Article

A Perspective on Sustainable Computational Chemistry Software Development and Integration

Rosa Di Felice, Maricris L. Mayes, Ryan M. Richard, David B. Williams-Young, Garnet Kin-Lic Chan, Wibe A. de Jong, Niranjan Govind, Martin Head-Gordon, Matthew R. Hermes, Karol Kowalski,* Xiaosong Li, Hans Lischka, Karl T. Mueller, Erdal Mutlu, Anders M. N. Niklasson, Mark R. Pederson, Bo Peng, Ron Shepard, Edward F. Valeev, Mark van Schilfgaarde, Bess Vlaisavljevich, Theresa L. Windus, Sotiris S. Xantheas, Xing Zhang, and Paul M. Zimmerman

Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations

ABSTRACT: The power of quantum chemistry to predict the ground and excited state properties of complex chemical systems has driven the development of computational quantum chemistry software, integrating advances in theory, applied mathematics, and computer science. The emergence of new computational paradigms associated with exascale technologies also poses significant challenges that require a flexible forward strategy to take full advantage of existing and forthcoming computational resources. In this context, the sustainability and interoperability of computational chemistry software development are among the most pressing issues. In this perspective, we discuss software infrastructure needs and investments with an eye to fully utilize exascale resources and provide unique computational tools for next-generation science problems and scientific discoveries.
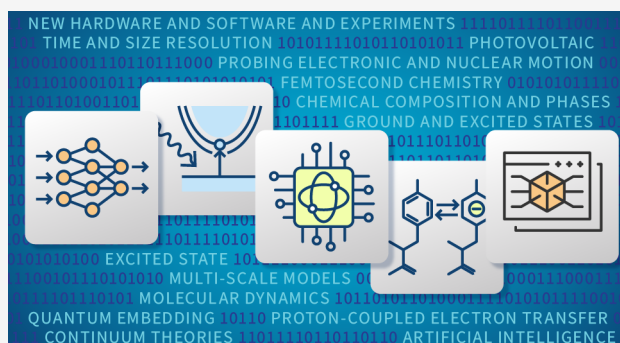
INTRODUCTION

The emergence of new computational technologies offers a unique opportunity to tackle important chemistry problems using methodologies and their implementations designed to take full advantage of exascale computational resources. Access to these state-of-the-art resources provides a means to define a new level of interoperability to propagate interaction-driven models across spatial and temporal scales. In this perspective, we discuss the role and importance of the sustainability of computational chemistry software development, focusing on scientific challenges, theoretical formulations, algorithms, languages/programming models, hardware, and developer communities. Computational chemistry suites developed over the past decade (for example, ABINIT, ACES, BAGEL, CASTEP, C2PK, CHARMM, CFOUR, COLUMBUS, DFTB +, DIRAC, e^T1.0, FLOSIC, GAMESS, Gaussian, Molcas, Molpro, MPQC, MRCC, NECI, NWChem, Octopus, ONE-TEP, ORCA, PSI4, PySCF, Q-Chem, QMCPACK, Quantum ESPRESSO, SIESTA, TeraChem, Tyrbomole, VASP, and others; see refs 1−32) have integrated novel electronic structure methods of increasing computational complexity with applied mathematics algorithms and efficient computer science tools to utilize existing and emerging computing architectures effectively. Co-design efforts have been key to integrating novel approaches in electronic structure theories, computer science, and applied mathematics. The latter has also been supported by new programming models, focusing on modularity, interlanguage operability, and application programming interface design, rather than the creation of monolithic programs. Taken together, all of these advances have allowed developers and users to tackle complex chemical problems, develop workflows, and provide insights into how existing methodologies can be extended to the next level of complexity.

Here, we concentrate on future directions for the coexistence and cross-fertilization of various computational models and emerging technologies. Possible roadblocks that may result in the loss of sustainability of the software development process are also considered. Special attention is given to the elements of the stewardship program that are directly related to the development, curation, hardening, and distribution of the scientific
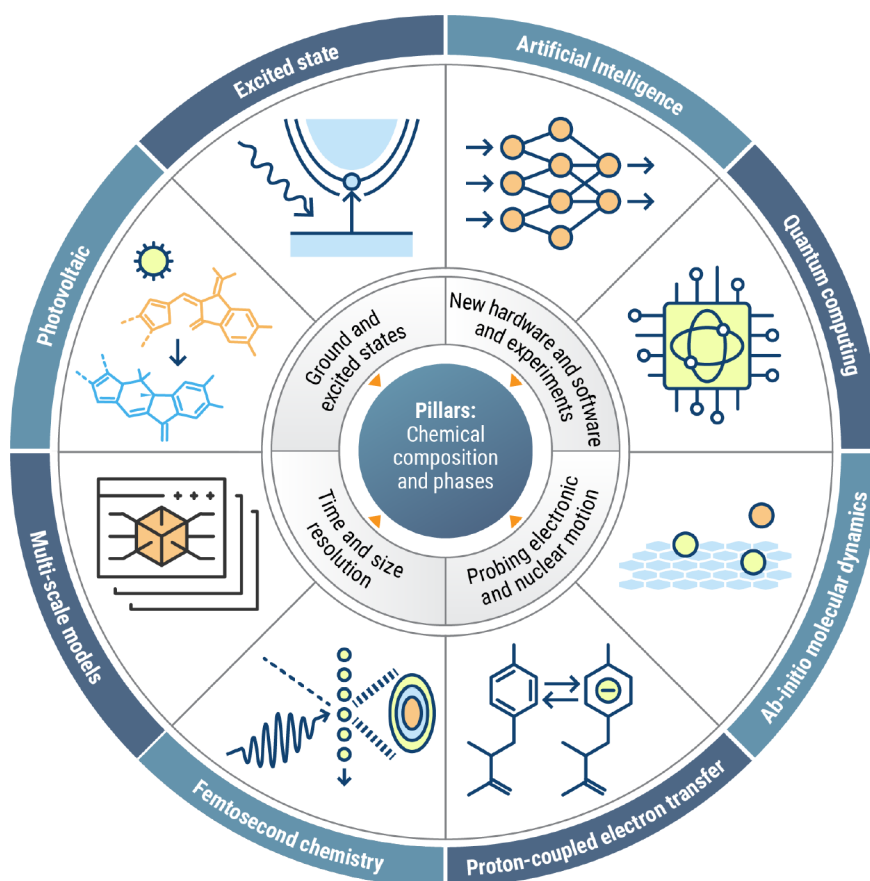
Figure 1. Visualization of layers of scientific challenges. The central core contains fundamental challenges based on the chemical elements that compose target systems, which determine their computational complexity. These core pillars are articulated in the Scientific Challenges and Discoveries section. The intermediate layer contains transversal issues to all pillars of different origins ranging from the quantum mechanical nature of the problems to the emergence of disruptive new technologies. The outer circle illustrates examples of challenging systems and problems from the recent scientific literature.

software needed for the efficient and productive utilization of next-generation high-performance computing (HPC) systems. Lastly, we emphasize the importance of training the next generation of domain and computer scientists.

Our strategy for developing software for simulating complex chemical processes is based on specialized modules and libraries that consider complicated dependencies between particular classes of methodologies and corresponding software components defined through solvers, models and algebraic complexity, numerical representations, runtimes, parallelization strategies, and specialized computational kernels. The following topics have to be properly addressed to meet the future needs in computational chemistry:

- Effective utilization of sparsity of multidimensional tensors in reduced-scaling models.
- Seamless integration of several levels of parallelism in the computational workflows.
- Operation/communication optimization of sparse models and workflows.
- New domain-specific languages (DSLs) to deal with the algebraic complexity of models that effectively utilize sparsity in high-accuracy computational models.
- Support for parallel models on new architectures and accelerators.
- Coupling of methods to enable multiscale modeling on complex chemical landscapes.

- Support for programming models that enable rapid prototyping and interfacing between new methods, algorithms, and applications, that utilize all the above advances.

The importance of these developments will significantly increase with the routine utilization of exascale platforms planned for the next decade.

To ensure software integrity, continuous integration with stringent testing requirements, code analysis tools, coding standards, and other best practices will be crucial. In the near future, extending profiling capabilities to record detailed execution information will allow for the investigation of bottlenecks while keeping a record of the execution choices made over the course of the application run. These data will be crucial for reproducibility as they will allow us to re-evaluate the problem with the same execution choices. Furthermore, the existence of an engaged user base and associated consulting support is crucial for ensuring the scientific reproducibility of the results obtained with the developed software and assists in long-term planning for development that caters to the needs of the scientific community at large.

The establishment of a sustainable software ecosystem in the future calls for the implementation of targeted support mechanisms to facilitate the design, development, maintenance, and scientific applications. Due to the rapid evolution of hardware technology, new factors must be taken into account to

guide the design of next-generation software. The integration of multidisciplinary research, including advanced theoretical formulations, applied mathematics, high-performance computing, and computer science, is an essential component of the strategy, with a focus on addressing outstanding problems in chemical and materials science. Simultaneously, the provision of training for the next generation of researchers is crucial. Collaboration with industrial partners in relevant areas such as cloud computing may offer an alternative means of enabling scalable research software for the broader user community. This perspective builds upon earlier discussions on software development strategies,[33−35] incorporating the experiences gained in the development of exa-scale software and broader aspects of scientific software sustainability.[1−32]

We first focus on scientific challenges that require novel computational chemistry approaches and tools; this is followed by theoretical methodologies, programming models, and a sustainable computational chemistry ecosystem.

■ SCIENTIFIC CHALLENGES AND DISCOVERIES

Many scientific challenges (see Figure 1 for examples) are triggered by technological demand, which sets the goals of computational chemistry/physics research. While the list of scientific applications is large, it is encouraging that they can be tackled with novel and robust theoretical, applied mathematics, and computational tools that provide predictive modeling capacities of the underlying electronic and structural properties of molecules and materials in tandem with emerging computational technologies. Here, we list a few areas that can capitalize on these advancements.

(1) Battery Technology: Research in this context is aimed at developing more power, through cheaper processes that produce denser and lighter batteries. A typical battery is comprised of two electrodes (anode and cathode), a separator between the electrodes, and an electrolyte. Ions emitted by one electrode reach the other electrode through the electrolyte, cyclically. New active materials are needed for electrodes and electrolytes,[36−38] while the use of multicharged ions is desired to continue increasing the power generation in a safe and efficient manner.

(2) Clean (solar) energy: Most commercial solar panels have an efficiency between 15% and 21%, with peaks around 50%. Research on materials and interfaces, as well as on electron and energy transfer, aided by computational tools,[34,39] aims at raising these numbers while reducing the cost for users. In addition, the development of liquid fuels using solar-driven processes has gained significant interest as a method to use our existing liquid fuel infrastructure in a compatible manner.[40]

(3) Design of Catalysts:[41] 90% of chemical processes producing commercial chemicals involve catalytic processes,[42] and catalysts are usually tailored for specific applications. Catalysts enable faster and more efficient chemical reactions, by lowering activation energies. By modifying branching ratios, they are also capable of enhancing the amount of desired products while at the same time reducing the amount of undesired products. Computational design of catalysts[34] has the potential to accelerate the affordable realization of desired reactions in search of new biodegradable plastics, new pharmaceuticals, environmentally safe fuels and fertilizers, and viable solutions to the critical materials problem.[43,44]

(4) Rational Materials Design:[45] High-throughput calculations of the structural and electronic properties of materials, along with accurate prediction of various spectroscopic features, produce databases that can be used to synthesize an advanced material that is optimal for a specific application such as, for instance, a material harder than diamond. For an optimal management of this strategy, the databases should include many different material categories, including as many chemical elements as possible in different abundances and in different crystal symmetries. Furthermore, they should include not only bulk materials but also surfaces and interfaces. Materials design includes the development of materials and chemicals for quantum hardware,[46] as well as the exploitation of quantum hardware to predict the properties of materials and chemical systems and reactions.[47]

(5) Biological Chemistry: Studies in this context may impact the production of microbes for energy applications and recycling processes,[48] as well as understanding fundamental genetic mechanisms[49] and finding solutions to genetic defects.[50,51]

(6) Separation Science:[52] Understanding the origin of and achieving selectivity in separation science is a fundamental issue, where it is important to discern small differences in weak interactions between molecules and substrates. The challenge is especially relevant to carbon dioxide and methane separation since, like the desirable atmospheric components (molecular nitrogen and oxygen), the interactions between substrate and adsorbate depend largely on the polarizability of the adsorbate. It is also essential in the separation of materials that are critical for our technological society, such as the rare earth elements.[53−55]

(7) Heavy Element Chemistry: The f-elements that appear at the bottom of the periodic table (lanthanides and actinides) are relevant for technologies related to energy and national security. f-electron systems are characterized by the simultaneous presence of itinerant (delocalized) and highly localized states and interactions between them.[54,56−58] While lanthanide chemistry can be mostly understood by studying the impact of changing the size of the metal atom to tune the properties of a molecular complex, actinides do not exhibit the same periodic trends, a fact that requires the use of advanced electronic structure methods beyond mean-field approaches and accurate treatments of relativistic and correlation effects.[59,60]

(8) Gas Phase Chemistry:[61−63] Most energy production processes involve combustion, a gas-phase chemical process even with liquid and solid fuels; those fuels may be either renewable (e.g., biofuels) or nonrenewable (e.g., fossil fuels). In addition to energy production, the characterizations of soot formation, nitrogen oxides, and other reaction products are also important to a broad range of scientific challenges.

(9) Strong Field Physics: Strong field interactions between ultrafast intense fields (attosecond pulses) and matter has led to a plethora of new physical phenomena, such as multiphoton ionization, above-threshold ionization, nonsequential double ionization, high harmonic generation, attosecond pulse generation, coherent X-ray generation, etc. This has led to the field of attosecond science, which represents a new frontier in fundamental ultrafast studies

in the atomic, molecular, and the condensed phases requiring new theoretical developments.[64−69] In this context, the challenge of describing unbound electronic states is particularly compelling.[70,71]

(10) Ultrafast Science: Emerging X-ray free electron laser (XFEL) sources like the Linac Coherent Light Source (LCLS)[72−76] offer new types of probes of matter with unprecedented spatial and temporal resolutions. The ability of these probes to spatially resolve coherent motion in complex systems is crucial for harnessing electronic, vibrational, vibronic (coupled electronic/vibrational) coherences, coherent solute−solvent motions, charge flow between electron/proton donor−acceptor sites, and control of intramolecular electron and proton motion on ultrafast time scales.

Addressing the challenges outlined in the examples above requires the close integration between state-of-the-art theoretical methodologies and high-performance-computing tools to take advantage of exascale architectures.

■ THEORETICAL METHODOLOGIES

In this Section, we provide a brief overview of computational chemistry methodologies that, due to their efficiency in capturing correlation effects, their ability to scale across time and spatial domains, and their potential in utilizing exascale computational resources, are vital elements in existing and forthcoming computing hardware infrastructures (see Figure 2).
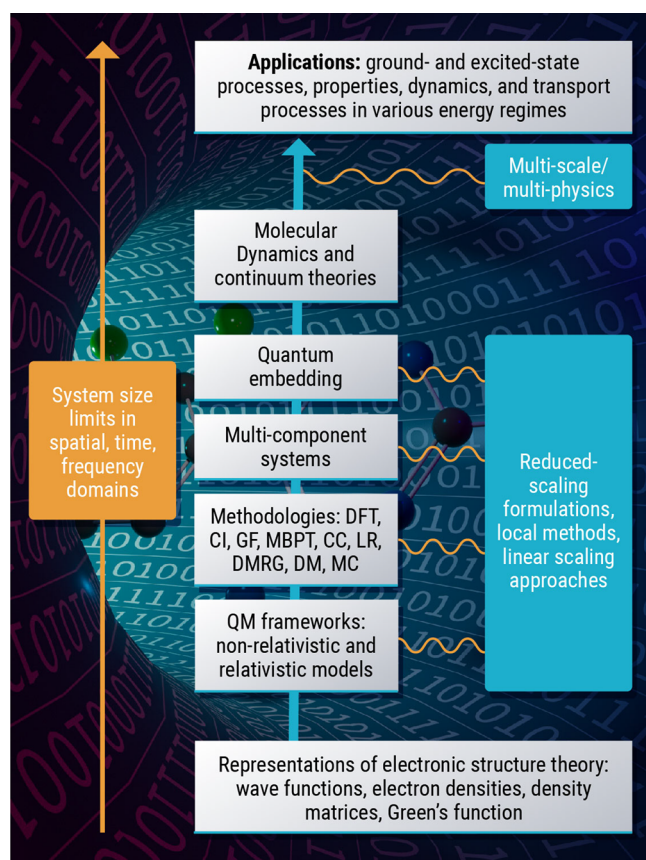
Figure 2. A "connected diagram" illustrates the hierarchy of methods and synergies between various formulations toward providing the required accuracy level for the ever-growing complexity and size of outstanding chemical problems.

These formulations are crucial in addressing the science problems discussed in the previous Section. The wealth of methodologies employed in modeling chemical processes also requires discussing use cases and possible limitations, illustrated by the juxtaposition of the salient features of electronic structure methods.

Quantum Chemistry Frameworks. The nonrelativistic and relativistic frameworks are the two foundations in electronic structure theory that define the fundamental interactions in chemical and materials systems:

The Schrödinger framework is the foundation of nonrelativistic electronic structure theory. The underlying one-component electronic wave functions, in the exact limit, are eigenfunctions of the total spin angular momentum S^2 as well as the spin projection along an arbitrary axis, S_z, which is built-in as a constraint to the electronic structure methods. In the Schrödinger framework, electron spin is a good quantum number, and the speed of light is treated as $c = \infty$.

The Dirac framework introduces relativistic effects through the Dirac-Coulomb-Breit Hamiltonian[77−79] that operates on a two- or four-component wave function. Relativistic effects are known to be extremely important for describing heavy-element chemistry and accurate prediction of spectral signatures; scalar relativistic effects cause significant contractions of the core electron shells, while vector-based relativistic effects (e.g., spin−orbit, spin−spin interactions) modulate the optical and magnetic properties of chemical complexes in response to external perturbations.

The quantum field framework is a next frontier in electronic structure theory with an even more detailed description of the interactions. In quantum field theory, photon-mediated electron-positron correlations are introduced, which can alter the energetic ordering of quantum states, such as the well-known Lamb shift in spectroscopy. For electronic systems, the time-dependent photons can be traced out via equal time integration, giving rise to an effective quantum field method. The energetic contribution from the quantum field is much smaller than the Dirac-Coulomb-Breit Hamiltonian but increases significantly toward heavy-element and highly charged states.

Classes of Many-Body Methodologies. Orbital-based models of electronic structure are the foundation of qualitative reasoning in chemistry and physics as well as the bedrock of the modern quantitative toolkit for electronic structure.

- The 1-body (mean-field) electronic structure models such as Kohn−Sham (KS) Density Functional Theory (DFT) and Hartree−Fock (HF) are the dominant models used for practical computations in chemistry and physics today. Although computationally efficient (with $O(N^3)$ computational cost in conventional form, with more sophisticated forms approaching linear scaling; N stands for the system size), the simplicity of 1-body models precludes their broad ability to describe complex electron correlations, such as those due to the coupling of spin degrees of freedom in open-shell systems and especially excited states, as well as nonlocal interactions. Nevertheless, the permanent workhorse status of such models demands continuing algorithmic developments of such methods to improve efficiency/portability, as well as new conceptual developments addressing known artifacts such as self-interaction errors.[80−82]

- The many-body models rely on the 1-body models (HF, DFT) as starting points to provide them with a reference

wave function, Green's function, nodal surfaces, etc. The rich landscape of many-body methods can be coarsely categorized into (a) wave function-based methods [configuration interaction (also known as exact diagonalization), density matrix renormalization group (DMRG) and other tensor network methods, variational Quantum Monte Carlo (QMC)], (b) wave operator-based methods (coupled-cluster (CC),[83,84] many-body perturbation theory (MBPT)), (c) Green's function-based methods (GW, Bethe-Salpeter Equation), and (d) diffusion and auxiliary-field QMC. These methods are fundamentally more expensive than the 1-body methods, either due to the prefactor (diffusion QMC) or complexity (high-order polynomial scaling, such as $O(N^7)$ for CCSD(T),[85] and up to $O(N^{n_e})$ for full-CI diagonalization; n_e designates the number of electrons). Although algorithmic developments and numerical approximations can reduce the cost/complexity of some many-body methods (even to linear[86]), the worst-case exponential scaling is unavoidable to approach exactness.

Although the reach of orbital-based descriptions of electronic structure is often augmented by coupling to the approximate models, such as classical atomistic force fields (QM/MM embedding) or orbital-free DFT methods (QM/QM embedding), addressing the scientific challenges (including those identified in Section II) will involve breakthrough developments of the orbital-based toolkit.

Specific Many-Body Classifications. *Choice of Fundamental Variable.*

- Wave function (ψ) methods offer the possibility to construct a hierarchical system of approximations for ground and excited states. Their main drawback is their high cost: intrinsic scaling with system size N is poor.
- Density-functional (ρ) methods are by far the most efficient because all of the information about the ground state is compressed into a simple scalar function, the density. The blessing of DFT is also its curse: functionals are simple but uncontrolled and, in several instances, not transferable across system sizes.
- Green's function (G) methods form the fundamental variable for many-body perturbation theory (MBPT). Although a Green's function can be constructed by itself, it can also incorporate the formulations based on ρ and ψ.

Single-Reference vs Multireference Formulations. It is useful to classify the various methods into single reference (SR) and multireference (MR).[87−91] SR approaches usually start from a single (often closed shell) Slater determinant, whereas MR methods allow the inclusion of several (many) configurations (determinants or configuration state functions (CSFs)) as a starting point. Many popular variational and perturbational methods are available, among which the Møller–Plesset perturbation theory and coupled cluster theory rank among the most popular SR ab initio methods. For the sake of convenience, DFT methods based on the Kohn–Sham theory may also be included as SR methods here.

These methods have demonstrated marvelous success in calculating ground state properties and electronic excitations in the Franck–Condon region, i.e., near the ground state minimum. In view of this success, why do we need MR methods? The answer is given by the increasing importance of treating more complex problems like the treatment of carbon-based polyradicaloid low-bandgap nanomaterials, transition

metal and lanthanide complexes, and bond-breaking processes. In all of these cases (and many more), strongly correlated, quasi-degenerate orbital schemes are encountered for which SR methods quickly reach their limits. When discussing MR methods, one often separates the treatment of the strongly coupled near-degenerate electronic configurations from the weakly coupled configurations. To handle the strongly coupled subspace, one can use a brute force CI treatment or systematically improvable types of ansatz, such as provided by the density matrix renormalization group[79,92−95] (DMRG) method and the graphically contracted functions (GCF) approach,[96−98] which can treat much larger CSF expansion sizes. On top of this, many methods have been introduced to handle the more weakly coupled configurations, including variational methods, such as the MR configuration interaction[99,100] (MRCI) method, MR perturbation theories (with CASPT2[101−104] and NEVPT2[105,106] as the prominent representatives), multiconfiguration pair-density functional theory[107,108] (MC-PDFT) (which combines a functional of the pair density with a multiconfiguration self-consistent field (MCSCF) wave function), and multi-Slater determinant auxiliary field quantum Monte Carlo methods.[109,110] Stochastic sampling techniques can also be used more broadly in the electronic structure problem, for example, to treat all correlations in the full-CI quantum Monte Carlo[111] (FCIQMC) method, or in hybrid approaches that combine stochastic and many-body methods.[112] Significant advancements have also been made in the extension of SR CC theories into the strongly correlated regime through the implementation of alternative design principles.[113−116]

Green's Function Formulation. The foundation of the Green's function approach is the Dyson equation, which encapsulates the properties of the many-body system through one-body operators. A conceptual merit of the one-particle many-body Green's function (MBGF) is its ability to directly calculate the key electronic properties of the ionization and attachment processes without resorting to separate calculations for different states. Typical one-particle MBGF approaches can usually be derived from the many-body perturbation expansions for the matrix of one-particle Green's function G and/or for the related self-energy Σ via the Dyson equation. For example, in the Hedin formulation,[117] the theory consists of a set of five closed equations, which can be expanded via perturbation theory as a set of Feynman diagrams in powers of the screened Coulomb interaction, W. Many-body effects are captured, formally exactly, in the self-energy Σ of the one-body Green's function. Much of the complexity is embedded in the frequency dependence of W and Σ. The *GW* approximation may be viewed as an extension of the Hartree–Fock approximation, with W replacing v in the exchange. Alternatively, Green's functions can be formulated using coupled-cluster theory (GFCC),[118−122] the algebraic diagrammatic construction (ADC) approach,[123−125] and perturbative many-body expansions of the self-energy (Σ).[126,127]

These approaches vary in the degree of complexity in the way the many-body effects are handled. When these effects become crucial, as often featured by satellite states in the ionization process out of the inner valence band, where poles will appear in the analytical structure of the self-energy, a proper description of the poles in the analytical structure of the self-energy is required.

Ground vs Excited State Methods. Currently, an overwhelming number of quantum chemical calculations are applied to the electronic ground state. The key aspect shared by all of the methods and required for predictive results is the ability to

capture complex electron correlation effects. In this family of many-body approaches, DFT is arguably the most popular method that is broadly applicable. However, although computationally more expensive, wave function based methods (for example, Møller−Plesset perturbation theory, coupled cluster, configuration interaction, DMRG, etc.) offer the possibility to construct systematically improvable classes of approximation, where the achievable accuracies are intimately tied to the available computational resources. Green's function methods provide a complementary way of addressing the ground-state electronic structure problem.[128−130]

Most of the above-mentioned methods are systematically extendable to compute excited states. The most commonly used approaches are linear-response (LR) formulations (for example, LR-TDDFT and LR-CC[131−135]), equation-of-motion coupled-cluster (EOMCC) methods,[136−140] spin-flip variants,[141] ADC formulations, multireference active space formulations, in either perturbative or iterative flavors, and many others. All these methods have been used extensively to compute a range of linear and nonlinear spectroscopies spanning broad energy ranges (IR, UV/vis, and X-ray) and nonadiabatic photodynamics simulations.

Over the last two decades, significant advances have been made in real-time (RT) approaches, which go beyond the perturbative regime (for example, RT-TDDFT, RT-CC, RT-Green's Function). Real-time approaches provide an unprecedented view of electron dynamics on the atto- and femto-second time scales, with vast potential to yield new insights into the complex electronic behavior of molecules and materials.[142]

Multicomponent Systems. Recent efforts in electronic structure theory have also focused on extensions to multicomponent systems, where additional components include spin degrees of freedom, a quantized electromagnetic field, and/or the nuclear wave function. For spin-driven electronic dynamics, such as intersystem crossing events, spin-couplings, and relativistic effects, variational treatments within the two- or four-component Dirac framework are needed.[77−79,143] The coupling of a molecule to a quantized electromagnetic field, real-time quantum electrodynamics (QED),[144] has led to studies of photon absorption and emission and simulations of cavity QED experiments. For many light-driven dynamics in chemical systems, quantum mechanical representations of proton dynamics have been demonstrated with the nuclear−electronic orbital (NEO) approach in the context of multicomponent RT-TDDFT and CC formulations for molecular systems.[145,146] By treating protons and electrons quantum mechanically at the same level of theory, NEO formulations automatically capture essential features such as vibrational zero-point energy, proton delocalization, vibrational anharmonicity, and non-Born−Oppenheimer effects. Although QED-enabled real-time methods have emerged as a useful tool to study novel photon-driven chemical processes, a complete theory requires a full first-principles QED electronic structure theory treatment with photon-mediated coupling between electrons and positrons and the treatment of retardation with the frequency-dependent Breit Hamiltonian. A key challenge in the development of first-principles methodologies for multicomponent systems is developing systematically improvable and accurate formulations on top of self-consistent field methods.[147−149]

Embedding Methods. One route to balancing computational cost and accuracy can be achieved by considering the relatively local nature of chemical interactions. Interesting chemical phenomena tend to be localized to the active areas of extended electronic systems. Following this observation, it is possible to split a larger system into an active subsystem, which can be tackled with an accurate level of theory that would be prohibitive for the full system, and the environment, which is treated at a cost-efficient level that provides broadly acceptable accuracy on larger systems. Over the years, a variety of schemes have been developed around the concept of embedding subsystems[150−155] at different electronic structure theory levels, including ONIOM,[156,157] DFT embedding,[151,158,159] partition DFT,[160,161] fragment methods,[162−165] potential-functional embedding,[166] embedded mean-field theory (EMFT),[167] Green's function embedding,[168,169] self-energy embedding,[170,171] EOMCC embedding,[172] density matrix embedding theory (DMET),[173−177] stochastic embedding DFT,[178] dynamical mean-field theory (DMFT),[179−181] and projector-based embedding.[182−186]

Molecular Dynamics. The time-scale problem in chemistry can often be solved with Born−Oppenheimer molecular dynamics (BOMD) simulations.[187,188] Typically, Newton's equations of motion for the (classical) atomic positions are integrated, step-by-step, to generate the molecular trajectories, where the interatomic forces are calculated on-the-fly from the ground state electronic structure for each new configuration. For light nuclei such as hydrogen, the case has been previously made[189] that for BOMD simulations it is more appropriate to consider their nuclear statistical nature via Feynman path integral simulations,[190] which are complex linear quantum superpositions of classical trajectories.[191] A major limitation of this approach is the large computational cost in the iterative optimization of the electronic ground state that is required prior to each force evaluation. Insufficient convergence may lead to nonphysical dynamics with nonconservative forces. The problem is particularly challenging in combination with low numerical precision or linear scaling methods. Time-reversible extrapolation methods have been developed to solve these problems. New formulations based on a backward error analysis or a shadow Hamiltonian approach in combination with Car−Parrinello-like extended Lagrangian techniques have also been introduced to reduce the computation cost and improve the accuracy and long-term stability.

To reach the time scale necessary for many problems, it is not possible to use direct molecular dynamics simulations. Instead, various accelerated molecular dynamics methods can be used to boost the effective time scale, often by multiple orders of magnitude. However, in general, these methods are applicable only to special rare event dynamics. For floppy dynamic systems that are common in chemistry and molecular biology, we can instead use accelerated sampling techniques or Monte Carlo methods, although the actual time scale is then lost.

Multiscale Methods. To computationally and efficiently model many important chemical and molecular processes on large length and time scales requires the seamless integration of degrees of freedom with different representations as well as external influences and environmental effects. Early efforts to couple quantum methods with classical and continuum methods have been demonstrated and reviewed[192,193] and will continue to be necessary even as variable-accuracy quantum methods begin to approach length scales that are currently associated with force-field models. As in the cases of quantum methods, it is anticipated that methodological domains will include the following: (1) regions where classical charge-transfer and polarizable response is needed,[194] (2) regions where only polarizability (electrical or magnetic) needs to be addressed, and
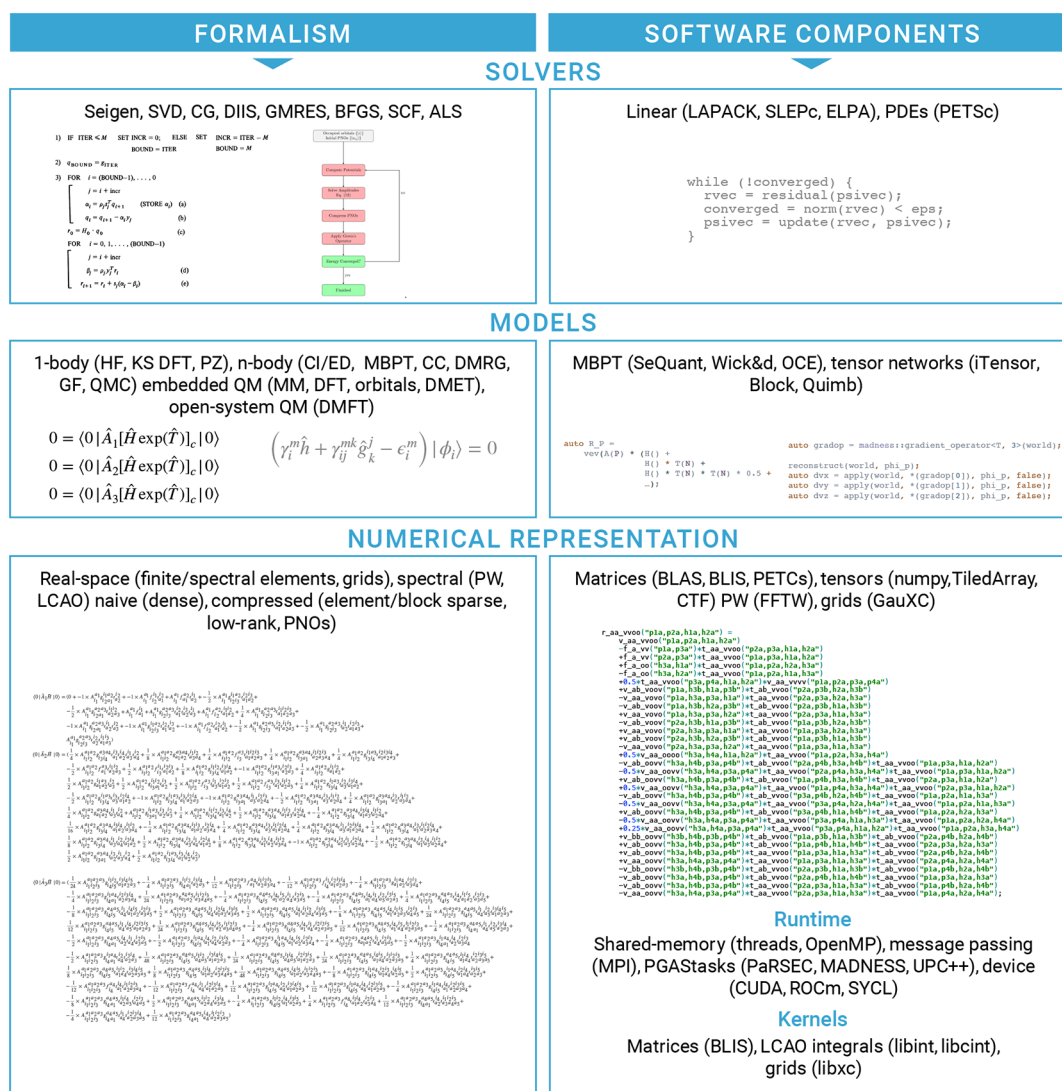
Figure 3. Schematic representation of the interdependencies between various theoretical formulations and software components needed for modern-era computational chemistry.

(3) other regions that simply required classical force fields or elastic continuum theories. Similar to the case of coupling quantum-mechanical methods, the requisite coarse-graining will require improved strategies for ensuring that the resulting interatomic forces associated with atoms near methodological boundaries are invariant to whether or not the forces are calculated with either the less- or more- sophisticated methods. Additionally, by coupling variable length-scale methodologies, dynamics simulations within these formulations will require variable time-scale simulations and the ability to realistically shift dynamical simulations when discontinuous stimuli from the external environment occur.

COMPUTATIONAL APPROACHES

Enabling predictive simulations on emerging architectures hinges upon effectively utilizing the computational resources to address the inherent numerical scaling of theoretical formulations, including efficient solvers, interoperable models, numerical representations, and novel programming models (Figure 3). These requirements need a stable and sustainable programming environment since solutions in terms of improvements of existing methods or the development of new methods

will be a continuous process that needs interaction between different programming paradigms and program developers.

Numerical Scaling. The computational cost of canonical (i.e., without additional numerical assumptions) quantum-mechanical electronic structure calculations scales, in general, with the cube, N^3, of the system size, N, or worse, which correlates with the expected accuracy of the predictions. For example, the ubiquitous CCSD(T) method scales as N^7 with the system size, which allows one to tackle systems composed of 10−100 light atoms.[195] In contrast, approaches such as DFT methodologies scale typically as N^3 and can handle much larger systems.

Despite this steep canonical scaling, it is often possible to reduce the scaling of electronic structure calculations due to additional physical features of the solution. Reduced scaling methods based on mean-field methods like HF and DFT typically rely on the locality of the orbitals (or Wannier functions, using Kohn's nearsightedness principle). The electronic locality occurs in nonmetallic systems or for materials at a high electronic temperature. Divide−and−conquer schemes or numerically thresholded sparse matrix algebra can then be used to take advantage of the electronic locality to achieve linear

scaling complexity. Some of the key problems with linear scaling electronic structure theory include the following: 1) a reduced numerical accuracy with errors that often are difficult to control; 2) a high computational prefactor, where the linear scaling advantage kicks in only for very large systems; and 3) additional overhead associated with parallelism.

In correlated wave function methods, reduced scaling is achieved by taking advantage of the rapid decay of many-electron correlations with distances in most chemical situations; truncation of these interactions can be achieved in several ways, including (1) truncation of incremental many-body expansions for the energy based on the partitioning of the system into fragments of the size larger than the quantum correlation length of the system[162−165] and (2) truncations of the operators and wave functions directly by representing them in a form that reveals their sparse structure. In molecular applications, this means using spatially localized basis sets (AOs, localized MOs, and finite/spectral elements), pair-natural orbitals (PNOs), or domain-local pair natural orbital (DLPNO) methods, to mention only a few strategies.[86,196−203] A significant effort has been devoted to extend the reduced-scaling problems beyond typical ground-state applications to excited-state or linear response methodologies formulated in the time or frequency domains.[204−206] The numerical error of many-body methods also suffers from slow asymptotic decay due to the singularity of the Coulomb electron−electron interaction and the resulting cusps in the electronic wave function. This translates into rapid asymptotic growth of the computational cost with the desired precision; most importantly, the use of small basis sets results in unacceptably significant errors. To address this problem, explicitly correlated formalisms, specifically in the form of the R12/F12 methods, build in the cusp-like structure into the wave function via terms dependent on the interelectronic distances. One of the pressing issues in computational chemistry will be integrating the F12/R12 methods with various reduced-scaling frameworks.[207−209]

Machine Learning. The ability to generate large amounts of high-quality data from *ab initio* theory is a prerequisite for data-driven machine-learning techniques. These new technologies are based on information theory and artificial intelligence and are rapidly evolving into a new field of research in computational chemistry. Such new machine learning methods often complement or can, in some instances, even replace many traditional approaches in computational chemistry.[210−216] Particularly fruitful areas of the application of machine learning in theoretical and computational chemistry include the design of interatomic potentials for molecular dynamics simulations, the development of new exchange-correlation functionals in DFT, and the prediction of properties of chemical systems. Despite the success of machine learning in computational chemistry, there are good reasons for some caution. Often the underlying physical mechanism for a predicted property, e.g., that generated by a deep neural network, is missing. The accumulation of knowledge from machine learning will therefore be limited, and errors will be hard to detect. Often machine learning models act like an interpolation between already explored data points and cannot be used to discover new unexpected phenomena. Interatomic forces are governed by long-range electrostatic interactions between the positive nuclear charges and the negative charges of the relaxed electron density,[217] which cannot be captured by commonly used machine learned force fields, in which the energy is cast as a sum of atomic terms. Incorporating electrostatic interactions into machine-learned force field

models of this type will then be like modified versions of traditional polarizable flexible charge models, where the long-range charge interactions and relaxations are included separately. Nevertheless, machine learning in computational chemistry is undoubtedly here to stay.

PROGRAMMING MODELS AND SOFTWARE INTEGRATION

Software Development Challenges for Modern Computing. While certain computational chemistry workloads can be supported by consumer grade computing platforms, the true power of quantum chemistry to affect change and influence scientific inquiry relies on its effective utilization of high-performance computational (HPC) resources, supported by funding agencies in the US and worldwide. A driving force for the success of computational chemistry has been its ability to adapt to an ever evolving computational landscape and to quickly adopt emerging performance-driven technologies. Historically, this adaptation has focused on central processing unit (CPU) architectures, which have been the dominant design feature in both consumer and HPC platforms across virtually all industries for the last half century. While the pervasiveness of CPUs has supported progress in this domain, it has also contributed to a degree of complacency among developers regarding the importance of performance-portability, extensibility, and other hardware-aware considerations in their software efforts. In recent years, it has been recognized[218−220] that addressing these challenges represents a sizable hurdle for software sustainability efforts in computational chemistry, and the identification of long-term solutions is critical for these efforts in the years to come.

With the inevitable demise of Moore's law, modern HPC has adopted the use of specialized hardware for performance critical computation over the general purpose, power-intensive capabilities of CPU processors. While the homogeneity of CPU design has historically allowed for a certain level of hardware-software codesign in computational chemistry, particularly in its use of numerical linear algebra, it is clear that the dominance of graphics, artificial intelligence (AI), and machine learning (ML) will be the primary driver for specialized hardware innovation in the years to come. This paradigm shift is best represented by the introduction of accelerators, such as graphics processing units (GPU) and, more recently, AI-driven hardware such as tensor cores and tensor processing units (TPU), into the HPC ecosystem. In addition to a need for the development of novel programming models, compiler technologies, and optimized libraries to target these platforms, the move to accelerators often requires the re-evaluation of algorithmic design due to fundamental differences in execution strategies being appropriate only for particular classes of workloads (e.g., vectorized, low precision, and high arithmetic-intensity). AI-hardware's low mixed-precision floating-point operations, in particular, add new challenges to the numerical accuracy, algorithm stability, and convergence estimates for quantum chemical methodologies.[221−224]

While there has been an enormous effort afforded to the incorporation of modern HPC platforms into the scientific computing and computational chemistry ecosystems,[220−227] these efforts have been fraught with challenges and cannot yet be considered as mature as their legacy counterparts targeting CPU architectures. Outstanding challenges and opportunities in these areas include how best to leverage low-precision arithmetic for computational chemistry applications and how to develop new

or map existing algorithms onto particular compute patterns (such as tensor contractions, convolutions, etc.) — especially for kernels which are traditionally not linear algebra based, and how to rebalance existing codes by navigating trade-offs between sparse and dense linear algebra. For computational chemistry to remain viable, it is critical for its associated software and method development efforts to continue to evolve with modern HPC and to adopt defensive, flexible programming philosophies to better prepare them for the current and future architectures. In this section, we review a number of topics salient to low- and high-level abstractions that are required to address these challenges in modern computational chemistry software.

Low-Level Abstractions. It has long been recognized in computational chemistry that compiler technology alone is often insufficient to achieve the peak performance of critical kernels on modern compute systems. As such, details regarding hardware design must be considered when attempting to develop performant software. This statement is equally true for software targeting either CPU- or accelerator-based architectures. However, the vast majority of the workflow design of chemistry software is not performance critical, and there typically exist only a handful of algorithmic kernels that need to be optimized for target architectures. Due to the size and complexity of typical chemistry software packages, to completely refactor code bases for each architecture of interest is impractical, and sustainable software efforts should strive to lessen developer effort to achieve performance portability to the largest extent possible. In this Section, we examine strategies to encapsulate and abstract low-level hardware-specific optimization for the development of sustainable computational chemistry software.

The development of performant computational chemistry software is particularly challenging in comparison to other scientific disciplines due to the relatively large number of performance critical kernels that comprise typical algorithmic workflows. A certain number of these kernels are generic in the sense that they are common to other areas of scientific computing, such as matrix and tensor algebra (multiplication, decompositions, etc.) and Fourier transforms to name a few. For these kernels, chemistry software can often rely on community software in the form of libraries (e.g., BLAS,[228] LAPACK,[229] ScaLAPACK,[230] and ELPA[231−233]) to act as sufficiently general low-level abstractions. Recent years have also seen the assembly of software collections for targeting HPC architectures, such as the Extreme-scale Scientific Software Stack (E4S)[234,235] and the Extreme-scale Scientific Software Development Kit (xSDK).[236] Typically, these libraries are released in a manner that targets a specific architecture of interest (CPU/accelerator, shared/distributed memory) and made accessible by standardized Application Programming Interfaces (APIs). Such specialization has led to the development of chemistry-community driven abstraction layers such as the Electronic Structure Infrastructure (ELSI),[237] the Basic Matrix Library for quantum chemistry (BML),[238] and the CECAM Electronic Structure Library (ESL).[239] For particular chemistry applications, many performance critical, low-level abstractions are able to be satisfied by these libraries alone, but for other applications, there exist a number of domain-specific kernels that cannot be satisfied by generic community software.

Many computational chemistry methods rely on domain-specific kernels to perform a number of performance critical tasks. These kernels are most common in applications working with basis representations (e.g., Gaussian and Slater type orbitals, numeric atomic orbitals, wavelets, etc.) of integrodifferential operators encountered in physical Hamiltonians. For example, in atomic-orbital-based electronic structure theory, the evaluation and manipulation of the electron repulsion integral (ERI) tensor, or its various decompositions, are highly sensitive to underlying hardware details and constitute a fair majority of the computational work encountered in these applications. While it would be possible to leverage community software in the implementation of many of these kernels, it is often the case that the development of highly specialized kernels and algorithms leads to significant performance improvements on modern hardware. For example, the development of highly optimized recursions[240−250] and chemistry-specific quadrature schemes[251−256] for operator integral evaluation and contraction[257,258] are known to outperform generic numerical integral machinery on both CPU and accelerator architectures. In addition, domain-driven tensor frameworks[259−262] typically outperform generic tensor frameworks for computational chemistry workloads. As such, many chemistry-community driven libraries have been developed for low-level operations such as analytical (e.g., libint,[263] libcint,[264] and simint[265]) and numerical (e.g., GauXC,[266−268] libGridXC[239]) operator integrals, functional evaluation (libxc,[269] XCfun,[270] ExchCXX[266]), and tensor algebra (e.g., TiledArray,[259,260] TAMM,[261] and the Cyclops Tensor Framework (CTF)[262]) to name a few. The development and optimization of these libraries is challenging; thus, it is of critical interest to the development of sustainable computational chemistry software that these kernels be implemented in a performance portable manner.

As with the implementation of many performance critical kernels, optimization of chemistry-specific kernels is highly hardware specific. For CPU-based architectures, leveraging single-instruction multiple-data (SIMD) and fused multiply add (FMA) capabilities is paramount to achieving peak performance. While typically generated by optimizing compilers, explicit SIMD optimization can be performed in a portable manner through the use of low-level abstractions such as vector intrinsics, as has been explored in several integral libraries.[263−265] The situation is complicated on accelerator based architectures which often rely on vendor-specific programming models (e.g., CUDA[271] for NVIDIA GPUs and HIP for AMD GPUs) to directly manipulate hardware capabilities. To avoid refactoring code bases for each accelerator, several attempts have been made by the compiler community to develop unified programming models over multiple accelerator backends (e.g., OpenMP,[272] OpenACC,[273] OpenCL,[274] SYCL). In addition, software-driven performance portability layers (e.g., Kokkos[275] and RAJA[276]) have also been explored. While such efforts have proven to be fruitful for some cases,[277] such approaches have limited features and performance portability to be universally applicable.

Despite significant efforts in developing performance portable software within particular classes of hardware, it is often the case that, in some sense, the execution strategies for chemistry workflows must be optimized themselves, apart from low-level hardware-specific implementation details. For example, CPU-based ERI libraries typically expose integral evaluation at the granularity of individual shell-quartets, while efficient accelerator algorithms require a much coarser granularity to achieve good performance. Design choices such as these lead to unsustainable development practices where large portions of code bases must be refactored for different classes of hardware regardless of the implementation details of performance critical kernels. The need to avoid (or at least minimize) such

Journal of Chemical Theory and Computation
pubs.acs.org/JCTC
Article
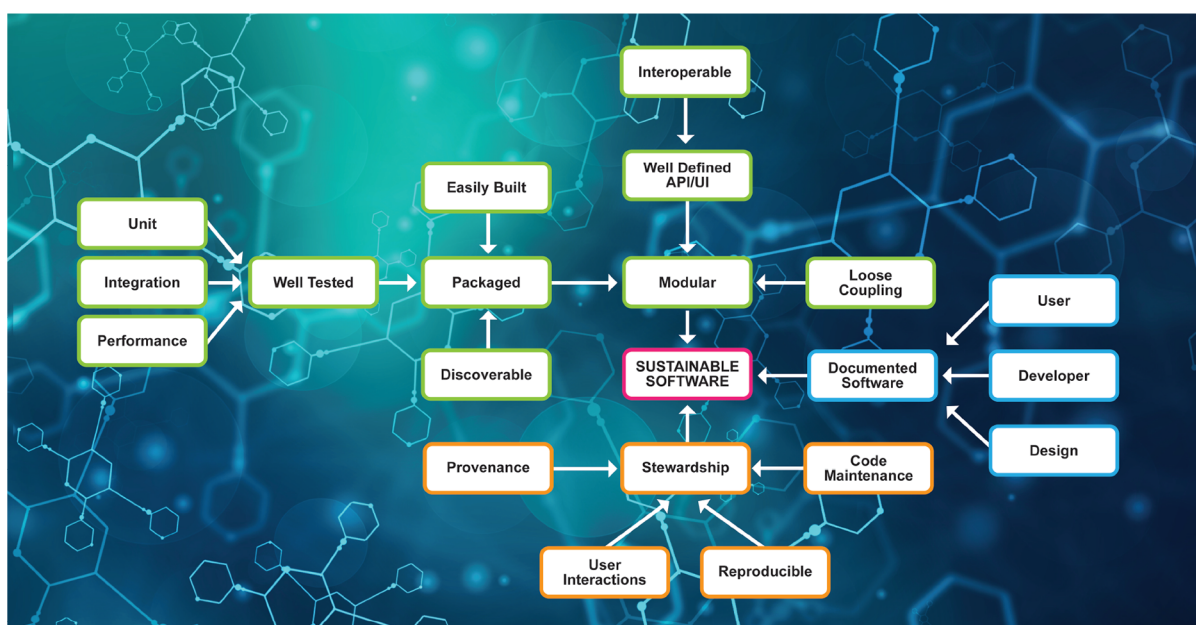
Figure 4. A "word web" highlighting the many considerations that go into designing, writing, and stewarding sustainable software.

refactoring creates a somewhat niche role that can be occupied by a "middle-layer". This middle-layer is designed to decouple the low-level and high-level layers from one another. As the quantum chemistry software stack complexity continues to increase, middle-layers like PluginPlay[278] — a framework for developing modular scientific software, where the developer chooses the module granularity — will become increasingly important to help bridge the gap between low-level and high-level abstractions, the distinction between which is becoming increasingly blurry.

High-Level Abstractions. Reflecting the hierarchical structure of the formal abstractions of our chemistry applications, it is natural to hide the low-level abstractions, which deal with the hardware at its lowest levels of granularity, under a layer of domain-specific high-level abstractions. Layered design can help improve sustainability by insulating implementations of models and algorithms from the disruptive changes in low-level implementation details (kernels, programming models, architecture). When properly executed, it also naturally makes composition easier by insulating the users of high-level abstractions from low-level implementation details.

The high-level abstractions can take many forms depending on the particulars of the methods, algorithms, and even details of numerical representations. Perhaps the most notable example from electronic structure is tensor algebra libraries/frameworks, which support or encompass the implementation of many-body electronic structure methods in algebraic (second-quantized) representation. Tensor Contraction Engine (TCE)[279,280] supported the composition of complex many-body methods from a high-level operator specification in a domain-specific language. The implementation of the resulting tensor algebra on a distributed partitioned global address space (PGAS) runtime Global Arrays (GA)[281] was also compiler-generated. In addition to many improvements and generalizations (SMITH,[282] SMITH3[283]) of TCE, a major refinement of its ideas is becoming possible by decoupling of the high-level operator algebra and tensor algebra layers, the latter including components for optimization (e.g., factorization) of the algebra of symmetric tensors[284] and its implementation using generic

tensor frameworks (including distributed and heterogeneous).[285−289] Similar high-level abstractions can be found in other areas of electronic structure, e.g., tensor network computation[290] and DFT.[291]

Yet higher up the abstraction ladder, the focus switches from the representation of electronic states to solvers for the associated quantum and classical equations-of-motion (e.g., nonequilibrium time-dependent electronic structure, Born−Oppenheimer and nonadiabatic dynamics, etc.). For performing such tasks there exist multiple "frameworks" (ASE,[292] NewtonX,[293] SHARC,[294] QCEngine,[295] among others[296,297]) for abstracting (interfacing to) standalone packages; such frameworks can be viewed as domain-specific specializations of generic workflow components. The highest level of abstraction is also where, typically, the computational chemistry and machine learning models meet.[298]

In practice, both high- and low-level software abstractions are critical for the development of performant computational chemistry software, particularly in the modern computing era. However, the development and maintenance of such abstractions are time- and resource-consuming, posing a considerable hurdle for *sustainable* software design. In the following Section, we examine development practices for sustainability in computational chemistry software.

■ TRANSITIONING TO SUSTAINABLE COMPUTATIONAL CHEMISTRY SOFTWARE

Most computational chemistry methods have very high space and time complexities. In order for these methods to be applicable to realistic rather than toy systems, developers need to spend a considerable amount of time optimizing the software. Historically, these optimizations have been done on a per code basis leading to a somewhat substantial amount of "reinventing the wheel"; e.g., nearly every electronic structure package contains an implementation of the self-consistent field (SCF) method, Møller−Plesset perturbation theory, and coupled cluster theory, which have been hand-tuned by the developers of the package. Even if a developer wanted to repurpose an algorithm from another package, this is often very difficult

because most algorithms have been developed with a "just get something working" mindset. Broadly speaking, this often means that the developers (i) did not write documentation (especially developer documentation), (ii) spent little to no time on design, (iii) accepted a very tight coupling with the rest of the package, and (iv) skimped on test coverage. The fact that the resulting software tends to contain a large amount of technical debt and anecdotal evidence suggests that unwillingness to address this technical debt is often the reason why developers choose to reimplement algorithms (often perpetuating the cycle). Here, we argue that with the rate at which computational chemistry, software engineering, and computer hardware are currently advancing, the field of computational chemistry can no longer afford to accept this mindset. If we want to stay at the forefront of scientific advancement, then we need to work together to develop and maintain sustainable and interoperable software. Figure 4 is meant to accompany the present discussion by providing a succinct summary of the main topics discussed in this section.

Actually defining what it means for computational chemistry software to be sustainable is a tricky task, especially since advancements in other fields "move the goal posts". Presently, we adopt the simple and practical definition that sustainable software outlives the original use cases and can survive change. This means that the software can readily be adapted to new use cases, leveraged by different teams, and ported to new hardware. The software may be either open or closed source, but its public-facing APIs and user interfaces must be stable and well documented. The software and the data it produces should be standardized and curated to ensure both remain accessible at later times. Throughout the process, it is imperative that the high space and time complexity of computational chemistry be kept in mind; namely, all of the above must be done in a performance-aware manner.

In practical terms, this suggests that computational chemistry software must be designed in a modular and encapsulated manner. The modules should be as decoupled as possible in order to facilitate refactoring, integration, and rapid-prototyping. The corollary to this point is that modularity should be used to ensure a separation-of-concerns. While many quantum-chemistry packages are modular at the level of computing an energy or an energy gradient, we need to go far beyond that, including, but not limited to nonadiabatic couplings, building integrals, forming the Fock matrix, scanning potential energy surfaces, and numeric solvers. Smaller, more fine-grained modules are easier to maintain, tune, and properly credit than monolithic modules. It is also much easier to ensure the correctness and reproducibility for fine-grained modules.

From the perspective of sustainability, the computer language a module is written in tends to be less important than the languages for which it provides APIs for. For example, C-bindings exist for many Fortran libraries, and an increasing number of C/C++ libraries also provide Python bindings. Generally speaking, computational science is moving away from Fortran. While software written in Fortran is likely to persist for some time, it is our present recommendation that developers prioritize providing C/C++ and Python APIs regardless of the language in which the module is written. C/C++ retains critical roles in the software implementation ecosystem as the most widely used languages for low-level implementation. However, we note that in scientific computing and machine learning Python has emerged as a glue language capable of calling disparate pieces of software, in a cohesive manner, even if the

software is written in different languages. Combined with the fact that many languages have the ability to interface with Python, this suggests that Python can conceivably serve as a common API. The performance limitations of Python remain a relevant consideration, even in its use as a glue language. However, this may be ameliorated by the growing availability of Python just-in-time (JIT) compilers, such as PyTorch,[299] JAX,[300] etc., which utilize the Python syntax, but do not actually execute code via the Python interpreter. Increasing examples of Python based quantum chemistry frameworks and APIs, such as PySCF,[301,302] NWChemEx,[303] Psi4,[304,305] Dalton,[306] Gator[307] etc., showcase the potential of this approach.

Beyond modularity, sustainability also suggests that the software be readily extensible and customizable, which in turn requires a flexible and general infrastructure. However, this can be challenging, particularly for computational chemistry software, which often has predefined functionalities and highly optimized, hard-to-modify implementations. Introducing de-coupled, fine-grained modules with proper interfaces for extensions can partially address this issue. A more compre-hensive solution is to develop modular software that is amenable to meta-programming (writing code that uses other code as input). Meta-programming allows other developers to pro-gramatically, and noninvasively, extend and customize existing code. A particularly relevant example is given by recent autodifferentiation efforts[308-316] which, as the name suggests, are able to automatically compute analytic derivatives of a numerical routine. Meta-programming offers a major advantage to users outside of the chemical sciences (e.g., workers in machine learning and quantum information theory) as they can not only perform chemistry simulations but also extend and incorporate parts of the software into their own projects. By increasing the number of people from diverse communities using or contributing to the software, the likelihood of the software becoming obsolete is reduced. Over time, an ecosystem may form around the platform, where an interconnected community of developers and users from various fields can provide support, feedback, and resources that help the software continue to improve and evolve. This may also help to address some of the challenges associated with retaining developers and securing funding, as discussed below.

Another major piece of sustainability is stewardship. Here, stewardship entails tasks beyond initial method development meant to ensure that the software remains viable over the long-term. To have staying power, software needs to be stable, accessible, reproducible, and reliable, which comes from good design, extensive testing, outstanding documentation, robust deployment strategies, and community engagement. In practice, designing for the dynamic nature of science is hard, but semantic versioning combined with good version control practices can help minimize the damage when designs need to change. Extensive testing requires more than unit testing and includes the following: integration, performance, deployment, and acceptance testing. Testing should be combined with code coverage to ensure that the code is indeed exhaustively tested. Developer and user documentation, tutorials, and resources are extremely important to ensure the software can outlive any particular developer's involvement. Without such resources, using or extending the software is time-consuming and difficult. The software should be deployed in a manner that facilitates an easy setup and reliable access. A somewhat underappreciated point pertaining to robust deployment is that it becomes essential to treat the software's infrastructure, such as the build

system, testing harnesses, code generators, and continuous integration workflows, as code too. Engaging the user community ensures that the software gets used, improved, and further vetted. It is worth noting that part of supporting the user community is providing computer science and engineering resources, since many members of the community are scientists by trade.

While one can envision sustainable software which lacks interoperability, the reality is that without interoperability, it is extremely time-consuming to develop and steward software. For our purposes, interoperability means that two pieces of software "just work" together. For components to exhibit true interoperability, it must be possible to swap the components with no work other than telling the framework to use the new component. This means no glue code, data conversions, language barriers, or additional configuration. In practice, this is more of an ideal to strive for than is a characteristic of the actual software. Nonetheless, it is something we as a community should strive for and work toward if we want to avoid reinventing the wheel. It is important to realize that interoperability must be a community effort since we must agree on common standards for data and APIs. Inevitably, because of the large amount of technical debt in most packages, there will be a large upfront cost to move to interoperability, so the conversion is best done piece wise.

The last piece of sustainability is growing and retaining the overall computational chemistry community. Without users or developers working together, the field will slowly die. Anecdotal evidence from the Exascale Computing Project shows that developers of scientific software are in high demand by industry, particularly by tech companies. Hence, sustainability requires retaining these developers. We note that we are specifically focused on retaining and expanding the number of developers in permanent positions (such as in professor and scientist roles); positions such as student and post doctoral researchers are by design supposed to turn over. Historically, retention has been difficult for many reasons, chief among them being credit, salary and benefits, and difficulty securing external funding. Credit is an issue since many universities or laboratories do not appreciate software development as much as they appreciate publishing, patents, and conference presentations. Industry, on the other hand, has a strong appreciation for how important software development is, which is reflected in the salaries and benefits that they are willing to offer to retain their developers. While it is unlikely that universities or government laboratories will ever match industry salaries, addressing the credit issue could go a long way toward retention as would lowering the barrier to securing external funding for software development activities.

Moving forward, it is the recommendation of the authors that a higher emphasis be placed on ensuring that computational chemistry software is sustainable and properly stewarded. Admittedly, not every research avenue pans out, so there still needs to be an initial proof-of-concept/"just get something working" phase. What we are arguing instead is that once an idea has been vetted, developers must disseminate the feature in a sustainable manner. At present, this can admittedly be a tall order; therefore, we also recommend that the community pursue true interoperability and build software infrastructure, such as reusable frameworks, which can leverage the interoperability to facilitate sustainability. We also note there are several existing scientific research software communities from which we can take cues, such as Research Software Alliance,[317] Better Scientific Software,[318] US Research Software

Sustainability Institute,[319] and the Molecular Sciences Software Institute.[320] Notably, these organizations have already considered many of the issues mentioned above in depth and provided suggestions. The last recommendation is to ensure that our stakeholders (funding agencies, journals, universities, etc.) are also prioritizing sustainable computational chemistry.

Finally, we note that there are many open questions and challenges related to sustainability. Perhaps the foremost of these concerns is the availability of sustainable funding. For larger projects, software stewardship is a full time job, and finding the funds to support stewardship positions is challenging. One potential solution is to monetize the software. This is not without its own challenges though. For example, interfacing and interacting with both open-source and closed-source software can be tricky, not just scientifically but also legally. Another potential hurdle faced by closed-source software is ensuring that the scientific results are reproducible. Admittedly, this is also a problem for open-source codes, too; however, users of open-source codes can always fall back to reading the source as a last resort. We refer the reader elsewhere for a deeper discussion of choosing the commercial route for the sustainability of a specific software package.[33]

Sustainability also faces challenges related to education. Since most scientific research software deals with highly technical subject matter, finding someone who understands the science and is skilled in software engineering is difficult. Furthermore, many existing software engineering solutions can not be immediately ported to scientific software simply because the software was never designed to accommodate them. The net result is that we need to develop a new approach to education that teaches fundamental science from the perspective of software engineering. That said, we do not purport to have the answers to these challenges, but they are challenges that the community will need to contend with if we want to transition to sustainable software in computational chemistry.

■ CONCLUSIONS AND FUTURE OUTLOOK

Computational chemistry forms a complicated ecosystem that embraces formulations driven by various design principles and scenarios for synergies and interoperability between them, leading to unique and predictive computational frameworks working across multiple scales and complexities. Therefore, a sustainable chemistry software development effort requires the interplay among electronic structure theory, applied mathematics, and computer science.

With the advent of exascale computing and new computational paradigms associated with rapid advances in cloud computing, edge computing, and broad utilization of machine learning models, computational chemistry is facing deep transformations and challenges, leading to a new level of computational abilities needed for understanding critical societal challenges. New computational tools will allow for modeling processes and properties for system sizes and accuracies, which were impossible to attain in the last decades. For example, exascale computing offers a unique chance to integrate various representations of quantum mechanics based on the wave functions, density, and Green's function approaches that capture complex correlation effects across spatial and temporal scales for the ground- and excited-states of complex systems.

For these challenges to be addressed, the scientific community needs to work synergistically with public and private funders to address problems related to sustainable computational chem-

istry development based on the integration of broad classes of parallel computing tools, algorithms, and programming models. As essential elements underpinning this effort, one should mention the need for integrating hardware and application kernels for various types of formalisms and using modular low-level abstractions encapsulating basic information to ensure the high efficiency of scientific software. High-level abstraction in the software's design, maintenance, and interoperability is needed to compose many-body methods through the use of domain-specific languages and flexible tools for distributed computing and runtimes. These factors play a critical role in making software readily extensible and customizable as well as user-friendly for complex chemical workflows.

The lasting effects of adapting to the unprecedented computational capabilities can only be achieved if a coordinated effort is implemented to retain a qualified workforce and provide communication/collaboration conduits for multidisciplinary teams to provide efficient frameworks capable not only of supporting but also driving new scientific efforts.

■ AUTHOR INFORMATION

Corresponding Author

Karol Kowalski — *Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States;* ⓘ orcid.org/0000-0001-6357-785X; Email: karol.kowalski@pnnl.gov

Authors

Rosa Di Felice — *Departments of Physics and Astronomy and Quantitative and Computational Biology, University of Southern California, Los Angeles, California 90089, United States; CNR-NANO Modena, Modena 41125,, Italy;* ⓘ orcid.org/0000-0002-7772-6550

Maricris L. Mayes — *Department of Chemistry and Biochemistry, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts 02747, United States*

Ryan M. Richard — *Ames Laboratory, Ames, Iowa 50011, United States*

David B. Williams-Young — *Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0003-2735-3706

Garnet Kin-Lic Chan — *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States*

Wibe A. de Jong — *Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0002-7114-8315

Niranjan Govind — *Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States;* ⓘ orcid.org/0000-0003-3625-366X

Martin Head-Gordon — *Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0002-4309-6669

Matthew R. Hermes — *Department of Chemistry, Chicago Center for Theoretical Chemistry, University of Chicago, Chicago, Illinois 60637, United States;* ⓘ orcid.org/0000-0001-7807-2950

Xiaosong Li — *Department of Chemistry, University of Washington, Seattle, Washington 98195, United States;* ⓘ orcid.org/0000-0001-7341-6240

Hans Lischka — *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409, United States;* ⓘ orcid.org/0000-0002-5656-3975

Karl T. Mueller — *Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington 99354, United States;* ⓘ orcid.org/0000-0001-9609-9516

Erdal Mutlu — *Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

Anders M. N. Niklasson — *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* ⓘ orcid.org/0000-0003-1856-4982

Mark R. Pederson — *Department of Physics, The University of Texas at El Paso, El Paso, Texas 79968, United States*

Bo Peng — *Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States;* ⓘ orcid.org/0000-0002-4226-7294

Ron Shepard — *Chemical Sciences and Engineering Division, Argonne National Laboratory, Lemont, Illinois 60439, United States;* ⓘ orcid.org/0000-0002-0272-0824

Edward F. Valeev — *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* ⓘ orcid.org/0000-0001-9923-6256

Mark van Schilfgaarde — *National Renewable Energy Laboratory, Golden, Colorado 80401, United States*

Bess Vlaisavljevich — *Department of Chemistry, University of South Dakota, Vermillion, South Dakota 57069, United States;* ⓘ orcid.org/0000-0001-6065-0732

Theresa L. Windus — *Department of Chemistry, Iowa State University and Ames Laboratory, Ames, Iowa 50011, United States;* ⓘ orcid.org/0000-0001-6065-3167

Sotiris S. Xantheas — *Advanced Computing, Mathematics and Data Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States; Department of Chemistry, University of Washington, Seattle, Washington 98195, United States;* ⓘ orcid.org/0000-0002-6303-1037

Xing Zhang — *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States*

Paul M. Zimmerman — *Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States;* ⓘ orcid.org/0000-0002-7444-1314

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.3c00419

Notes

The authors declare no competing financial interest.
E-mail: difelice@usc.edu; maricris.mayes@umassd.edu; ryanmrichard1@gmail.com; dbwy@lbl.gov; gkc1000@gmail.com; wadejong@lbl.gov; niri.govind@pnnl.gov; mhg@cchem.berkeley.edu; mrhermes@uchicago.edu; xsli@uw.edu; hans.lischka@ttu.edu; Karl.Mueller@pnnl.gov; erdal.mutlu@pnnl.gov; amn@lanl.gov; mrpederson@utep.edu; peng398@pnnl.gov; shepard@anl.gov; evaleev@vt.edu; Mark.vanSchilfgaarde@nrel.gov; Bess.Vlaisavljevich@usd.edu; twindus@iastate.edu; sotiris.xantheas@pnnl.gov; xzhang8@caltech.edu; paulzim@umich.edu.

tion" meeting held in November 2022 in Seattle, Washington, USA. The authors are indebted to the agencies and programs responsible for funding their individual research efforts, without which this article would not have been possible. The authors hope that this paper provides a viewpoint, that can contribute to the ongoing and pressing discussions on the future of computational chemistry software and its integration with emerging technologies to enable new scientific advances. The Computational and Theoretical Chemistry Institute (CTCI) at PNNL is also gratefully acknowledged.

■ REFERENCES

(1) Romero, A. H.; et al. ABINIT: Overview and focus on selected capabilities. *J. Chem. Phys.* **2020**, *152*, 124102.

(2) Perera, A.; Bartlett, R. J.; Sanders, B. A.; Lotrich, V. F.; Byrd, J. N. Advanced concepts in electronic structure (ACES) software programs. *J. Chem. Phys.* **2020**, *152*, No. 184105.

(3) Shiozaki, T. BAGEL: brilliantly advanced general electronic-structure library. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2018**, *8*, No. e1331.

(4) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I.; Refson, K.; Payne, M. C. First principles methods using CASTEP. *Zeitschrift für kristallographie-crystalline materials* **2005**, *220*, 567−570.

(5) Kühne, T. D.; et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, No. 194103.

(6) Lee, J.; Hitzenberger, M.; Rieger, M.; Kern, N. R.; Zacharias, M.; Im, W. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **2020**, *153*, No. 035103.

(7) Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-cluster techniques for computational chemistry: The CFOUR program package. *J. Chem. Phys.* **2020**, *152*, No. 214108.

(8) Lischka, H.; et al. The generality of the GUGA MRCI approach in COLUMBUS for treating complex quantum chemistry. *J. Chem. Phys.* **2020**, *152*, No. 134110.

(9) Hourahine, B.; et al. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, No. 124101.

(10) Saue, T.; et al. The DIRAC code for relativistic molecular calculations. *J. Chem. Phys.* **2020**, *152*, No. 204104.

(11) Folkestad, S. D.; et al. eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, No. 184103.

(12) Yang, Z.-h.; Pederson, M. R.; Perdew, J. P. Full self-consistency in the Fermi-orbital self-interaction correction. *Phys. Rev. A* **2017**, *95*, No. 052505.

(13) Barca, G. M. J.; et al. Recent developments in the general atomic and molecular electronic structure system. *J. Chem. Phys.* **2020**, *152*, No. 154102.

(14) Frisch, M. e.; Trucks, G.; Schlegel, H. B.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Petersson, G.; Nakatsuji, H. et al. *Gaussian 16*; 2016.

(15) Aquilante, F.; et al. Modern quantum chemistry with [Open] Molcas. *J. Chem. Phys.* **2020**, *152*, No. 214117.

(16) Werner, H.-J.; et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, No. 144107.

(17) Peng, C.; Lewis, C. A.; Wang, X.; Clement, M. C.; Pierce, K.; Rishi, V.; Pavošević, F.; Slattery, S.; Zhang, J.; Teke, N.; Kumar, A.; Masteran, C.; Asadchev, A.; Calvin, J. A.; Valeev, E. F. Massively Parallel Quantum Chemistry: A high-performance research platform for electronic structure. *J. Chem. Phys.* **2020**, *153*, No. 044120.

(18) Kállay, M.; et al. The MRCC program system: Accurate quantum chemistry from water to proteins. *J. Chem. Phys.* **2020**, *152*, No. 074107.

(19) Guther, K.; et al. NECI: N-Electron Configuration Interaction with an emphasis on state-of-the-art stochastic methods. *J. Chem. Phys.* **2020**, *153*, No. 034107.

(20) Aprá, E.; et al. NWChem: Past, present, and future. *J. Chem. Phys.* **2020**, *152*, No. 184102.

(21) Tancogne-Dejean, N.; et al. Octopus, a computational framework for exploring light-driven phenomena and quantum dynamics in extended and finite systems. *J. Chem. Phys.* **2020**, *152*, No. 124119.

(22) Prentice, J. C. A.; et al. The ONETEP linear-scaling density functional theory program. *J. Chem. Phys.* **2020**, *152*, No. 174111.

(23) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, No. 224108.

(24) Smith, D. G. A.; et al. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, No. 184108.

(25) Sun, Q.; et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, No. 024109.

(26) Epifanovsky, E.; et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, No. 084801.

(27) Kent, P. R. C.; et al. QMCPACK: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion quantum Monte Carlo. *J. Chem. Phys.* **2020**, *152*, No. 174105.

(28) Giannozzi, P.; Baseggio, O.; Bonfá, P.; Brunato, D.; Car, R.; Carnimeo, I.; Cavazzoni, C.; de Gironcoli, S.; Delugas, P.; Ferrari Ruffino, F.; Ferretti, A.; Marzari, N.; Timrov, I.; Urru, A.; Baroni, S. Quantum ESPRESSO toward the exascale. *J. Chem. Phys.* **2020**, *152*, No. 154105.

(29) García, A.; et al. Siesta: Recent developments and applications. *J. Chem. Phys.* **2020**, *152*, No. 204108.

(30) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J.; James, W.; Song, C.; Titov, A. V.; Ufimtsev, I. S.; Martínez, T. J. TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. *J. Chem. Phys.* **2020**, *152*, No. 224110.

(31) Balasubramani, S. G.; et al. TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, No. 184107.

(32) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169−11186.

(33) Krylov, A.; Herbert, J.; Furche, F.; Head-Gordon, M.; Knowles, P.; Lindh, R.; Manby, F.; Pulay, P.; Skylaris, C.; Werner, H.-J. What Is the Price of Open-Source Software? *J. Phys. Chem. Lett.* **2015**, *6*, 2751−2754.

(34) Krylov, A.; Windus, T. L.; Barnes, T.; Marin-Rimoldi, E.; Nash, J. A.; Pritchard, B.; Smith, D. G. A.; Altarawy, D.; Saxe, P.; Clementi, C.; Crawford, T. D.; Harrison, R. J.; Jha, S.; Pande, V. S.; Head-Gordon, T. Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science. *J. Chem. Phys.* **2018**, *149*, No. 180901.

(35) Sherrill, C. D.; Manolopoulos, D. E.; Martínez, T. J.; Michaelides, A. Electronic structure software. *J. Chem. Phys.* **2020**, *153*, 070401.

(36) Spotte-Smith, E. W. C.; Blau, S.; Xie, X.; Patel, H.; Wen, M.; Wood, B.; Dwaraknath, S.; Auslag Persoon, K. Quantum chemical calculations of lithium-ion battery electrolyte and interphase species. *Scientific Data* **2021**, *8*, 203.

(37) Visayas, B. R. B.; Pahari, S. K.; Gokoglan, T. C.; Golen, J. A.; Agar, E.; Cappillino, P. J.; Mayes, M. L. Computational and experimental investigation of the effect of cation structure on the solubility of anionic flow battery active-materials. *Chem. Sci.* **2021**, *12*, 15892−15907.

(38) Asha, A. S.; Iroegbu, J. N.; Visayas, B. R. B.; Mayes, M.; Shen, C. Molecular insights into the electric double-layer structure at a polymer electrolyte-electrode interface. *Electrochim. Acta* **2023**, *446*, No. 142131.

(39) Choudhary, K.; Bercx, M.; Jiang, J.; Pachter, R.; Lamoen, D.; Tavazza, F. Accelerated Discovery of Efficient Solar Cell Materials Using Quantum and Machine-Learning Methods. *Chem. Mater.* **2019**, *31*, 5900−5908.

(40) *Basic Energy Sciences Roundtable: Liquid Solar Fuels.* https://science.osti.gov/-/media/bes/pdf/reports/2020/Liquid_Solar_Fuels_Report.pdf (accessed 2023-08-25).

(41) Ess, D.; Gagliardi, L.; Hammes-Schiffer, S. Introduction: Computational Design of Catalysts from Molecules to Materials. *Chem. Rev.* **2019**, *119*, 6507−6508.

(42) Ludwig, J. R.; Schindler, C. S. Catalyst: Sustainable Catalysis. *Catalysis* **2017**, *2*, 313−316.

(43) Bullock, R. Reaction: Earth-abundant metal catalysts for energy conversions. *Catalysis* **2017**, *2*, 444−446.

(44) Constable, D. J. Reaction: Sustainable catalysis without metals. *Catalysis* **2017**, *2*, 446−447.

(45) Curtarolo, S.; Hart, G.; Buongiorno Nardelli, M.; Mingo, N.; Sanvito, S.; Levy, O. The high-throughput highway to computational materials design. *Nat. Mater.* **2013**, *12*, 191−201.

(46) Wasielewski, M. R.; Forbes, M. D.; Frank, N. L.; Kowalski, K.; Scholes, G. D.; Yuen-Zhou, J.; Baldo, M. A.; Freedman, D. E.; Goldsmith, R. H.; Goodson, T., III; et al. Exploiting chemistry and molecular systems for quantum information science. *Nature Reviews Chemistry* **2020**, *4*, 490−504.

(47) Cao, Y.; Romero, J.; Olson, J.; Degroote, M.; P.D, J.; Kieferov'a, M.; Kivlichan, I.; Menke, T.; Peropadre, B.; Sawaya, N.; Sim, S.; Veis, L.; Aspuru-Guzik, A. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **2019**, *119*, 10856−10915.

(48) Das, S.; Natarajan, G.; Ting, Y. Bio-extraction of precious metals from urban soil waste. *AIP Conf. Proc.* **2017**, *1805*, No. 020004.

(49) O'Brien, E.; Holt, M. E.; Thompson, M. K.; Salay, L. E.; Ehlinger, A. C.; Chazin, W. J.; Barton, J. K. The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* **2017**, *355*, No. eaag1789.

(50) Saha, A.; Arantes, P.; Palermo, G. Dynamics and Mechanism of CRISPR-Cas9 through the Lens of Computational Methods. *Current Opinions in Structural Biology* **2022**, *75*, No. 102400.

(51) Ray, A.; Di Felice, R. Protein-Mutation-Induced Conformational Changes of the DNA and Nuclease Domain in CRISPR/Cas9 Systems by Molecular Dynamics Simulations. *J. Phys. Chem. B* **2020**, *124*, 2168−2179.

(52) National Academies of Sciences, E.; Medicine, *Research Agenda for Transforming Separation Science*; 2019.

(53) Prodius, D.; Klocke, M.; Smetana, V.; Alammar, T.; Perez Garcia, M.; Windus, T. L.; Nlebedim, I. C.; Mudring, A.-V. Rationally designed rare earth separation by selective oxalate solubilization. *Chem. Commun.* **2020**, *56*, 11386−11389.

(54) Penchoff, D. A.; Valeev, E.; Jagode, H.; Luszczek, P.; Danalis, A.; Bosilca, G.; Harrison, R. J.; Dongarra, J.; Windus, T. L. *Rare Earth Elements and Actinides: Progress in Computational Science Applications*; Chapter 1, pp 3−53.

(55) García Alejo, A.; Silva, N. D.; Liu, Y.; Windus, T. L.; García, M. P. Solvent Phase Optimizations Improve Correlations with Experimental Stability Constants for Aqueous Lanthanide Complexes. *Solvent Extr. Ion Exch.* **2023**, *41*, 241−251.

(56) Gagliardi, L. The study of actinide chemistry with multi-configurational quantum chemical methods. *Int. J. Quantum Chem.* **2011**, *111*, 3302−3306.

(57) Kaltsoyannis, N. Transuranic Computational Chemistry. *Chem.—Eur. J.* **2018**, *24*, 2815−2825.

(58) Wooles, A. J.; Mills, D. P.; Tuna, F.; McInnes, E. J.; Law, G. T.; Fuller, A. J.; Kremer, F.; Ridgway, M.; Lewis, W.; Gagliardi, L.; et al. Uranium (III)-carbon multiple bonding supported by arene δ-bonding in mixed-valence hexauranium nanometre-scale rings. *Nat. Commun.* **2018**, *9*, 2097.

(59) Dolg, M., Ed.; *Computational Methods in Lanthanide and Actinide Chemistry*; John Wiley & Sons, Ltd.: West Sussex, 2015.

(60) Schädel, M. Chemistry of the superheavy elements. *Philosophical Transactions A* **2015**, *373*, No. 20140191.

(61) Jacob, D. J. *Introduction to atmospheric chemistry*; Princeton University Press: 1999.

(62) Gardiner, W. C., Ed.; *Gas-phase combustion chemistry*; Springer: 2000.

(63) Kohse-Höinghaus, K. Combustion, Chemistry, and Carbon Neutrality. *Chem. Rev.* **2023**, *123*, 5139.

(64) Young, L.; Ueda, K.; Gühr, M.; Bucksbaum, P. H.; Simon, M.; Mukamel, S.; Rohringer, N.; Prince, K. C.; Masciovecchio, C.; Meyer, M.; et al. Roadmap of ultrafast x-ray atomic and molecular physics. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2018**, *51*, No. 032003.

(65) Kraus, P. M.; Zürch, M.; Cushing, S. K.; Neumark, D. M.; Leone, S. R. The ultrafast X-ray spectroscopic revolution in chemical dynamics. *Nature Reviews Chemistry* **2018**, *2*, 82−94.

(66) Ghimire, S.; Reis, D. A. High-harmonic generation from solids. *Nat. Phys.* **2019**, *15*, 10−16.

(67) Palacios, A.; Martín, F. The quantum chemistry of attosecond molecular science. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2020**, *10*, No. e1430.

(68) Li, J.; Lu, J.; Chew, A.; Han, S.; Li, J.; Wu, Y.; Wang, H.; Ghimire, S.; Chang, Z. Attosecond science based on high harmonic generation from gases and solids. *Nat. Commun.* **2020**, *11*, 2748.

(69) Goulielmakis, E.; Brabec, T. High harmonic generation in condensed matter. *Nat. Photonics* **2022**, *16*, 411−421.

(70) Jagau, T.-C.; Bravaya, K. B.; Krylov, A. I. Extending quantum chemistry of bound states to electronic resonances. *Annu. Rev. Phys. Chem.* **2017**, *68*, 525−553.

(71) Jagau, T.-C. Theory of electronic resonances: fundamental aspects and recent advances. *Chem. Commun.* **2022**, *58*, 5205−5224.

(72) Bostedt, C.; Boutet, S.; Fritz, D. M.; Huang, Z.; Lee, H. J.; Lemke, H. T.; Robert, A.; Schlotter, W. F.; Turner, J. J.; Williams, G. J. Linac coherent light source: The first five years. *Rev. Mod. Phys.* **2016**, *88*, No. 015007.

(73) Pellegrini, C. X-ray free-electron lasers: from dreams to reality. *Phys. Scr.* **2016**, *T169*, No. 014004.

(74) Bergman, U.; Yachandra, V. K.; Yano, J. *X-ray free electron lasers: applications in materials, chemistry and biology*; Royal Society of Chemistry: 2017; Vol. 18.

(75) Schoenlein, R.; Boutet, S.; Minitti, M.; Dunne, A. The Linac Coherent Light Source: recent developments and future plans. *Applied Sciences* **2017**, *7*, 850.

(76) Bergmann, U.; Kern, J.; Schoenlein, R. W.; Wernet, P.; Yachandra, V. K.; Yano, J. Using X-ray free-electron lasers for spectroscopy of molecular catalysts and metalloenzymes. *Nature Reviews Physics* **2021**, *3*, 264−282.

(77) Sun, S.; Stetina, T. F.; Zhang, T.; Hu, H.; Valeev, E. F.; Sun, Q.; Li, X. Efficient Four-Component Dirac-Coulomb-Gaunt Hartree-Fock in the Pauli Spinor Representation. *J. Chem. Theory Comput.* **2021**, *17*, 3388−3402.

(78) Sun, S.; Ehrman, J. N.; Sun, Q.; Li, X. Efficient Evaluation of the Breit Operator in the Pauli Spinor Basis. *J. Chem. Phys.* **2022**, *157*, No. 064112.

(79) Hoyer, C. E.; Lu, L.; Hu, H.; Shumilov, K. D.; Sun, S.; Knecht, S.; Li, X. Correlated Dirac-Coulomb-Breit Multiconfigurational Self-Consistent-Field Methods. *J. Chem. Phys.* **2023**, *158*, No. 044101.

(80) Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **1981**, *23*, 5048−5079.

(81) Pederson, M.; Ruzsinszky, J. P.; Perdew, A. Communication: Self-interaction correction with unitary invariance in density functional theory. *J. Chem. Phys.* **2014**, *140*, No. 121103.

(82) Teale, A. M.; Helgaker, T.; Savin, A.; Adamo, C.; Aradi, B.; Arbuznikov, A. V.; Ayers, P. W.; Baerends, E. J.; Barone, V.; Calaminici, P.; et al. DFT exchange: sharing perspectives on the workhorse of quantum chemistry and materials science. *Physical chemistry chemical physics* **2022**, *24*, 28700−28781.

(83) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256−4266.

(84) Bartlett, R.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291−352.

(85) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A FIFTH-ORDER PERTURBATION COMPARISON OF ELECTRON CORRELATION THEORIES. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(86) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse Maps—A Systematic Infrastructure for Reduced-Scaling Electronic Structure Methods. II. Linear Scaling Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Phys.* **2016**, *144*, No. 024109.

(87) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483−5488.

(88) Pulay, P. A perspective on the CASPT2 method. *Int. J. Quantum Chem.* **2011**, *111*, 3273−3279.

(89) Ma, D.; Li Manni, G.; Gagliardi, L. The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* **2011**, *135*, No. 044128.

(90) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669−3680.

(91) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference nature of chemistry: The coupled-cluster view. *Chem. Rev.* **2012**, *112*, 182−243.

(92) Baiardi, A.; Reiher, M. The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges. *J. Chem. Phys.* **2020**, *152*, No. 040903.

(93) Zhai, H.; Chan, G. K.-L. Low communication high performance ab initio density matrix renormalization group algorithms. *J. Chem. Phys.* **2021**, *154*, No. 224116.

(94) Lee, S.; Zhai, H.; Sharma, S.; Umrigar, C. J.; Chan, G. K.-L. Externally corrected ccsd with renormalized perturbative triples (R-ecCCSD (T)) and the density matrix renormalization group and selected configuration interaction external sources. *J. Chem. Theory Comput.* **2021**, *17*, 3414−3425.

(95) Hoyer, C. E.; Hu, H.; Lu, L.; Knecht, S.; Li, X. Relativistic Kramers-Unrestricted Exact-Two-Component Density Matrix Renormalization Group. *J. Phys. Chem. A* **2022**, *126*, 5011−5020.

(96) Shepard, R. A general nonlinear expansion form for electronic wave functions. *J. Phys. Chem. A* **2005**, *109*, 11629−11641.

(97) Shepard, R.; Gidofalvi, G.; Brozell, S. R. The multifacet graphically contracted function method. I. Formulation and implementation. *J. Chem. Phys.* **2014**, *141*, No. 064105.

(98) Lestrange, P. J.; Hoffmann, M. R.; Li, X. Time-Dependent Configuration Interaction using the Graphical Unitary Group Approach: Nonlinear Electric Properties. *Adv. Quantum Chem.* **2018**, *76*, 295−313.

(99) Lischka, H.; Nachtigallova, D.; Aquino, A. J.; Szalay, P. G.; Plasser, F.; Machado, F. B.; Barbatti, M. Multireference approaches for excited states of molecules. *Chem. Rev.* **2018**, *118*, 7293−7361.

(100) Hu, H.; Jenkins, A. J.; Liu, H.; Kasper, J. M.; Frisch, M. J.; Li, X. Relativistic Two-Component Multireference Configuration Interaction Method with Tunable Correlation Space. *J. Chem. Theory Comput.* **2020**, *16*, 2975−2984.

(101) Malmqvist, P. Å.; Rendell, A.; Roos, B. O. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *J. Phys. Chem.* **1990**, *94*, 5477−5482.

(102) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

(103) Lu, L.; Hu, H.; Jenkins, A. J.; Li, X. Exact-Two-Component Relativistic Multireference Second-Order Perturbation Theory. *J. Chem. Theory Comput.* **2022**, *18*, 2983−2992.

(104) Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. Multireference Electron Correlation Methods: Journeys along Potential Energy Surfaces. *Chem. Rev.* **2020**, *120*, 5878−5909.

(105) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252−10264.

(106) Guo, S.; Li, Z.; Chan, G. K.-L. A perturbative density matrix renormalization group algorithm for large active spaces. *J. Chem. Theory Comput.* **2018**, *14*, 4063−4071.

(107) Sharma, P.; Bao, J. J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration pair-density functional theory. *Annu. Rev. Phys. Chem.* **2021**, *72*, 541−564.

(108) Sharma, P.; Jenkins, A. J.; Scalmani, G.; Frisch, M. J.; Truhlar, D. G.; Gagliardi, L.; Li, X. Exact-Two-Component Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2022**, *18*, 2947−2954.

(109) Landinez Borda, E. J.; Gomez, J.; Morales, M. A. Non-orthogonal multi-Slater determinant expansions in auxiliary field quantum Monte Carlo. *J. Chem. Phys.* **2019**, *150*, No. 074105.

(110) Mahajan, A.; Lee, J.; Sharma, S. Selected configuration interaction wave functions in phaseless auxiliary field quantum Monte Carlo. *J. Chem. Phys.* **2022**, *156*, No. 174111.

(111) Booth, G. H.; Thom, A. J.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, No. 054106.

(112) Deustua, J. E.; Shen, J.; Piecuch, P. Converging high-level coupled-cluster energetics by Monte Carlo sampling and moment expansions. *Physical review letters* **2017**, *119*, No. 223003.

(113) Degroote, M.; Henderson, T. M.; Zhao, J.; Dukelsky, J.; Scuseria, G. E. Polynomial similarity transformation theory: A smooth interpolation between coupled cluster doubles and projected BCS applied to the reduced BCS Hamiltonian. *Phys. Rev. B* **2016**, *93*, No. 125124.

(114) Qiu, Y.; Henderson, T. M.; Zhao, J.; Scuseria, G. E. Projected coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 064111.

(115) Motta, M.; et al. Towards the Solution of the Many-Electron Problem in Real Materials: Equation of State of the Hydrogen Chain with State-of-the-Art Many-Body Methods. *Phys. Rev. X* **2017**, *7*, No. 031059.

(116) Kats, D.; Manby, F. R. Communication: The distinguishable cluster approximation. *J. Chem. Phys.* **2013**, *139*, 021102.

(117) Hedin, L. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.* **1965**, *139*, A796.

(118) Nooijen, M.; Snijders, J. G. Coupled cluster approach to the single-particle Green's function. *Int. J. Quantum Chem.* **1992**, *44*, 55−83.

(119) Nooijen, M.; Snijders, J. G. Coupled cluster Green's function method: Working equations and applications. *Int. J. Quantum Chem.* **1993**, *48*, 15−48.

(120) Meissner, L.; Bartlett, R. J. Electron propagator theory with the ground state correlated by the coupled-cluster method. *Int. J. Quantum Chem.* **1993**, *48*, 67−80.

(121) McClain, J.; Lischner, J.; Watson, T.; Matthews, D. A.; Ronca, E.; Louie, S. G.; Berkelbach, T. C.; Chan, G. K.-L. Spectral functions of the uniform electron gas via coupled-cluster theory and comparison to the G W and related approximations. *Phys. Rev. B* **2016**, *93*, No. 235139.

(122) Peng, B.; Bauman, N. P.; Gulania, S.; Kowalski, K. Coupled cluster Green's function: Past, present, and future. In *Annual Reports in Computational Chemistry*; Dixon, D. A., Ed.; Elsevier: 2021; Vol. *17*, pp 23−53.

(123) von Niessen, W.; Schirmer, J.; Cederbaum, L. S. Computational methods for the one-particle green's function. *Computer Physics Reports* **1984**, *1*, 57−125.

(124) Mertins, F.; Schirmer, J. Algebraic propagator approaches and intermediate-state representations. I. The biorthogonal and unitary coupled-cluster methods. *Phys. Rev. A* **1996**, *53*, 2140−2152.

(125) Schirmer, J.; Trofimov, A. Intermediate state representation approach to physical properties of electronically excited molecules. *J. Chem. Phys.* **2004**, *120*, 11449−11464.

(126) Hirata, S.; Hermes, M. R.; Simons, J.; Ortiz, J. General-order many-body green's function method. *J. Chem. Theory Comput.* **2015**, *11*, 1595−1606.

(127) Hirata, S.; Doran, A. E.; Knowles, P. J.; Ortiz, J. One-particle many-body Green's function theory: Algebraic recursive definitions,

linked-diagram theorem, irreducible-diagram theorem, and general-order algorithms. *J. Chem. Phys.* **2017**, *147*, No. 044108.

(128) Galitskii, V. M.; Migdal, A. B. Application of quantum field theory methods to the many body problem. *Sov. Phys. JETP* **1958**, *7*, 18.

(129) Luttinger, J. M.; Ward, J. C. Ground-state energy of a many-fermion system. II. *Phys. Rev.* **1960**, *118*, 1417.

(130) Baym, G.; Kadanoff, L. P. Conservation laws and correlation functions. *Phys. Rev.* **1961**, *124*, 287.

(131) Monkhorst, H. J. Calculation of Properties with the Coupled-Cluster Method. *Int. J. Quantum Chem.* **1977**, *12*, 421−432.

(132) Koch, H.; Jørgensen, P. Coupled Cluster Response Functions. *J. Chem. Phys.* **1990**, *93*, 3333−3344.

(133) Sekino, H.; Bartlett, R. J. A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem.* **1984**, *26*, 255−265.

(134) Gauss, J. Calculation of NMR chemical shifts at second-order many-body perturbation theory using gauge-including atomic orbitals. *Chemical physics letters* **1992**, *191*, 614−620.

(135) Russ, N. J.; Crawford, T. D. Local correlation in coupled cluster calculations of molecular response properties. *Chemical physics letters* **2004**, *400*, 104−111.

(136) Geertsen, J.; Rittby, M.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method: Excitation Energies of Be and CO. *Chem. Phys. Lett.* **1989**, *164*, 57−62.

(137) Comeau, D. C.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method. Applications to Open- and Closed-Shell Reference States. *Chem. Phys. Lett.* **1993**, *207*, 414−423.

(138) Stanton, J. F.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method. A Systematic Biorthogonal Approach to Molecular Excitation Energies, Transition Probabilities, and Excited State Properties. *J. Chem. Phys.* **1993**, *98*, 7029−7039.

(139) Piecuch, P.; Bartlett, R. J. In *EOMXCC: A New Coupled-Cluster Method for Electronic Excited States*; Lowdin, P.-O., Sabin, J. R., Zerner, M. C., Brandas, E., Eds.; Advances in Quantum Chemistry Supplement C; Academic Press: 1999; Vol. *34*, pp 295−380.

(140) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433−462.

(141) Casanova, D.; Krylov, A. I. Spin-flip methods in quantum chemistry. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4326−4342.

(142) Li, X.; Govind, N.; Isborn, C.; DePrince, A. E., III; Lopata, K. Real-time time-dependent electronic structure theory. *Chem. Rev.* **2020**, *120*, 9951−9993.

(143) Goings, J. J.; Kasper, J. M.; Egidi, F.; Sun, S.; Li, X. Real Time Propagation of the Exact Two Component Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2016**, *145*, No. 104107.

(144) Ruggenthaler, M.; Mackenroth, F.; Bauer, D. Time-dependent Kohn-Sham approach to quantum electrodynamics. *Phys. Rev. A* **2011**, *84*, No. 042107.

(145) Zhao, L.; Tao, Z.; Pavošević, F.; Wildman, A.; Hammes-Schiffer, S.; Li, X. Real-Time Time-Dependent Nuclear-Electronic Orbital Approach: Dynamics beyond the Born−Oppenheimer Approximation. The. *J. Phys. Chem. Lett.* **2020**, *11*, 4052−4058.

(146) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent quantum chemistry: Integrating electronic and nuclear quantum effects via the nuclear-electronic orbital method. *Chem. Rev.* **2020**, *120*, 4222−4253.

(147) Haugland, T. S.; Ronca, E.; Kjønstad, E. F.; Rubio, A.; Koch, H. Coupled cluster theory for molecular polaritons: Changing ground and excited states. *Physical Review X* **2020**, *10*, No. 041043.

(148) Riso, R. R.; Haugland, T. S.; Ronca, E.; Koch, H. Molecular orbital theory in cavity QED environments. *Nat. Commun.* **2022**, *13*, 1368.

(149) Liebenthal, M. D.; Vu, N.; DePrince, A. E., III Equation-of-motion cavity quantum electrodynamics coupled-cluster theory for electron attachment. *J. Chem. Phys.* **2022**, *156*, No. 054105.

(150) Seijo, L.; Barandiaran, Z. The Ab Initio Model Potential Method: A Common Strategy for Effective Core Potential and Embedded Cluster Calculations. *Computational chemistry: reviews of current trends*; World Scientific: 1999; pp 55−152.

(151) Jacob, C. R.; Neugebauer, J. Subsystem density-functional theory. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 325−362.

(152) Libisch, F.; Huang, C.; Carter, E. A. Embedded correlated wavefunction schemes: Theory and applications. *Acc. Chem. Res.* **2014**, *47*, 2768−2775.

(153) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* **2015**, *115*, 5891−5928.

(154) Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705−2712.

(155) Sushko, P. V.; Huang, C.; Govind, N.; Kowalski, K. Embedding Methods in Materials Discovery. *Computational Materials Discovery* **2018**, 87−116.

(156) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and Pt(P(t-Bu)3)2 + H2 Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(157) Chung, L. W.; Sameera, W.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; et al. The ONIOM method and its applications. *Chem. Rev.* **2015**, *115*, 5678−5796.

(158) Govind, N.; Wang, Y.; Da Silva, A.; Carter, E. Accurate ab initio energetics of extended systems via explicit correlation embedded in a density functional environment. *Chem. Phys. Lett.* **1998**, *295*, 129−134.

(159) Huang, C.; Pavone, M.; Carter, E. A. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.* **2011**, *134*, No. 154110.

(160) Elliott, P.; Burke, K.; Cohen, M. H.; Wasserman, A. Partition density-functional theory. *Phys. Rev. A* **2010**, *82*, No. 024501.

(161) Nafziger, J.; Wasserman, A. Density-based partitioning methods for ground-state molecular calculations. *J. Phys. Chem. A* **2014**, *118*, 7623−7639.

(162) Stoll, H. Correlation energy of diamond. *Phys. Rev. B* **1992**, *46*, 6700.

(163) Stoll, H.; Paulus, B.; Fulde, P. On the accuracy of correlation-energy expansions in terms of local increments. *J. Chem. Phys.* **2005**, *123*, No. 144108.

(164) Fedorov, D. G.; Kitaura, K. Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *J. Phys. Chem. A* **2007**, *111*, 6904−6914.

(165) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632−672.

(166) Huang, C.; Carter, E. A. Potential-functional embedding theory for molecules and materials. *J. Chem. Phys.* **2011**, *135*, No. 194104.

(167) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III Embedded mean-field theory. *J. Chem. Theory Comput.* **2015**, *11*, 568−580.

(168) Inglesfield, J. A method of embedding. *J. Phys. C* **1981**, *14*, 3795.

(169) Chibani, W.; Ren, X.; Scheffler, M.; Rinke, P. Self-consistent Green's function embedding for advanced electronic structure methods based on a dynamical mean-field concept. *Phys. Rev. B* **2016**, *93*, No. 165106.

(170) Lan, T. N.; Kananenka, A. A.; Zgid, D. Communication: Towards ab initio self-energy embedding theory in quantum chemistry. *J. Chem. Phys.* **2015**, *143*, No. 241102.

(171) Rusakov, A. A.; Iskakov, S.; Tran, L. N.; Zgid, D. Self-energy embedding theory (SEET) for periodic systems. *J. Chem. Theory Comput.* **2019**, *15*, 229−240.

(172) Parravicini, V.; Jagau, T.-C. Embedded equation-of-motion coupled-cluster theory for electronic excitation, ionisation, electron attachment, and electronic resonances. *Mol. Phys.* **2021**, *119*, No. e1943029.

(173) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Physical review letters* **2012**, *109*, No. 186404.

(174) Knizia, G.; Chan, G. K.-L. Density Matrix Embedding: A Strong-Coupling Quantum Embedding Theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428−1432.

(175) Cui, Z.-H.; Zhu, T.; Chan, G. K.-L. Efficient implementation of ab initio quantum embedding in periodic systems: Density matrix embedding theory. *J. Chem. Theory Comput.* **2020**, *16*, 119−129.

(176) Pham, H. Q.; Bernales, V.; Gagliardi, L. Can Density Matrix Embedding Theory with the Complete Activate Space Self-Consistent Field Solver Describe Single and Double Bond Breaking in Molecular Systems? *J. Chem. Theory Comput.* **2018**, *14*, 1960.

(177) Pham, H. Q.; Hermes, M. R.; Gagliardi, L. Periodic Electronic Structure Calculations with the Density Matrix Embedding Theory. *J. Chem. Theory Comput.* **2020**, *16*, 130−140.

(178) Li, W.; Chen, M.; Rabani, E.; Baer, R.; Neuhauser, D. Stochastic embedding DFT: Theory and application to p-nitroaniline in water. *J. Chem. Phys.* **2019**, *151*, No. 174115.

(179) Georges, A.; Kotliar, G.; Krauth, W.; Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.* **1996**, *68*, 13−125.

(180) Held, K. Electronic structure calculations using dynamical mean field theory. *Advances in physics* **2007**, *56*, 829−926.

(181) Zhu, T.; Cui, Z.-H.; Chan, G. K.-L. Efficient formulation of ab initio quantum embedding in periodic systems: Dynamical mean-field theory. *J. Chem. Theory Comput.* **2020**, *16*, 141−153.

(182) Huzinaga, S.; Cantu, A. A. Theory of Separability of Many-Electron Systems. *J. Chem. Phys.* **1971**, *55*, 5543−5549.

(183) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.

(184) Claudino, D.; Mayhall, N. J. Automatic Partition of Orbital Spaces Based on Singular Value Decomposition in the Context of Embedding Theories. *J. Chem. Theory Comput.* **2019**, *15*, 1053−1064.

(185) Hégely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact density functional and wave function embedding schemes based on orbital localization. *J. Chem. Phys.* **2016**, *145*, No. 064107.

(186) Waldrop, J. M.; Windus, T. L.; Govind, N. Projector-based quantum embedding for molecular systems: An investigation of three partitioning approaches. *J. Phys. Chem. A* **2021**, *125*, 6384−6393.

(187) Marx, D.; Hutter, J. Ab initio molecular dynamics: Theory and implementation. *Modern methods and algorithms of quantum chemistry*; 2000; Vol. *1*, p 141.

(188) Mouvet, F.; Villard, J.; Bolnykh, V.; Rothlisberger, U. Recent advances in first-principles based molecular dynamics. *Acc. Chem. Res.* **2022**, *55*, 221−230.

(189) Burnham, C. J.; Xantheas, S. S. Development of transferable interaction models for water. I. Prominent features of the water dimer potential energy surface. *J. Chem. Phys.* **2002**, *116*, 1479−1492.

(190) Feynman, R. P. Space-time approach to non-relativistic Quantum Mechanics. *Rev. Mod. Phys.* **1948**, *20*, 367−387.

(191) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.

(192) Broughton, J. Q.; Abraham, F. F.; Bernstein, N.; Kaxiras, E. Concurrent coupling of length scales: Methodology and application. *Phys. Rev. B* **1999**, *60*, 2391−2403.

(193) Tavazza, F.; Levine, L. E.; Chaka, A. M. Hybrid methods for Atomic-Level simulations spanning multiple-length scales in Solid State. *Reviews in Computational Chemistry*; 2008; Vol. *26*, p 287,.

(194) Vassetti, D.; Labat, F. Towards a transferable nonelectrostatic model for continuum solvation: The electrostatic and nonelectrostatic energy correction model. *J. Comput. Chem.* **2022**, *43*, 1372−1387.

(195) Heindel, J. P.; Herman, K. M.; Aprá, E.; Xantheas, S. S. Guest-host interactions in clathrate hydrates: Benchmark MP2 and CCSD (T)/CBS binding energies of CH_4, CO_2, and H_2S in (H_2O) 20 cages. The. *J. Phys. Chem. Lett.* **2021**, *12*, 7574−7582.

(196) Ayala, P. Y.; Scuseria, G. E. Linear Scaling Second-Order M?ller−Plesset Theory in the Atomic Orbital Basis for Large Molecular Systems. *J. Chem. Phys.* **1999**, *110*, 3660.

(197) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691−5705.

(198) Schütz, M.; Werner, H.-J. Local perturbative triples correction (T) with linear cost scaling. *Chem. Phys. Lett.* **2000**, *318*, 370−378.

(199) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and Accurate Approximations to the Local Coupled Cluster Singles Doubles Method Using a Truncated Pair Natural Orbital Basis. *J. Chem. Phys.* **2009**, *131*, No. 064103.

(200) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. *J. Chem. Phys.* **2013**, *139*, No. 134101.

(201) Riplinger, C.; Neese, F. An Efficient and near Linear Scaling Pair Natural Orbital Based Local Coupled Cluster Method. *J. Chem. Phys.* **2013**, *138*, No. 034106.

(202) Pavošević, F.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps-A Systematic Infrastructure for Reduced-Scaling Electronic Structure Methods. IV. Linear-scaling Second-Order Explicitly Correlated Energy with Pair Natural Orbitals. *J. Chem. Phys.* **2016**, *144*, No. 144109.

(203) Pavošević, F.; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps-A Systematic Infrastructure for Reduced Scaling Electronic Structure Methods. V. Linear Scaling Explicitly Correlated Coupled-Cluster Method with Pair Natural Orbitals. *J. Chem. Phys.* **2017**, *146*, No. 174108.

(204) Crawford, T. D.; Kumar, A.; Bazanté, A. P.; Di Remigio, R. Reduced-scaling coupled cluster response theory: Challenges and opportunities. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, *9*, No. e1406.

(205) D'Cunha, R.; Crawford, T. D. Pno++: Perturbed pair natural orbitals for coupled cluster linear response theory. *J. Chem. Theory Comput.* **2021**, *17*, 290−301.

(206) Wang, Z.; Peyton, B. G.; Crawford, T. D. Accelerating real-time coupled cluster methods with single-precision arithmetic and adaptive numerical integration. *J. Chem. Theory Comput.* **2022**, *18*, 5479−5491.

(207) Kutzelnigg, W. R. 12-Dependent Terms in the Wave Function as Closed Sums of Partial Wave Amplitudes for Large l. *Theor. Chim. Acta* **1985**, *68*, 445−469.

(208) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly correlated R12/F12 methods for electronic structure. *Chem. Rev.* **2012**, *112*, 75−107.

(209) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4−74.

(210) Snyder, J. C.; Rupp, M.; Hansen, K.; Müller, K.-R.; Burke, K. Finding Density Functionals with Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, No. 253002.

(211) Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **2018**, *361*, 360−365.

(212) Welborn, M.; Cheng, L.; Miller, T. F., III Transferability in machine learning for electronic structure via the molecular orbital basis. *J. Chem. Theory Comput.* **2018**, *14*, 4772−4779.

(213) Clark, A. E.; Adams, H.; Hernandez, R.; Krylov, A. I.; Niklasson, A. M.; Sarupria, S.; Wang, Y.; Wild, S. M.; Yang, Q. middle science: Traversing scale in complex many-body systems. *middle science: Traversing scale in complex many-body systems* **2021**, *7*, 1271.

(214) Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the electronic structure problem with machine learning. *npj Computational Materials* **2019**, *5*, 22.

(215) Dral, P. O. Quantum chemistry in the age of machine learning. *journal of physical chemistry letters* **2020**, *11*, 2336−2347.

(216) Kulik, H. J. Roadmap on Machine learning in electronic structure. *Electronic Structure* **2022**, *4*, 023004.

(217) Feynman, R. P. Forces in molecules. *Physical review* **1939**, *56*, 340.

(218) Windus, T. et al. *Basic Energy Sciences Exascale Requirements Review*; An Office of Science review sponsored jointly by Advanced Scientific Computing Research and Basic Energy Sciences, November 3−5, 2015, Rockville, Maryland; 2017.

(219) Dubey, A.; McInnes, L. C.; Thakur, R.; Draeger, E. W.; Evans, T.; Germann, T. C.; Hart, W. E. Performance portability in the exascale computing project: exploration through a panel series. *Computing in Science & Engineering* **2021**, *23*, 46−54.

(220) Gordon, M. S.; Windus, T. L. Editorial: Modern Architectures and Their Impact on Electronic Structure Theory. *Chem. Rev.* **2020**, *120*, 9015−9020.

(221) Finkelstein, J.; Smith, J. S.; Mniszewski, S. M.; Barros, K.; Negre, C. F. A.; Rubensson, E. H.; Niklasson, A. M. N. Mixed Precision Fermi-Operator Expansion on Tensor Cores from a Machine Learning Perspective. *J. Chem. Theory Comput.* **2021**, *17*, 2256−2265.

(222) Finkelstein, J.; Smith, J. S.; Mniszewski, S. M.; Barros, K.; Negre, C. F. A.; Rubensson, E. H.; Niklasson, A. M. N. Quantum-Based Molecular Dynamics Simulations Using Tensor Cores. *J. Chem. Theory Comput.* **2021**, *17*, 6180−6192.

(223) Finkelstein, J.; Rubensson, E. H.; Mniszewski, S. M.; Negre, C. F. A.; Niklasson, A. M. N. Quantum Perturbation Theory Using Tensor Cores and a Deep Neural Network. *J. Chem. Theory Comput.* **2022**, *18*, 4255−4268.

(224) Pederson, R.; Kozlowski, J.; Song, R.; Beall, J.; Ganahl, M.; Hauru, M.; Lewis, A. G. M.; Yao, Y.; Mallick, S. B.; Blum, V.; Vidal, G. Large Scale Quantum Chemistry with Tensor Processing Units. *J. Chem. Theory Comput.* **2023**, *19*, 25−32.

(225) Pederson, R.; Kozlowski, J.; Song, R.; Beall, J.; Ganahl, M.; Hauru, M.; Lewis, A. G. M.; Yao, Y.; Mallick, S. B.; Blum, V.; Vidal, G. Large Scale Quantum Chemistry with Tensor Processing Units. *J. Chem. Theory Comput.* **2023**, *19*, 25−32.

(226) Götz, A. W.; Wölfle, T.; Walker, R. C. Chapter 2 - Quantum Chemistry on Graphics Processing Units. Wheeler, R. A., Ed.; In *Annual Reports in Computational Chemistry*; Elsevier: 2010; Vol. 6, pp 21−35.

(227) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9*, 213−221.

(228) Blackford, S. L. *ACM Trans. Math. Softw.* **2002**, *28*, 135−151.

(229) Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Du Croz, J.; Greenbaum, A.; Hammarling, S.; McKenney, A.; Sorensen, D. *LAPACK Users' Guide*, 3rd ed.; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1999.

(230) Blackford, L. S.; Choi, J.; Cleary, A.; D'Azeuedo, E.; Demmel, J.; Dhillon, I.; Hammarling, S.; Henry, G.; Petitet, A.; Stanley, K.; Walker, D.; Whaley, R. C.; Dongarra, J. J. *ScaLAPACK User's Guide*; Society for Industrial and Applied Mathematics: USA, 1997.

(231) Kŭs, P.; Marek, A.; Köcher, S.; Kowalski, H.-H.; Carbogno, C.; Scheurer, C.; Reuter, K.; Scheffler, M.; Lederer, H. Optimizations of the eigensolvers in the ELPA library. *Parallel Computing* **2019**, *85*, 167−177.

(232) Marek, A.; Blum, V.; Johanni, R.; Havu, V.; Lang, B.; Auckenthaler, T.; Heinecke, A.; Bungartz, H.-J.; Lederer, H. The ELPA library: scalable parallel eigenvalue solutions for electronic structure theory and computational science. *J. Phys.: Condens. Matter* **2014**, *26*, No. 213201.

(233) Yu, V. W.-z.; Moussa, J.; Kŭs, P.; Marek, A.; Messmer, P.; Yoon, M.; Lederer, H.; Blum, V. GPU-acceleration of the ELPA2 distributed eigensolver for dense symmetric and hermitian eigenproblems. *Comput. Phys. Commun.* **2021**, *262*, No. 107808.

(234) Heroux, M. A.; McInnes, L. C.; Thakur, R.; Vetter, J. S.; Li, X. S.; Aherns, J.; Munson, T.; Mohror, K. *ECP Software Technology Capability Assessment Report*; 2020.

(235) The Extreme-scale Scientific Software Stack (E4S). https://e4s-project.github.io (accessed 2023-03-04).

(236) Extreme-scale Scientific Software Development Kit (xSDK). https://xsdk.info (accessed 2023-03-04).

(237) Yu, V. W.-z.; et al. ELSI—An open infrastructure for electronic structure solvers. *Comput. Phys. Commun.* **2020**, *256*, No. 107459.

(238) Bock, N.; Mniszewski, S.; Aradi, B.; Wall, M. E.; Mohd-Yusof, C. F. A. N. J.; Niklasson, A. N. M. The basic matrix library (BML) for quantum chemistry. *J. Supercomput.* **2018**, *74*, 6201−6219.

(239) Oliveira, M. J. T.; et al. The CECAM electronic structure library and the modular software development paradigm. *J. Chem. Phys.* **2020**, *153*, No. 024117.

(240) Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **1986**, *84*, 3963−3974.

(241) Head-Gordon, M.; Pople, J. A. A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. *J. Chem. Phys.* **1988**, *89*, 5777−5786.

(242) Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. An efficient algorithm for the generation of two-electron repulsion integrals over gaussian basis functions. *Int. J. Quantum Chem.* **1989**, *36*, 269−280.

(243) McMurchie, L. E.; Davidson, E. R. One- and two-electron integrals over cartesian gaussian functions. *J. Comput. Phys.* **1978**, *26*, 218−231.

(244) King, H. F.; Dupuis, M. Numerical integration using rys polynomials. *J. Comput. Phys.* **1976**, *21*, 144−165.

(245) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222−231.

(246) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004−1015.

(247) Asadchev, A.; Allada, V.; Felder, J.; Bode, B. M.; Gordon, M. S.; Windus, T. L. Uncontracted Rys Quadrature Implementation of up to G Functions on Graphical Processing Units. *J. Chem. Theory Comput.* **2010**, *6*, 696−704.

(248) Miao, Y.; Merz, K. M. J. Acceleration of Electron Repulsion Integral Evaluation on Graphics Processing Units via Use of Recurrence Relations. *J. Chem. Theory Comput.* **2013**, *9*, 965−976.

(249) Asadchev, A.; Valeev, E. F. Memory-Efficient Recursive Evaluation of 3-Center Gaussian Integrals. 2022, arXiv:2210.03192. *arXiv Preprint*. https://arxiv.org/abs/2210.03192 (accessed 2023-08-25).

(250) Johnson, K. G.; Mirchandaney, S.; Hoag, E.; Heirich, A.; Aiken, A.; Martínez, T. J. Multinode Multi-GPU Two-Electron Integrals: Code Generation Using the Regent Language. *J. Chem. Theory Comput.* **2022**, *18*, 6522−6536.

(251) Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547−2553.

(252) Murray, C. W.; Handy, N. C.; Laming, G. J. Quadrature schemes for integrals of density functional theory. *Mol. Phys.* **1993**, *78*, 997−1014.

(253) Treutler, O.; Ahlrichs, R. Efficient molecular numerical integration schemes. *J. Chem. Phys.* **1995**, *102*, 346−354.

(254) Mura, M. E.; Knowles, P. J. Improved radial grids for quadrature in molecular density-functional calculations. *J. Chem. Phys.* **1996**, *104*, 9848−9858.

(255) Gill, P. M. W.; Chien, S.-H. Radial quadrature for multi-exponential integrands. *J. Comput. Chem.* **2003**, *24*, 732−740.

(256) Aprà, E.; et al. NWChem: Past, Present, and Future. *J. Chem. Phys.* **2020**, *152*, No. 184102.

(257) Shao, Y.; Head-Gordon, M. An improved J matrix engine for density functional theory calculations. *Chem. Phys. Lett.* **2000**, *323*, 425−433.

(258) Laqua, H.; Dietschreit, J. C. B.; Kussmann, J.; Ochsenfeld, C. Accelerating Hybrid Density Functional Theory Molecular Dynamics Simulations by Seminumerical Integration, Resolution-of-the-Identity Approximation, and Graphics Processing Units. *J. Chem. Theory Comput.* **2022**, *18*, 6010−6020.

(259) Calvin, J. A.; Lewis, C. A.; Valeev, E. F. Scalable Task-Based Algorithm for Multiplication of Block-Rank-Sparse Matrices. *Proceedings of the 5th Workshop on Irregular Applications: Architectures and Algorithms*; New York, NY, USA, 2015.

(260) Calvin, J. A.; Valeev, E. F. Task-based algorithm for matrix multiplication: A step towards block-sparse tensor computing. 2015, arXiv:1504.05046. *arXiv Preprint*. https://arxiv.org/abs/1504.05046 (accessed 2023-08-25).

(261) Mutlu, E.; Panyala, A.; Kowalski, K.; Bauman, N.; Peng, B.; Brabec, J.; Krishnamoorthy, S. TAMM: Tensor Algebra for Many-body Methods. 2022, arXiv:2201.01257. *arXiv Preprint*. https://arxiv.org/abs/2201.01257 (accessed 2023-08-25).

(262) Solomonik, E.; Matthews, D.; Hammond, J.; Demmel, J. Cyclops Tensor Framework: Reducing Communication and Eliminating Load Imbalance in Massively Parallel Contractions. *2013 IEEE 27th International Symposium on Parallel and Distributed Processing*; 2013; pp 813−824.

(263) Valeev, E. F. *Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions*, version 2.7.0-beta.6; 2020. http://libint.valeyev.net/, (accessed 2023-08-25).

(264) Sun, Q. Libcint: An efficient general integral library for Gaussian basis functions. *J. Comput. Chem.* **2015**, *36*, 1664−1671.

(265) Pritchard, B. P.; Chow, E. Horizontal vectorization of electron repulsion integrals. *J. Comput. Chem.* **2016**, *37*, 2537−2546.

(266) Williams-Young, D. B.; de Jong, W. A.; van Dam, H. J.; Yang, C. On the Efficient Evaluation of the Exchange Correlation Potential on Graphics Processing Unit Clusters. *Frontiers in Chemistry* **2020**, *8*, No. 581058.

(267) Williams-Young, D. B.; Bagusetty, A.; de Jong, W. A.; Doerfler, D.; van Dam, H. J.; Vázquez-Mayagoitia, Á.; Windus, T. L.; Yang, C. Achieving performance portability in Gaussian basis set density functional theory on accelerator based architectures in NWChemEx. *Parallel Computing* **2021**, *108*, No. 102829.

(268) Petrone, A.; Williams-Young, D. B.; Sun, S.; Stetina, T. F.; Li, X. An Efficient Implementation of Two-Component Relativistic Density Functional Theory with Torque-Free Auxiliary Variables. *European Physical Journal B* **2018**, *91*, 169.

(269) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in libxc−A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1−5.

(270) Ekström, U. *XCFun: A library of exchange-correlation functionals with arbitrary-order derivatives*; 2020. https://zenodo.org/record/3946698 (accessed 2023-08-25).

(271) Nickolls, J.; Buck, I.; Garland, M.; Skadron, K. Scalable parallel programming with cuda: Is cuda the parallel programming model that application developers have been waiting for? *Queue* **2008**, *6*, 40−53.

(272) de Supinski, B. R.; Scogland, T. R.; Duran, A.; Klemm, M.; Bellido, S. M.; Olivier, S. L.; Terboven, C.; Mattson, T. G. The ongoing evolution of openmp. *Proceedings of the IEEE* **2018**, *106*, 2004−2019.

(273) Herdman, J. A.; Gaudin, W. P.; Perks, O.; Beckingsale, D. A.; Mallinson, A. C.; Jarvis, S. A. Achieving Portability and Performance through OpenACC. *2014 First Workshop on Accelerator Programming using Directives*; 2014; pp 19−26.

(274) Stone, J. E.; Gohara, D.; Shi, G. OpenCL: A Parallel Programming Standard for Heterogeneous Computing Systems. *Computing in Science & Engineering* **2010**, *12*, 66−73.

(275) Trott, C. R.; et al. Kokkos 3: Programming Model Extensions for the Exascale Era. *IEEE Transactions on Parallel and Distributed Systems* **2022**, *33*, 805−817.

(276) Beckingsale, D. A.; Burmark, J.; Hornung, R.; Jones, H.; Killian, W.; Kunen, A. J.; Pearce, O.; Robinson, P.; Ryujin, B. S.; Scogland, T. R. RAJA: Portable performance for large-scale scientific applications. *2019 ieee/acm international workshop on performance, portability and productivity in hpc (p3hpc)*; 2019; pp 71−81.

(277) Eriksen, J. J. Efficient and Portable Acceleration of Quantum Chemical Many-Body Methods in Mixed Floating Point Precision Using OpenACC Compiler Directives. *Mol. Phys.* **2017**, *115*, 2086−2101.

(278) Team, N. *PluginPlay: A framework for developing modular software*; 2018. https://github.com/NWChemEx-Project/PluginPlay (accessed 2023-08-25).

(279) Hirata, S. Tensor Contraction Engine: Abstraction and Automated Parallel Implementation of Configuration-Interaction, Coupled-Cluster, and Many-Body Perturbation Theories. *J. Phys. Chem. A* **2003**, *107*, 9887−9897.

(280) Baumgartner, G.; et al. Synthesis of High-Performance Parallel Programs for a Class of Ab Initio Quantum Chemistry Models. *Proc. IEEE* **2005**, *93*, 276−292.

(281) Nieplocha, J.; Harrison, R. J.; Littlefield, R. J. Global Arrays: A Portable "Shared-Memory" Programming Model for Distributed Memory Computers. Supercomput. *Supercomputing '94: Proceedings of the 1994 ACM/IEEE conference on Supercomputing*; 1994; pp 340−349.

(282) Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. Explicitly Correlated Coupled-Cluster Singles and Doubles Method Based on Complete Diagrammatic Equations. *J. Chem. Phys.* **2008**, *129*, No. 071101.

(283) MacLeod, M. K.; Shiozaki, T. Communication: Automatic Code Generation Enables Nuclear Gradient Computations for Fully Internally Contracted Multireference Theory. *J. Chem. Phys.* **2015**, *142*, No. 051103.

(284) Lai, P.-W.; Zhang, H.; Rajbhandari, S.; Valeev, E.; Kowalski, K.; Sadayappan, P. Effective Utilization of Tensor Symmetry in Operation Optimization of Tensor Contraction Expressions. *Chem. Phys. Lett.* **2012**, *9*, 412−421.

(285) Epifanovsky, E.; Wormit, M.; Kuś, T.; Landau, A.; Zuev, D.; Khistyaev, K.; Manohar, P.; Kaliman, I.; Dreuw, A.; Krylov, A. I. New Implementation of High-Level Correlated Methods Using a General Block Tensor Library for High-Performance Electronic Structure Calculations. *J. Comput. Chem.* **2013**, *34*, 2293−2309.

(286) Solomonik, E.; Matthews, D.; Hammond, J. R.; Stanton, J. F.; Demmel, J. A Massively Parallel Tensor Contraction Framework for Coupled-Cluster Computations. *J. Parallel Distrib. Comput.* **2014**, *74*, 3176−3190.

(287) Calvin, J. A.; Lewis, C. A.; Valeev, E. F. Scalable Task-Based Algorithm for Multiplication of Block-Rank-Sparse Matrices. *IA3 15 5th Workshop Irregul. Appl. Archit. Algorithms. New York, New York, USA*; 2015; 1−8.

(288) Calvin, J. A.; Peng, C.; Rishi, V.; Kumar, A.; Valeev, E. F. Many-Body Quantum Chemistry on Massively Parallel Computers. *Chem. Rev.* **2021**, *121*, 1203−1231.

(289) Mutlu, E.; Panyala, A.; Gawande, N.; Bagusetty, A.; Kim, J.; Kowalski, K.; Bauman, N.; Peng, B.; Brabec, J.; Krishnamoorthy, S. TAMM: Tensor Algebra for Many-body Methods. *J. Chem. Phys.* **2023**, *159*, 024801.

(290) Fishman, M.; White, S. R.; Stoudenmire, E. M. *The ITensor Software Library for Tensor Network Calculations*; 2021. http://arxiv.org/abs/2007.14822 (accessed 2023-08-25).

(291) Sundararaman, R.; Letchworth-Weaver, K.; Schwarz, K. A.; Gunceler, D.; Ozhabes, Y.; Arias, T. JDFTx: Software for Joint Density-Functional Theory. *SoftwareX* **2017**, *6*, 278−284.

(292) Hjorth Larsen, A.; et al. The Atomic Simulation Environment— a Python Library for Working with Atoms. *J. Phys.: Condens. Matter* **2017**, *29*, No. 273002.

(293) Barbatti, M.; Ruckenbauer, M.; Plasser, F.; Pittner, J.; Granucci, G.; Persico, M.; Lischka, H. Newton- X: A Surface-hopping Program for Nonadiabatic Molecular Dynamics. *WIREs Comput. Mol. Sci.* **2014**, *4*, 26−33.

(294) Mai, S.; Marquetand, P.; González, L. Nonadiabatic Dynamics: The SHARC Approach. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1370.

(295) Smith, D. G. A.; et al. Quantum Chemistry Common Driver and Databases (QCDB) and Quantum Chemistry Engine (QCE Ngine): Automation and Interoperability among Computational Chemistry Programs. *J. Chem. Phys.* **2021**, *155*, No. 204801.

(296) O'boyle, N. M.; Tenderholt, A. L.; Langner, K. M. Cclib: A Library for Package-Independent Computational Chemistry Algorithms. *J. Comput. Chem.* **2008**, *29*, 839−845.

(297) Zapata, F.; Ridder, L.; Hidding, J.; Jacob, C. R.; Infante, I.; Visscher, L. QMflows: ATool Kit for Interoperable Parallel Workflows in Quantum Chemistry. *J. Chem. Inf. Model.* **2019**, *59*, 3191−3197.

(298) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chem. Rev.* **2021**, *121*, 9816−9872.

(299) Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in PyTorch. *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*; 2017.

(300) Bradbury, J.; Frostig, R.; Hawkins, P.; Johnson, M. J.; Leary, C.; Maclaurin, D.; Necula, G.; Paszke, A.; VanderPlas, J.; Wanderman-Milne, S.; Zhang, Q. *JAX: composable transformations of Python+NumPy programs*; 2018. http://github.com/google/jax (accessed 2023-08-25).

(301) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. PySCF: the Python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1340.

(302) Sun, Q.; et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, No. 024109.

(303) Kowalski, K.; et al. From NWChem to NWChemEx: Evolving with the Computational Chemistry Landscape. *Chem. Rev.* **2021**, *121*, 4962−4998.

(304) Smith, D. G. P SI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.

(305) Smith, D. G.; et al. P si 4N um P y: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theory Comput.* **2018**, *14*, 3504−3511.

(306) Olsen, J. M. H.; et al. Dalton Project: A Python platform for molecular- and electronic-structure simulations of complex systems. *J. Chem. Phys.* **2020**, *152*, No. 214115.

(307) Rehn, D. R.; Rinkevicius, Z.; Herbst, M. F.; Li, X.; Scheurer, M.; Brand, M.; Dempwolff, A. L.; Brumboiu, I. E.; Fransson, T.; Dreuw, A.; Norman, P. Gator: APython-driven program for spectroscopy simulations using correlated wave functions. *WIREs Computational Molecular Science* **2021**, *11*, No. e1528.

(308) Zhang, X.; Chan, G. K.-L. Differentiable quantum chemistry with PySCF for molecules and materials at the mean-field level and beyond. *J. Chem. Phys.* **2022**, *157*, No. 204801.

(309) Kasim, M. F.; Lehtola, S.; Vinko, S. M. DQC: A Python program package for differentiable quantum chemistry. *J. Chem. Phys.* **2022**, *156*, 084801.

(310) Arrazola, J. M. et al. Differentiable quantum computational chemistry with PennyLane. 2021, arXiv:2111.09967. *arXiv Preprint.* https://arxiv.org/abs/2111.09967 (accessed 2023-08-25).

(311) Li, L.; Hoyer, S.; Pederson, R.; Sun, R.; Cubuk, E. D.; Riley, P.; Burke, K. Kohn-Sham Equations as Regularizer: Building Prior Knowledge into Machine-Learned Physics. *Phys. Rev. Lett.* **2021**, *126*, 036401.

(312) Kasim, M. F.; Vinko, S. M. Learning the exchange-correlation functional from nature with fully differentiable density functional theory. *Phys. Rev. Lett.* **2021**, *127*, 126403.

(313) Abbott, A. S.; Abbott, B. Z.; Turney, J. M.; Schaefer, H. F. Arbitrary-Order Derivatives of Quantum Chemical Methods via Automatic Differentiation. *J. Phys. Chem. Lett.* **2021**, *12*, 3232−3239.

(314) Xie, H.; Liu, J. G.; Wang, L. Automatic differentiation of dominant eigensolver and its applications in quantum physics. *Phys. Rev. B* **2020**, *101*, 245139.

(315) Liao, H. J.; Liu, J. G.; Wang, L.; Xiang, T. Differentiable Programming Tensor Networks. *Physical Review X* **2019**, *9*, 031041.

(316) Tamayo-Mendoza, T.; Kreisbeck, C.; Lindh, R.; Aspuru-Guzik, A. Automatic Differentiation in Quantum Chemistry with Applications to Fully Variational Hartree−Fock. *ACS Central Science* **2018**, *4*, 559−566.

(317) *Research Software Alliance.* https://www.researchsoft.org (accessed 2023-08-25).

(318) *Better Scientific Software.* https://bssw.io/ (accessed 2023-08-25).

(319) *US Research Software Sustainability Institute.* https://urssi.us/ (accessed 2023-08-25).

(320) *Molecular Sciences Software Institute.* https://molssi.org/ (accessed 2023-08-25).