (33,317 atoms) and quantitatively estimated the number of proton transfer events related to the 6-ns capsomer simulation (Fig. 4), providing a dynamic portrait of the molecular response to acidification.

Extending this approach to the fully solvated STMV nucleocapsid (of $\sim$ 1 million atoms), we were able to conduct a 1.5-ns on-the-fly viral response to acidification, making a significant step forward in realistic, reactive virus modeling under varying pH conditions. Compared to the isolated capsomer, the current results reveal notable differences (Fig. 5). Specifically, we observed an increased number of protein-protein interactions within the capsid, particularly among residues that attempt to relax within the initial frames of the simulation. This behavior underscores the complexity of the full viral system. Given the large size and intricate nature of STMV, a longer simulation timescale is essential to fully capture its dynamic behavior and is currently underway on the Aurora machine.

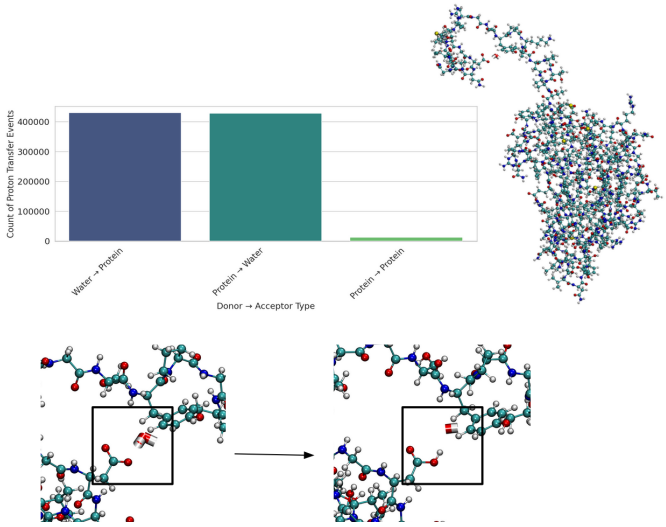

Figure 4. Number of proton transfer events among key donor–acceptor pairs: water–protein, protein–water, and protein–protein. The protein component corresponds to the STMV capsomer, illustrated to the right of the graph. The lower panel visualizations highlight a representative proton transfer event from a water molecule to a carboxyl group of a capsomer residue.

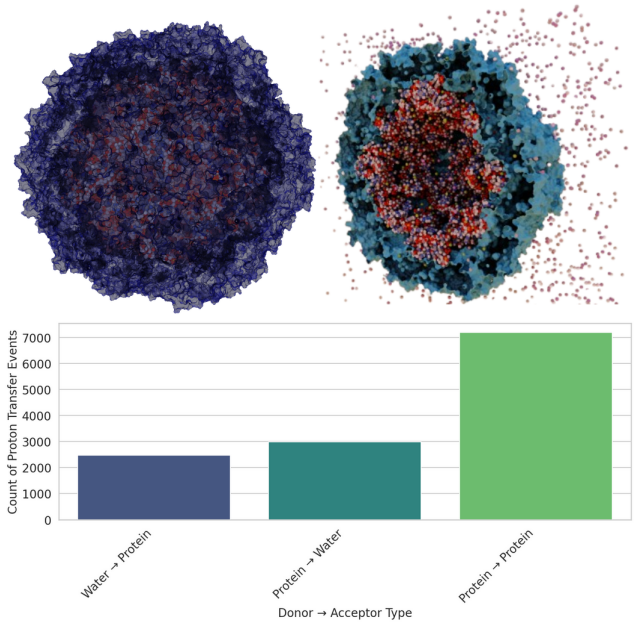

Figure 5. Number of proton transfer events among key donor–acceptor pairs: water–protein, protein–water, and protein–protein. The protein component corresponds to the STMV capsid. The top panel corresponds to the full STMV nucleocapsid (left) and a cross-sectional view of capsid–RNA assembly with surrounding ions.

6. Conclusion

We presented an end-to-end integrated computational strategy for generating synthetic quantum chemistry datasets that significantly enhance the accuracy and predictive power of atomistic Foundation Machine Learning (ML) models. Our method pushes the boundaries of quantum chemical calculations, especially Quantum Monte Carlo (QMC) and selected Configuration Interaction (sCI), to unprecedented scales. Leveraging the power of exascale computing resources, we computed energies and forces with near-exact quantum accuracy for tens of thousands of molecular configurations, establishing a new state-of-the-art in dataset generation. In a near future, the next step for such high quality dataset is to complete it with exact Full-CI/full basis quality quantum computing results once logical qubits become available. As an Easter egg, we already included in the dataset a few ”chemically accurate” fully converged energy results obtained with our GPU-accelerated quantum emulator [92]. This will generalize in the future.

Key algorithmic and computational innovations include a highly optimized multi-level batching scheme, a novel unified memory management strategy, and the optimization for GPUs of zero-variance force estimators within the QMC method. Together, these advancements enabled efficient and scalable calculations on leadership-class computing facilities such as Aurora. Specifically, our optimized GPU implementations achieved unparalleled performance, reaching sustained multi-petaflops computational throughput and overcoming significant bottlenecks traditionally associated with QMC force computations.

We demonstrated the practical applicability of our computational pipeline using FeNNix-Bio1, a new generation foundation neural network model trained with our high-fidelity datasets. Coupled with path integral-based Ring Polymer Molecular Dynamics (RPMD) and adaptive sampling, FeNNix-Bio1 successfully simulated condensed-phase quantum dynamics of a complex biological system comprising one million atoms with 32 RPMD beads to include nuclear quantum effects. This capability provides an unprecedented view into biochemical processes involving bond formation, proton transfers, and chemical reactions in realistic biosystems at quantum chemical accuracy. It is important to note that thanks to the generated dataset, the foundation model is not the only product that can be directly generated. Indeed, thanks to a recently proposed directed graph attention networks approach [19], the discussed FeNNol library can generate, within hours, new version of classical force fields such as GAFF (Generalized Amber Force Field) [98] that can be used to prepare initial simulation setups at limited cost. This work signifies a major step forward, bridging the gap between quantum accuracy and molecular dynamics simulations for biologically relevant scales. The ability to simulate multi-nanosecond quantum molecular dynamics on such large systems will profoundly impact our understanding of fundamental biochemical processes, potentially revolutionizing fields such as drug discovery, enzyme catalysis, and biomolecular engineering. Future work will

explore even larger scales and more diverse chemical spaces, laying the foundation for general-purpose ML potentials of unprecedented accuracy and versatility towards predictive molecular modeling in Chemistry and Biology.

Acknowledgments

QMC, sCI calculations, DFT computations on Aurora and some molecular dynamics simulation performed in this research used an award of computer time provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program (AB, KG, JPP). DFT Calculations on Polaris used an award of computer time provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program (RCC, JTK and YL).This research has used resources (Aurora and Polaris machines) of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. RCC, JTK and YL were supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials. CK and VA were supported by the Computational Chemical Sciences Program, under Award DE-SC0023382, funded by the U.S. Department of Energy, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division. This work was also made possible thanks to funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 810367), project EMC2 (JPP). We acknowledge EuroHPC Joint Undertaking for awarding the project IDs EHPC-DEV-2024D07-044 (E. P.) and EHPC-AI-2024A04-085 (E. P.) access to Leonardo at CINECA, Italy for DFT computations. The learning of the FeNNix-Bio1 model and some molecular dynamics simulations were performed at IDRIS, GENCI (Jean Zay machine, France) on grants No. A0150712052 (J.-P.P.) and grant GC010815453 (Grand Challenge H100 Jean Zay, J.-P.P.).

References

[1] Olivier Adjoua, Louis Lagardère, Luc-Henri Jolly, Arnaud Durocher, Thibaut Very, Isabelle Dupays, Zhi Wang, Théo Jaffrelot Inizan, Frédéric Célerse, Pengyu Ren, Jay W. Ponder, and Jean-Philip Piquemal. Tinker-hp: Accelerating molecular dynamics simulations of large complex systems with advanced point dipole polarizable force fields using gpus and multi-gpu systems. *Journal of Chemical Theory and Computation*, 17(4):2034–2053, 2021.

[2] James B. Anderson. Quantum chemistry by random walk: Higher accuracy. *The Journal of Chemical Physics*, 73(8):3897–3899, 1980.

[3] A. Annaberdiyev, G. Wang, C. A. Melton, M. C. Bennett, L. Shulenburger, and L. Mitas. A new generation of effective core potentials from correlated calculations: 3d transition metal series. *The Journal of Chemical Physics*, 149(13):134108, 2018.

[4] Roland Assaraf and Michel Caffarel. Computing forces with quantum monte carlo. *The Journal of Chemical Physics*, 113(10):4028–4034, 09 2000.

[5] Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.

[6] Albert P Bartók, Mike C Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Physical Review Letters*, 104(13):136403, 2010.

[7] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Fabio Falcioni, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdau, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, and Gábor Csányi. A foundation model for atomistic materials chemistry, 2024.

[8] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):1–11, 2022.

[9] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98(14):146401, 2007.

[10] A. Benali, L. Shulenburger, J. T. Krogel, X. Zhong, P. R. C. Kent, and O. Heinonen. Quantum monte carlo analysis of a charge ordered insulating antiferromagnet: the ti_4o_7 magnéli phase. *Phys. Chem. Chem. Phys.*, 18:18323–18335, 2016.

[11] A. Benali, L. Shulenburger, N. A. Romero, J. Kim, and O. A. von Lilienfeld. Application of diffusion monte carlo to materials dominated by van der waals interactions. *J. Chem. Theory Comput.*, 10(8):3417–3422, 2014.

[12] Anouar Benali, Kevin Gasperich, Kenneth D. Jordan, Thomas Applencourt, Ye Luo, M. Chandler Bennett, Jaron T. Krogel, Luke Shulenburger, Paul R. C. Kent, Pierre-François Loos, Anthony Scemama, and Michel Caffarel. Toward a systematic improvement of the fixed-node approximation in diffusion monte carlo for solids—a case study in diamond. *The Journal of Chemical Physics*, 153(18):184111, 2020.

[13] M. C. Bennett, C. Wang, A. Annaberdiyev, C. A. Melton, L. Shulenburger, and L. Mitas. A new generation of effective core potentials from correlated calculations: 2nd row elements. *The Journal of Chemical Physics*, 147(22):224106, 2017.

[14] Filippo Bigi, Sergey N Pozdnyakov, and Michele Ceriotti. Wigner kernels: body-ordered equivariant machine learning without a basis. *arXiv preprint arXiv:2303.04124*, 2023.

[15] Marharyta Blazhynska, Olivier Adjoua, Zoe Vetter, Louis Lagardère, and Jean-Philip Piquemal. Satellite tobacco mosaic virus: Revealing environmental drivers of capsid and nucleocapsid stability using high-resolution simulations. *bioRxiv*, pages 2025–01, 2025.

[16] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.

[17] Pavel Buslaev, Noora Aho, Anton Jansen, Paul Bauer, Berk Hess, and Gerrit Groenhof. Best practices in constant ph md simulations: Accuracy and sampling. *Journal of Chemical Theory and Computation*, 18(10):6134–6147, 2022. PMID: 36107791.

[18] Michele Ceriotti, Wei Fang, Peter G Kusalik, Ross H McKenzie, Angelos Michaelides, Miguel A Morales, and Thomas E Markland. Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges. *Chemical Reviews*, 116(13):7529–7550, 2016.

[19] Gong Chen, Théo Jaffrelot Inizan, Thomas Plé, Louis Lagardère, Jean-Philip Piquemal, and Yvon Maday. Advancing force fields parameterization: A directed graph attention networks approach. *Journal of Chemical Theory and Computation*, 20(13):5558–5569, 2024. PMID: 38875012.

[20] Stefan Chmiela, Huziel E Sauceda, Klaus-Robert Müller, and Alexandre Tkatchenko. Towards exact molecular dynamics simulations with machine-learned force fields. *Nature Communications*, 9(1):1–10, 2018.

[21] Christian Devereux, Justin S. Smith, Kate K. Huddleston, Kipton Barros, Roman Zubatyuk, Olexandr Isayev, and Adrian E. Roitberg. Extending the applicability of the ani deep learning molecular potential to sulfur and halogens. *Journal of Chemical Theory and Computation*, 16(7):4192–4202, 2020.

[22] Pavlo O Dral, Fuchun Ge, Yi-Fan Hou, Peikun Zheng, Yuxinxin Chen, Mario Barbatti, Olexandr Isayev, Cheng Wang, Bao-Xin Xue, Max Pinheiro Jr, et al. Mlatom 3: A platform for machine learning-enhanced computational chemistry simulations and workflows. *Journal of Chemical Theory and Computation*, 20(3):1193–1213, 2024.

[23] Matúš Dubecký, Lubos Mitas, and Petr Jurečka. Noncovalent Interactions by Quantum Monte Carlo. *Chemical Reviews*, 116(9):5188–5215, May 2016.

[24] Peter Eastman, Pavan Kumar Behara, David L. Dotson, Zhi Meng, Camilo Velez-Vega, Shuzhe Zhang, Aditya Jain, Christian Kramer, Michael J. Robertson, Jason Swails, John D. Chodera, Bentley Sheets, Ron O. Drorr, and Vijay S. Pande. Spice, a dataset of drug-like molecules and peptides for training machine learning potentials. *Scientific Data*, 10(1):193, 2023.

[25] Claudia Filippi, Roland Assaraf, and Saverio Moroni. Simple formalism for efficient derivatives and multi-determinant expansions in quantum monte carlo. *The Journal of Chemical Physics*, 144(19):194105, 2016.

[26] Claudia Filippi and C. J. Umrigar. Correlated sampling in quantum monte carlo: A route to forces. *Phys. Rev. B*, 61:R16291–R16294, Jun 2000.

[27] W. M. C Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33–83, 2001.

[28] K. Foyevtsova, J. T. Krogel, J. Kim, P. R. C. Kent, E. Dagotto, and F. A. Reboredo. Ab initio quantum monte carlo calculations of spin superexchange in cuprates: The benchmarking case of ca_2cuo_3. *Phys. Rev. X*, 4:031003, 2014.

[29] Scott M. Garner and Eric Neuscamman. A variational monte carlo approach for core excitations. *The Journal of Chemical Physics*, 153(14):144108, 2020.

[30] Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34:6790–6802, 2021.

[31] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pages 1263–1272. PMLR, 2017.

[32] Emmanuel Giner, Anthony Scemama, and Michel Caffarel. Using perturbatively selected configuration interaction in quantum monte carlo calculations. *Canadian Journal of Chemistry*, 91(9):879–885, 2013.

[33] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, 2011.

[34] Andrea Grisafi and Michele Ceriotti. Incorporating long-range physics in atomic-scale machine learning. *The Journal of Chemical Physics*, 151(20):204105, 2019.

[35] Andrea Grisafi, Jigyasa Nigam, and Michele Ceriotti. Multi-scale approach for the prediction of atomic scale properties. *Chemical Science*, 12(6):2078–2090, 2021.

[36] Scott Habershon, David E Manolopoulos, Thomas E Markland, and Thomas F Miller III. Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space. *Annual review of physical chemistry*, 64(1):387–413, 2013.

[37] Adam A. Holmes, C. J. Umrigar, and Sandeep Sharma. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.*, 147(16):164111, Oct 2017.

[38] B. Huron, J.P. Malrieu, and P. Rancurel. Perturbation calculation of transition moments. application to h2 and mgo. *Chemical Physics*, 3(2):277–283, 1974.

[39] Théo Jaffrelot Inizan, Thomas Plé, Olivier Adjoua, Pengyu Ren, Hatice Gokcan, Olexandr Isayev, Louis Lagardère, and Jean-Philip Piquemal. Scalable hybrid deep neural networks/polarizable potentials biomolecular simulations including long-range effects. *Chemical Science*, 14:5438–5452, 2023.

[40] Théo Jaffrelot Inizan, Frédéric Célerse, Olivier Adjoua, Dina El Ahdab, Luc-Henri Jolly, Chengwen Liu, Pengyu Ren, Matthieu Montes, Nathalie Lagarde, Louis Lagardère, Pierre Monmarché, and Jean-Philip Piquemal. High-resolution mining of the sars-cov-2 main protease conformational space: supercomputer-driven unsupervised adaptive sampling. *Chem. Sci.*, 12:4889–4907, 2021.

[41] Weile Jia, Han Wang, Mohan Chen, Denghui Lu, Lin Lin, Roberto Car, E Weinan, and Linfeng Zhang. Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning. In *SC20: International conference for high performance computing, networking, storage and analysis*, pages 1–14. IEEE, 2020.

[42] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[43] P. R. C. Kent, Abdulgani Annaberdiyev, Anouar Benali, M. Chandler Bennett, Edgar Josué Landinez Borda, Peter Doak, Hongxia Hao, Kenneth D. Jordan, Jaron T. Krogel, Ilkka Kylänpää, Joonho Lee, Ye Luo, Fionn D. Malone, Cody A. Melton, Lubos Mitas, Miguel A. Morales, Eric Neuscamman, Fernando A. Reboredo, Brenda Rubenstein, Kayahan Saritas, Shiv Upadhyay, Guangming Wang, Shuai Zhang, and Luning Zhao. QMCPACK: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion quantum monte carlo. *The Journal of Chemical Physics*, 152(17):174105, May 2020.

[44] Danish Khan, Anouar Benali, Scott Y. H. Kim, Guido Falk von Rudorff, and O. Anatole von Lilienfeld. Quantum mechanical dataset of 836k neutral closed shell molecules with upto 5 heavy atoms from cnofsipsclbr, 2024.

[45] J. Kim, A. Baczewski, T. Beaudet, A. Benali, C. Bennett, M. Berrill, N. Blunt, E. J. L. Borda, M. Casula, D. Ceperley, S. Chiesa, B. K. Clark, R. Clay, K. Delaney, M. Dewing, K. Esler, H. Hao, O. Heinonen, P. R. C. Kent, J. T. Krogel, I. Kylanpaa, Y. W. Li, M. G. Lopez, Y. Luo, F. Malone, R. Martin, A. Mathuriya,

J. McMinis, C. Melton, L. Mitas, M. A. Morales, E. Neuscamman, W. Parker, S. Flores, N. A. Romero, B. Rubenstein, J. Shea, H. Shin, L. Shulenburger, A. Tillack, J. Townsend, N. Tubman, B. van der Goetz, J. Vincent, D. C. Yang, Y. Yang, S. Zhang, and L. Zhao. Qmcpack : An open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules, and solids. *J. Phys.: Condens. Matter*, 30:195901, 2018.

[46] Emir Kocer, Tsz Wai Ko, and Jörg Behler. Neural network potentials: A concise overview of methods. *Annual review of physical chemistry*, 73:163–186, 2022.

[47] Jindřich Kolorenč and Lubos Mitas. Applications of quantum monte carlo methods in condensed systems. *Reports on Progress in Physics*, 74(2):026502, 2011.

[48] I. Kylänpää, J. Balachandran, G. Panchapakesan, O. Heinonen, P. R. C. Kent, and J. T. Krogel. Accuracy of ab initio electron correlation and electron densities in vanadium dioxide. *Phys. Rev. Materials*, 1:065408, 2017.

[49] Louis Lagardère, Luc-Henri Jolly, Filippo Lipparini, Félix Aviat, Benjamin Stamm, Zhifeng F. Jing, Matthew Harger, Hedieh Torabifard, G. Andrés Cisneros, Michael J. Schnieders, Nohad Gresh, Yvon Maday, Pengyu Y. Ren, Jay W. Ponder, and Jean-Philip Piquemal. Tinker-hp: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chemical Science*, 9:956–972, 2018.

[50] Rui Li, Qiming Sun, Xing Zhang, and Garnet Kin-Lic Chan. Introducing gpu acceleration into the python-based simulations of chemistry framework. *The Journal of Physical Chemistry A*, 129(5):1459–1468, 2025. PMID: 39846468.

[51] P. F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, and D. Jacquemin. A mountaineering strategy to excited states: highly-accurate energies and benchmarks for medium size molecules,. *J. Chem. Theory Comput.*, 16:1711, 2020.

[52] Pierre-François Loos, Anthony Scemama, Aymeric Blondel, Yann Garniron, Michel Caffarel, and Denis Jacquemin. A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.*, 14(8):4360–4379, jul 2018.

[53] Pierre-François Loos, Martial Boggio-Pasqua, Anthony Scemama, Michel Caffarel, and Denis Jacquemin. Reference energies for double excitations. *Journal of Chemical Theory and Computation*, 15(3):1939–1956, 2019.

[54] Pierre-François Loos, Anthony Scemama, and Denis Jacquemin. The quest for highly accurate excitation energies: A computational perspective. *The Journal of Physical Chemistry Letters*, 11(6):2374–2383, 2020. PMID: 32125872.

[55] Nicholas Lubbers, Justin S Smith, and Kipton Barros. Hierarchical modeling of molecular energies using a deep neural network. *The Journal of Chemical Physics*, 148(24), 2018.

[56] Narbe Mardirossian and Martin Head-Gordon. ωb97m-v: A combinatorially optimized, range-separated hybrid, meta-gga density functional with vv10 nonlocal correlation. *The Journal of Chemical Physics*, 144(21):214110, 2016.

[57] Narbe Mardirossian and Martin Head-Gordon. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics*, 115(19):2315–2372, 2017.

[58] Amrita Mathuriya, Ye Luo, Anouar Benali, Luke Shulenburger, and Jeongnim Kim. Optimization and parallelization of b-spline based orbital evaluations in qmc on multi/many-core shared memory processors. In *2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 213–223, 2017.

[59] Amrita Mathuriya, Ye Luo, Raymond C. Clay, Anouar Benali, Luke Shulenburger, and Jeongnim Kim. Embracing a new era of highly efficient and productive quantum monte carlo simulations. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '17, New York, NY, USA, 2017. Association for Computing Machinery.

[60] Nastasia Mauger, Thomas Plé, Louis Lagardère, Sara Bonella, Étienne Mangaud, Jean-Philip Piquemal, and Simon Huppert. Nuclear quantum effects in liquid water at near classical computational cost using the adaptive quantum thermal bath. *The Journal of Physical Chemistry Letters*, 12(34):8285–8291, 2021.

[61] Miguel A. Morales, Jeremy McMinis, Bryan K. Clark, Jeongnim Kim, and Gustavo E. Scuseria. Multideterminant wave functions in quantum monte carlo. *Journal of Chemical Theory and Computation*, 8(7):2181–2188, 2012.

[62] Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications*, 14(1):579, 2023.

[63] Leon Otis, Isabel Craig, and Eric Neuscamman. A hybrid approach to excited-state-specific variational monte carlo and doubly excited states. *The Journal of Chemical Physics*, 153(23):234105, 2020.

[64] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[65] Sergio D. Pineda Flores and Eric Neuscamman. Excited state specific multi-slater jastrow wave functions. *The Journal of Physical Chemistry A*, 123(8):1487–1497, 2019. PMID: 30702890.

[66] Max Pinheiro, Fuchun Ge, Nicolas Ferré, Pavlo O Dral, and Mario Barbatti. Choosing the right molecular machine learning potential. *Chemical Science*, 12(43):14396–14413, 2021.

[67] Thomas Plé, Olivier Adjoua, Louis Lagardère, and Jean-Philip Piquemal. FeNNol: An efficient and flexible library for building force-field-enhanced neural network potentials. *The Journal of Chemical Physics*, 161(4):042502, 07 2024.

[68] Thomas Plé, Anouar Benali, Olivier Adjoua, Evgeny Posenitskiy, Corentin Villot, Louis Lagardère, and Jean-Philip Piquemal. A foundation model for accurate atomistic simulations in drug design, 2025.

[69] Thomas Plé, Nastasia Mauger, Olivier Adjoua, Théo Jaffrelot Inizan, Louis Lagardère, Simon Huppert, and Jean-Philip Piquemal. Routine molecular dynamics simulations including nuclear quantum effects: from force fields to machine learning potentials. *Journal of Chemical Theory and Computation*, 19(5):1432–1445, 2023.

[70] Thomas Plé, Nastasia Mauger, Olivier Adjoua, Théo Jaffrelot Inizan, Louis Lagardère, Simon Huppert, and Jean-Philip Piquemal. Routine molecular dynamics simulations including nuclear quantum effects: From force fields to machine learning potentials. *Journal of Chemical Theory and Computation*, 19(5):1432–1445, 2023.

[71] Zhuoran Qiao, Anders S Christensen, Matthew Welborn, Frederick R Manby, Anima Anandkumar, and Thomas F Miller III. Unite: Unitary n-body tensor equivariant network with applications to quantum chemistry. *arXiv preprint arXiv:2105.14655*, 2021.

[72] Krishnan Raghavachari, Gary W. Trucks, John A. Pople, and Martin Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, 157(6):479–483, 1989.

[73] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester. Fixed-node quantum monte carlo for molecules. *J. Chem. Phys.*, 77:5593–5603, Jun 1982.

[74] Peter J. Reynolds, Jan Tobochnik, and Harvey Gould. Diffusion quantum monte carlo. *Computers in Physics*, 4(6):662–668, 1990.

[75] Anthony Scemama, Anouar Benali, Denis Jacquemin, Michel Caffarel, and Pierre-François Loos. Excitation energies from diffusion monte carlo using selected configuration interaction nodes. *J. Chem. Phys.*, 149(3):034108, jul 2018.

[76] Anthony Scemama, Michel Caffarel, Anouar Benali, Denis Jacquemin, and Pierre-François Loos. Influence of pseudopotentials on excitation energies from selected configuration interaction and diffusion Monte Carlo. *Res. Chem.*, 1:100002, 2019.

[77] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30, 2017.

[78] Kristof T Schütt, Farhad Arbabzadah, Stefan Chmiela, Klaus R Müller, and Alexandre Tkatchenko. Quantum-chemical insights from deep tensor neural networks. *Nature communications*, 8(1):13890, 2017.

[79] Kristof T Schütt, Stefaan SP Hessmann, Niklas WA Gebauer, Jonas Lederer, and Michael Gastegger. Schnetpack 2.0: A neural network toolbox for atomistic machine learning. *The Journal of Chemical Physics*, 158(14), 2023.

[80] Tobias Schäfer, Andreas Irmler, Alejandro Gallo, and Andreas Grüneis. Understanding discrepancies of wavefunction theories for large molecules, 2024.

[81] Khosrow Shakouri, Jörg Behler, Jörg Meyer, and Geert-Jan Kroes. Accurate neural network description of surface phonons in reactive gas–surface dynamics: N2+ ru (0001). *The journal of physical chemistry letters*, 8(10):2131–2136, 2017.

[82] Jinbo Shen, Yonglun Zeng, Xiaohong Zhuang, Lei Sun, Xiaoqiang Yao, Peter Pimpl, and Liwen Jiang. Organelle ph in the arabidopsis endomembrane system. *Molecular Plant*, 6(5):1419–1437, 2013.

[83] H. Shin, A. Benali, Y. Luo, E. Crabb, A. Lopez-Bezanilla, L. E. Ratcliff, A. M. Jokisaari, and O. Heinonen. Zirconia and hafnia polymorphs: Ground-state structural properties from diffusion monte carlo. *Phys. Rev. Materials*, 2:075001, 2018.

[84] H. Shin, J. Kim, H. Lee, O. Heinonen, A. Benali, and Y. Kwon. Nature of interlayer binding and stacking of $sp-sp^2$ hybridized carbon layers: A quantum monte carlo study. *J. Chem. Theory Comput.*, 13(11):5639–5646, 2017.

[85] H. Shin, Y. Luo, P. Ganesh, J. Balachandran, J. T. Krogel, P. R. C. Kent, A. Benali, and O. Heinonen. Electronic properties of doped and defective nio: A quantum monte carlo study. *Phys. Rev. Materials*, 1:173603, 2017.

[86] Andreas Singraber, Jörg Behler, and Christoph Dellago. Library-based lammps implementation of high-dimensional neural network potentials. *Journal of Chemical Theory and Computation*, 15(3):1827–1840, 2019. PMID: 30677296.

[87] Erik Slootman, Igor Poltavsky, Ravindra Shinde, Juan Cocomello, Saverio Moroni, Alexandre Tkatchenko, and Claudia Filippi. Accurate quantum monte carlo forces for machine-learned force fields: Ethanol as a benchmark. *Journal of Chemical Theory and Computation*, 20(14):6020–6027, 2024.

[88] Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chemical Science*, 8(4):3192–3203, 2017.

[89] Justin S. Smith, Benjamin T. Nebgen, Roman Zubatyuk, Nicholas Lubbers, Christian Devereux, Kipton Barros, Sergei Tretiak, Olexandr Isayev, and Adrian E. Roitberg. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nature Communications*, 10(1):2903, 2019.

[90] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. Pyscf: the python-based simulations of chemistry framework, 2017.

[91] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. PySCF: the Python-Based Simulations of Chemistry Framework. *WIREs Comput. Mol. Sci.*, 8(1):e1340, 2018.

[92] Diata Traore, Olivier Adjoua, César Feniou, Ioanna-Maria Lygatsika, Yvon Maday, Evgeny Posenitskiy, Kerstin Hammernik, Alberto Peruzzo, Julien Toulouse, Emmanuel Giner, et al. Shortcut to chemically accurate quantum computing via density-based basis-set correction. *Communications Chemistry*, 7(1):269, 2024.

[93] Norm M Tubman, Joonho Lee, Tyler Y Takeshita, Martin Head-Gordon, and K Birgitta Whaley. A deterministic alternative to the full configuration interaction quantum monte carlo method. *The Journal of Chemical Physics*, 145(4):044112, 2016.

[94] Oliver T Unke, Stefan Chmiela, Michael Gastegger, Kristof T Schütt, Huziel E Sauceda, and Klaus-Robert Müller. Spookynet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nature Communications*, 12(1):1–14, 2021.

[95] Konstantinos D. Vogiatzis, Dongxia Ma, Jeppe Olsen, Laura Gagliardi, and Wibe A. de Jong. Pushing configuration-interaction to the limit: Towards massively parallel mcscf calculations. *The Journal of Chemical Physics*, 147(18):184111, 11 2017.

[96] Lucas K Wagner and Dhawal Pathak. Improved quantum Monte Carlo forces for interaction energy curves and gap energetics. *The Journal of Chemical Physics*, 141(6):064102, 2014.

[97] Guangming Wang, Abdulgani Annaberdiyev, Cody A. Melton, M. Chandler Bennett, Luke Shulenburger, and Lubos Mitas. A new generation of effective core potentials from correlated calculations: 4s and 4p main group elements and first row additions. *The Journal of Chemical Physics*, 151(14):144110, 10 2019.

[98] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[99] Xiaojie Wu, Qiming Sun, Zhichen Pu, Tianze Zheng, Wenzhi Ma, Wen Yan, Yu Xia, Zhengxiao Wu, Mian Huo, Xiang Li, Weiluo Ren, Sheng Gong, Yumin Zhang, and Weihao Gao. Enhancing gpu-acceleration in the python-based simulations of chemistry frameworks. *WIREs Computational Molecular Science*, 15(2):e70008, 2025. e70008 CMS-1146.R2.

[100] Yinuo Yang, Shuhao Zhang, Kavindri D Ranasinghe, Olexandr Isayev, and Adrian E Roitberg. Machine learning of reactive potentials. *Annual Review of Physical Chemistry*, 75(1):371–395, 2024.

[101] Viktor Zaverkin and J Kästner. Gaussian moments as physically inspired molecular descriptors for accurate and scalable machine learning potentials. *Journal of Chemical Theory and Computation*, 16(8):5410–5421, 2020.

[102] Jinzhe Zeng, Duo Zhang, Denghui Lu, Pinghui Mo, Zeyu Li, Yixiao Chen, Marián Rynik, Li'ang Huang, Ziyao Li, Shaochen Shi, et al. Deepmd-kit v2: A software package for deep potential models. *The Journal of Chemical Physics*, 159(5), 2023.

[103] Yukun Zhang, Xiaoming Zhang, Jinzhao Sun, Heng Lin, Yifei Huang, Dingshun Lv, and Xiao Yuan. Fault-tolerant quantum algorithms for quantum molecular systems: A survey. *arXiv preprint arXiv:2502.02139*, 2025.

[104] Roman Zubatyuk, Justin S Smith, Jerzy Leszczynski, and Olexandr Isayev. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. *Science advances*, 5(8):eaav6490, 2019.

[105] Roman Zubatyuk, Justin S. Smith, Jerzy Leszczynski, and Olexandr Isayev. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. *Science Advances*, 5(8):eaav6490, 2019.

[106] Örs Legeza, Andor Menczer, Ádám Ganyecz, Miklós Antal Werner, Kornél Kapás, Jeff Hammond, Sotiris S. Xantheas, Martin Ganahl, and Frank Neese. Orbital optimization of large active spaces via ai-accelerators, 2025.