Methods and software for electronic structure based simulations of chemistry and materials

Volker Blum

Department of Mechanical Engineering & Materials Science, Duke University, Durham, NC, U.S.A.

Theresa L. Windus

Department of Chemistry, Iowa State University, Chemical and Biological Sciences Division, Ames National Laboratory, Ames, IA, U.S.A.

Electronic structure methods form the core of today's quantitative understanding of chemistry and materials science by simulations from the atomic scale upwards. The appeal of the field rests upon the fact that quantum mechanics provides an, in principle, exact mathematical prescription to predictively simulate any phenomenon of relevance that is related to the composition of matter as we know it. However, any practical computational implementations must necessarily be approximate, since the exact quantum mechanical formalism scales as a combinatorial explosion with problem size. The result is a rich field with immense existing capabilities, but one that is not even close to being finished. New computational capabilities and methodological developments continue to open up new scientific vistas to the science community.

This roadmap summarizes theoretical and methodological progress and current challenges in methods and software for electronic structure methods for materials science and chemistry, as implemented on current and foreseeable future computers. While not exhaustive, topics covered range from foundations of practical approximations to current software technology challenges and use cases in industry. Emphasis is on pathways to practical implementation of these concepts in current software and hardware environments, making them accessible to a broad community of scientists. At the outset, it may be important to clarify the terms "methods" and "software" – which are distinct but, in practice, inevitably intertwined. By "method," we refer to the level of formal, i.e., theoretical approximation that can be applied to make a particular phenomenon tractable (e.g., a particular density functional approximation in electronic structure theory or a particular way of treating atomic nuclei as classical or quantum particles). In contrast, "software" refers to the actual implementation as coded on a computer, whether in the form of a standalone code or as part of a larger software package. Frequently, the effectiveness of a given mathematical approximation (i.e., method) is determined by the degree to which it can be made applicable to real-world problems on an existing computer. This roadmap therefore reviews both “methods and software”, with an emphasis on practical solutions that are developed – or, in the words of Dirac (1929) “should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.” [1] Remarkably, even 95 years after Dirac’s paper, this summary remains accurate in the development of software and methods for simulations in chemistry, physics and materials science on the most up to date computational hardware.

Each section was contributed by a team of leaders in the field who were asked to contribute their ideas on four areas:

(i) Status of the Field,
(ii) Current and future challenges, and
(iii) Advances in Science and Technology to Meet Challenges.
(iv) Concluding Remarks

As one can imagine, trying to capture all of these in a very short section is challenging. For further reading, key references are included to foundational articles, current developments, and reviews. The sections are also meant to provide insights into the new challenges in the field for future stakeholders in these areas.

While the different sections provide insights into the specific subfields of electronic structure, several concepts showed up in multiple sections:

- We are at a crossroads where great achievements have been accomplished, but where innovations in new methods and algorithms for more accurate calculations to address more challenging chemical challenges in a shorter period of time are on the horizons.
- Interoperability of different methods and software were seen as solutions toward addressing complex scientific challenges. For example, embedding in its many different forms was discussed multiple times.
- Changes in hardware and software languages, while enabling technologies, create turbulence and reformulations of methods and algorithms. Reengineering our thinking and approaches is perhaps the only constant in the field.
- Electronic structure developers are exploring ways to use all of the new computing platforms including GPUs, cloud computing, machine learning, and quantum computing to accelerate solutions. Each of these has its own challenges that need to be overcome for electronic structure method implementations.
- Best practices in software engineering are seen as cornerstones to addressing many of the software challenges faced by the community. In particular, separation of concerns (more modular programming), development of abstract programming interfaces to allow for easier use and interoperability, usage of packaging tools, and provisions for support, documentation, and tutorials were seen as essential.
- The complexity of the software is such that multiple investigators are developing automated derivation and implementation tools to facilitate rapid development of methods and algorithms.

- Software usability in a manner that allows complex workflows for high throughput simulations or for ”trivial” parallelization is necessary for complex science.
- In a related point, verification, validation, and reproducibility are becoming even more important and solutions such as workflow tools may help with this.
- Another challenge in electronic structure methods is a better systematic understanding of errors in methods to facilitate collaborations with experimentalists.
- Training of the next generation of computational chemists and material scientists for new and evolving programming models, software paradigms, mathematical foundations, and electronic structure methods is challenging. However, it is key to making new advances in the field.
- Career paths for computational scientists who have the essential software engineering skills must be developed in a more open environment and recognized by metrics other than publications.

Overall, it is an exciting time to be a developer of electronic structure methods and there are many promising future directions to pursue. We hope that this roadmap provides not only windows, but actually a doorway into this future, for those who wish to shape it.

Acknowledgements

TLW acknowledges support from the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration, and by the U. S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Division of Chemical Sciences, Biological Sciences, and Geosciences Division at the Ames National Laboratory, which is operated by Iowa State University under contract No. DE-AC02-07CH11358. VB acknowledges support by the National Science Foundation (NSF) under award No. 1450280. Any opinions, findings, and conclusions or recommendations expressed here are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] P.A.M. Dirac. Quantum mechanics of many-electron systems. *Proc. Roy. Soc. London A*, 123:714–733, 1929.

Density Functional Theory: Formalism and Current Directions

Volker Blum

Thomas Lord Department of Mechanical Engineering and Materials Science and
Department of Chemistry, Duke University, Durham, NC, USA

Susi Lehtola

Department of Chemistry, University of Helsinki, P.O. Box 55, FI-00014 University of
Helsinki, Finland

Status

Density-functional theory (DFT),[1] particularly Kohn–Sham [2] (DFT) and generalized Kohn–Sham DFT, is responsible for the bulk of today's electronic structure applications. Approximate DFT methods provide sufficient accuracy for many applications in materials science and chemistry at highly competitive computational cost compared to higher-level, more accurate approximations to the Schrödinger or Dirac equations.

This chapter focuses on ground state DFT; time-dependent DFT is addressed in a later chapter of this Roadmap. Ground state DFT rests upon the Hohenberg–Kohn theorem of 1964 and the more rigorous Levy–Lieb demonstration [3, 4] that the ground-state energy E_{tot} can be expressed uniquely as a functional of the density $n(\mathbf{r})$, i.e., the probability density of finding a particle (usually an electron) at point $\mathbf{r}$. In the vast majority of cases, practical applications of DFT assume the Born–Oppenheimer approximation, separating the treatment of the nuclear coordinates $\{\mathbf{R}_I\}$ from the electronic ones. The result is the Born–Oppenheimer ground-state total energy $E_0[n(\mathbf{r})]$ as expressed in the shaded box in Fig. 1, where $n(\mathbf{r})$ is now the ground-state electron density.

The strength of the Kohn–Sham formalism of DFT is that the minimization of E_{tot} can be formulated as a problem of finding the states of auxiliary non-interacting particles moving in an effective field, allowing one to obtain the single-particle kinetic energy $T_s[n(\mathbf{r})]$ and thus E_{tot} and $n(\mathbf{r})$ by iterating over a set of straightforward three-dimensional partial differential equations. Orbital-free DFT methods relying on explicit kinetic energy functionals $T[n(\mathbf{r})]$ are faster than Kohn–Sham DFT that employs $T_s[n(\mathbf{r})]$. Although the development of orbital-free methods continues, they are arguably still less accurate and thus more restricted in their application space than Kohn–Sham DFT.

In Kohn–Sham DFT, the shaded expression for E_{tot} in Fig. 1 contains only one term

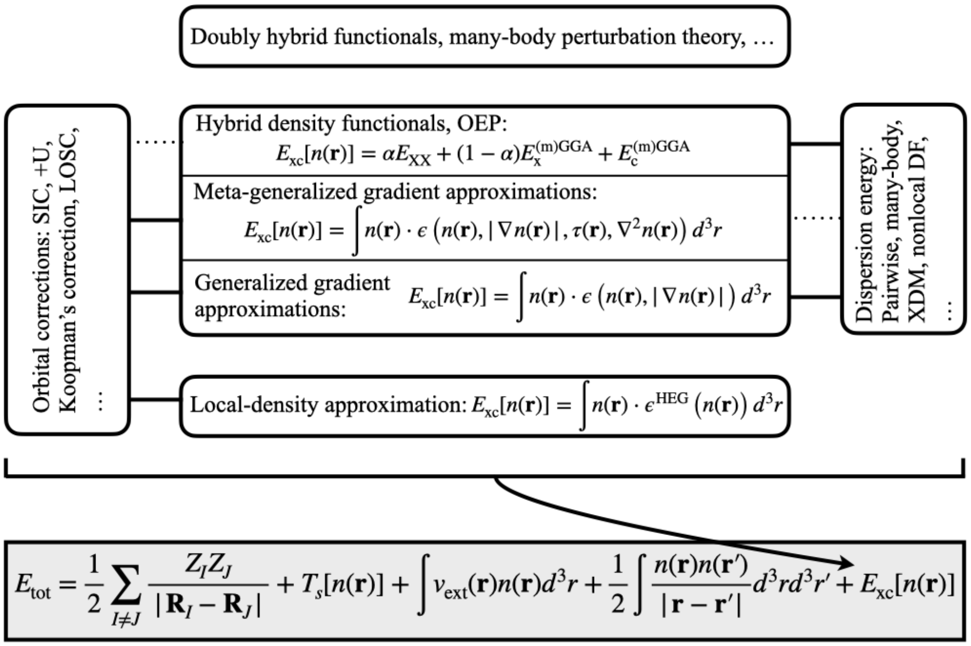


Figure 1. Kohn–Sham DFT expression for E_{tot} (shaded box) and hierarchy of mathematical forms of exchange-correlation functionals $E_{\mathrm{xc}}[n(\mathbf{r})]$. HEG: Homogeneous electron gas, SIC: Self-interaction correction, LOSC: Local-orbital scaling correction, OEP: Optimized effective potential, XDM: Exchange-dipole moment, DF: Density functional, (m)GGA: (meta-)generalized gradient approximation, XX: Exact exchange. Depictions of this hierarchy are now established in the literature (e.g., Ref. [5]); in the present version, the central box highlights the three arguably most productive levels of theory (GGA, mGGA, hybrid DFT).

that is unknown in its exact form, i.e., the exchange-correlation (xc) energy functional $E_{\mathrm{xc}}[n(\mathrm{r})]$. As a result, $E_{\mathrm{xc}}[n(\mathrm{r})]$ must be approximated. An established pathway to construct practical density functionals is to rely on known constraints and norms [6], but establishing the "best" parameterization of density functionals remains an active work in progress [6] and the optimum choice almost certainly depends on the specific application sought. A hierarchy of the most common types of approximations to $E_{\mathrm{xc}}[n(\mathrm{r})]$ found in the literature today is also shown in Fig. 1.

The simplest approximate form of E_{xc} is the local-density approximation (LDA), introduced in Refs. [1, 2], which relies on the known exchange-correlation energy density $\epsilon^{\mathrm{HEG}}(n(\mathbf{r}))$ of the homogeneous electron gas (HEG). Already the LDA accounts for the early successes of DFT: properties of metals, simple semiconductors, and other systems are captured with remarkable accuracy with the simple, parameter-free approximation. A key insight enabling these applications was the theory's extension to finite temperature via fractional occupations of the Kohn–Sham auxiliary orbitals [7].

Next, the central box in Fig. 1 shows the three most widely used approximations in production DFT calculations today: First, generalized-gradient approximations [8] (GGAs), which introduce a dependence on the density gradient $|\nabla n(\mathbf{r})|$; second, hybrid density functionals, which reintroduce a fraction of the "exact exchange" (XX) term of the Hartree–Fock equations [9]; and third, meta-GGA (mGGA) functionals, which add further derivative terms related to the Kohn–Sham density—the single-particle kinetic energy density $\tau(\mathbf{r})$ [10] and/or the density Laplacian $\nabla^2 n(\mathbf{r})$; these terms arise naturally from a Taylor expansion of the Kohn–Sham functional [11], giving rise to the well-known Jacob's ladder of DFT.[12]

One major challenge in DFT is that while some properties follow a clear pattern along Jacob's ladder—for instance, thermochemistry results tend to improve going from LDAs to GGAs, and from GGAs to meta-GGAs—the same does not apply to all properties. A related issue is that the accuracy of functionals is not necessarily transferable from one property to another. There is a huge number of GGA and meta-GGA functionals derived or fitted in dissimilar fashions, and different GGA (or meta-GGA) functionals can sometimes predict significantly different properties. A concise discussion of such issues is provided, e.g., in Ref. [6]. In applications, a practical approach to navigating the landscape of different functionals is therefore to benchmark results from a given functional against sound, experimentally obtained reference values for similar problem classes in order to ensure reliable results.

Despite the advances made in functional development (see below), the density functionals most widely employed by practitioners are thus often determined by the immense body of experience that exists regarding their successes and limitations—e.g., that of the PBE GGA functional [13] in materials science (overall, tremendously successful but known to overestimate equilibrium lattice parameters of solids by ∼1-2%) or of the B3LYP hybrid functional [14] in quantum chemistry (even though much more accurate functionals are nowadays available [15]). Nevertheless, many important advances have since been made on several fronts, as we will briefly summarize.

A critical advance was the realization that GGA/meta-GGA/hybrid functionals do not capture long-range dispersion interactions, which spawned multiple successful schemes to add the missing dispersion interactions to E_{xc} in either classical or density functional form (right box in Fig. 1). While weak compared to primary bonds, dispersion effects accumulate in large systems and can be critical determinants of the structure of large molecules, molecular assemblies or molecular crystals.

The issue that the (auxiliary) Kohn–Sham orbitals are not immune to self-interaction and incorrect occupation in strongly correlated systems has spawned a host of schemes including self-interaction corrections (SICs), correction schemes to restore Koopmans' theorem, the so-called +U parameterizations to shift orbital energies based on projectors and effective occupations, or, more generally, local-orbital scaling corrections to ensure the analytically correct linear behavior of $E_{\mathrm{tot}}(N)$ between integer electron counts N (left box in Fig. 1). A related issue is the systematic underestimation of energy band gaps by LDA and GGA functionals, which stems from this deviation from linearity.

Another route to improve upon the accuracy of "mainstream" GGA/meta-GGA-hybrid

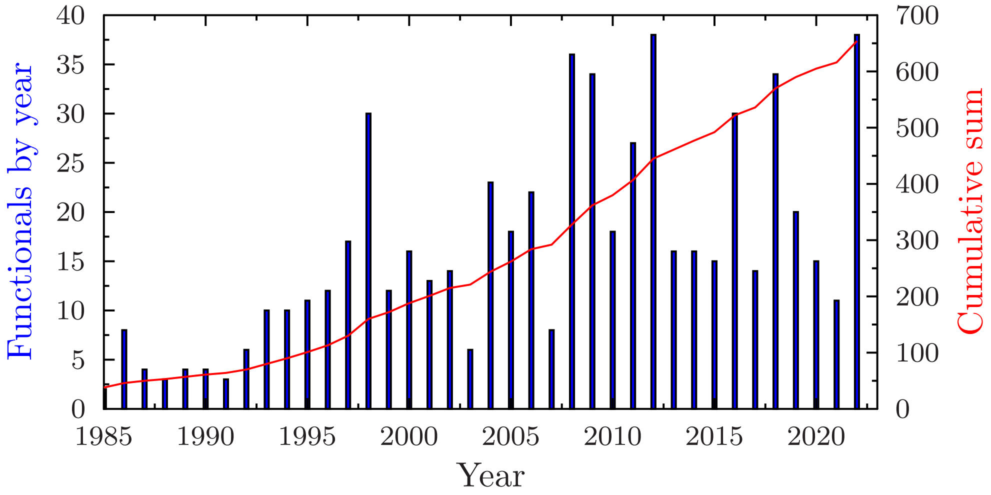


Figure 2. Statistics on the number of functionals by year of (latest) publication and the cumulative number of functionals based on the functionals implemented in Libxc version 6.2.0 [17].

group of functionals is to introduce more costly analytical approximations from the many-body realm, in terms of "doubly-hybrid" functionals, the random-phase approximation, or more general variants; however, doing so results in a significant increase of the computational cost. Local hybrids that determine the local fraction of exact exchange from the electron density offer a further avenue that has not yet become mainstream.

Finally, a recent pathway is to realize that the existing, computationally affordable approximations to DFT are not fully analytically derived, and to turn to machine learning (ML) the functional instead [16]. The realm of ML functionals is poised to expand greatly in coming years, already because access to ML technology and tools is nowadays widespread.

Current and Future Challenges

The open nature of the form of $E_{\mathrm{xc}}[n(\mathbf{r})]$ continues to encourage many developments trying to go beyond the state of the art. Beyond the obvious challenge of approaching "the exact functional" for all conceivable chemical situations, a host of technical challenges remain.

The number of proposed density functionals has increased strongly in the past several decades, as captured by the publication dates of functionals included in the Libxc library [17] shown in Fig. 2, for instance. Even though the overwhelming majority of these functionals do not become widely used, the proliferation of functionals causes problems of its own, as the implementation of "all" functionals in any given density functional program is a constantly and rapidly moving target. Moreover, many publications on new density functionals do not provide sufficient information to make the functionals fully reproducible. It is therefore not guaranteed that the implementations of a given functional in different programs afford mutually reproducible results.

To be useful for a broad range of applications, functionals—and computational implementations thereof—need to provide a number of higher derivatives. Applications to geometry optimization of excited states within the time-dependent DFT approach require fourth derivatives; machine learning of functionals for such properties will require even higher ones. This poses both theoretical and practical challenges: not all functional forms afford such derivatives, and the question on the efficient implementation of the large number of derivative terms is a difficult one, given that many functionals—especially machine learned ones—are complicated and heavily parameterized, resulting in a large number of contributions to each derivative term. A further issue is that the numerical well-behavedness of the functionals has not been given adequate attention by the developer community [18]. Because functionals are supposed to represent universal physics, they should be be transferable between systems, as well as numerical methods. However, functionals designed for small Gaussian basis sets may fail to work in a more flexible basis set such as finite elements [19], raising questions on whether they really afford *ab initio* approaches.

Returning to the challenge of approaching the "exact functional," one major current area of concern (for the ground state) is "strong correlation". Most strikingly, there is no guarantee that the occupation of the auxiliary Kohn–Sham orbitals (determined by their single-particle energies) matches the density of the actual many-electron wave function, when density functional approximations are employed. In some situations, a solution can be found by explicitly including the underlying physical degrees of freedom in the functional itself (e.g., separate spin densities in spin-polarized systems with localized states or the superconducting order parameter in DFT of superconductivity and its extensions[20]). Similarly, mapping the density and energy associated with a truly multi-determinantal many-electron wave function poses ongoing challenges. As one example, density functional studies of f-electron systems are hampered by this issue: although progress for some f-electron systems is being made [21], benchmark density functional studies involving f-electron elements with two or more holes or electrons in their filled f shell (e.g., Pr, Nd, or Pm) remain rare. Finding adequate density functional approximations for such systems remains an open area.

Advances in Science and Technology to Meet Challenges

A continually fruitful direction of functional development is to incorporate exact, analytical limits in the construction of practical density functionals. Already the PBE functional and earlier GGAs were motivated by satisfying known constraints, nevertheless leaving enough freedom for a proliferation of derived functionals. The addition of only one ($\tau(\mathbf{r})$) or two ($\tau(\mathbf{r})$ and $\nabla^2 n(\mathbf{r})$) additional functions of the density as ingredients in meta-GGAs seems like a simple thing, but the construction of such functionals that are both numerically robust and satisfy additional constraints is arguably not yet finished [18, 19].

Machine learning density functionals from higher-level data seems an obvious idea, but this raises the questions: Which higher level, and data on what systems? Determining sufficiently accurate many-electron wave functions for functional construction is itself a challenge for many types of chemical systems. For solids, for example, coupled-cluster theories even at the level of singles and doubles are extremely challenging, whereas molecular chemistry has clear examples of multireference systems where even the much more demanding level of triples is not enough. Quantum Monte Carlo (QMC) derived wave functions may offer a practical alternative but it is not yet clear how other aspects of this technique (and, in fact, the remaining inaccuracies of QMC wave functions) will impact their utility for future density functional development. The choice of the training data is also problematic. The ranking of functionals in any static benchmark is often determined by a small number of poorest performing systems [22]. As a result, the arising rankings are extremely sensitive to the exclusion or inclusion of few systems [22]. This suggests that ML functionals are also sensitive to the training dataset.

In the relativistic theory necessary to capture many interesting phenomena in heavier-element solids, it turns out that even classical electrodynamics is not satisfied with the density alone as a variable. Instead, the scalar electron density and the three-dimensional current density should appear on equal footing to conform to Lorentz covariance, and this observation should also be reflected in the relativistic density functional (in fact, similar developments may be incorporated in non-relativistic functionals as well). Developing appropriate current-density functionals is an open area at present in non-relativistic, scalar relativistic, as well as relativistic calculations, with relativistic calculations typically still relying on functionals borrowed from scalar- or non-relativistic constructions.

Finally, multicomponent density functional theory (e.g., considering density functional constructions of the nuclear density in addition to the electron density) is an active area that is gaining steam, as the quantum nature of the nuclei may be introduced in an efficient manner with such approaches, as discussed in a later section of this Roadmap.

As the above directions suggest, the considerable software engineering challenge involved in keeping implementations up to date with the rapidly increasing number of published functionals can be expected to continue into the future. Such challenges can be met by reusable, modular open source software libraries. Libxc [17] currently enables reliable reproduction of results across some 40 software packages by allowing the exact same (numerical implementation of a) density functional to be used regardless of the basis set or numerical approximation employed to solve the Kohn–Sham equations in these programs, eliminating the possibility of implementation-specific differences or bugs between the functionals in these programs. Libxc thereby greatly improves the cross-program reproducibility of computational results. The existence of Libxc also significantly simplifies the problem of integrating new density functionals into packages: in most cases, introducing new functionals requires simply linking to the newest version of Libxc.

Because fourth derivatives with respect to all density functional ingredients are necessary for some applications, and as many functional forms are prohibitively complicated to allow manual implementations, the use of automated approaches to form the derivatives is a practical necessity. Various methods may be used to compute such derivatives: in addition to the use of computer algebra systems to generate the derivatives as employed in Libxc [17], automated differentiation as originally proposed in the XCFun library [23] can also be employed; however, the computational cost for the latter approach is higher, as the derivatives need to be redetermined at every evaluation instead of being pregenerated and compiled into efficient binaries.

Concluding Remarks

Density functional theory and the early successes of simple density functional approximations such as the LDA, various GGAs and hybrid functionals have revolutionized our ability to simulate materials and molecules with a degree of generality that was simply inaccessible by empirical parameterizations. Despite the successes of the robust approximate many-body quantum mechanics baked into the framework of Kohn–Sham and generalized Kohn–Sham DFT, subtle energy differences encountered in the correct energy hierarchy of low-lying structures and even some qualitative failures still elude us. Much exciting physics and chemistry is thereby still left to be covered by applications of and new developments in DFT.

Acknowledgements

SL thanks the Academy of Finland for financial support under project numbers 350282 and 353749.

References

[1] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[2] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[3] Mel Levy. Electron densities in search of Hamiltonians. *Phys. Rev. A*, 26:1200–1208, 1982.

[4] Elliott H. Lieb. Density functionals for Coulomb systems. *Int. J. Quantum Chem.*, 24:243–277, 1983.

[5] Reinhard J. Maurer, Christoph Freysoldt, Anthony M. Reilly, Jan Gerit Brandenburg, Oliver T. Hofmann, Torbjörn Björkman, Sébastien Lebègue, and Alexandre Tkatchenko. Advances in density-functional calculations for materials modeling. *Annu. Rev. Mater. Res.*, 49(1):1–30, 2019.

[6] Aaron D. Kaplan, Mel Levy, and John P. Perdew. The predictive power of exact constraints and appropriate norms in density functional theory. *Annu. Rev. Phys. Chem.*, 74(1):193–218, 2023.

[7] N. David Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, 137:A1441–A1443, 1965.

[8] David C. Langreth and M. J. Mehl. Easily implementable nonlocal exchange-correlation energy functional. *Phys. Rev. Lett.*, 47:446–450, 1981.

[9] Axel D. Becke. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.*, 98:1372–1377, 1993.

[10] A. D. Becke and M. R. Roussel. Exchange holes in inhomogeneous systems: A coordinate-space model. *Phys. Rev. A*, 39:3761–3767, 1989.

[11] A. D. Becke. Hartree–Fock exchange energy of an inhomogeneous electron gas. *Int. J. Quantum Chem.*, 23:1915–1922, 1983.

[12] John P. Perdew and Karla Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.*, 577:1–20, 2001.

[13] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[14] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98:11623–11627, 1994.

[15] Narbe Mardirossian and Martin Head-Gordon. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.*, 115:2315–2372, 2017.

[16] John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, and Kieron Burke. Finding Density Functionals with Machine Learning. *Phys. Rev. Lett.*, 108:253002, 2012.

[17] Susi Lehtola, Conrad Steigemann, Micael J. T. Oliveira, and Miguel A. L. Marques. Recent developments in LIBXC—a comprehensive library of functionals for density functional theory. *SoftwareX*, 7:1–5, 2018.

[18] Susi Lehtola and Miguel A. L. Marques. Many recent density functionals are numerically ill-behaved. *J. Chem. Phys.*, 157:174114, 2022.

[19] Susi Lehtola. Meta-GGA density functional calculations on atoms with spherically symmetric densities in the finite element formalism. *J. Chem. Theory Comput.*, 19:2502–2517, 2023.

[20] A. Davydov, A. Sanna, C. Pellegrini, J. K. Dewhurst, S. Sharma, and E. K. U. Gross. Ab initio theory of plasmonic superconductivity within the Eliashberg and density-functional formalisms. *Phys. Rev. B*, 102:214508, 2020.

[21] Ruiqi Zhang, Bahadur Singh, Christopher Lane, Jamin Kidd, Yubo Zhang, Bernardo Barbiellini, Robert S. Markiewicz, Arun Bansil, and Jianwei Sun. Critical role of magnetic moments in heavy-fermion materials: Revisiting SmB_6. *Phys. Rev. B*, 105:195134, 2022.

[22] Thomas Weymuth and Markus Reiher. The transferability limits of static benchmarks. *Phys. Chem. Chem. Phys.*, 24:14692–14698, 2022.

[23] Ulf Ekström, Lucas Visscher, Radovan Bast, Andreas J. Thorvaldsen, and Kenneth Ruud. Arbitrary-order density functional response theory from automatic differentiation. *J. Chem. Theory Comput.*, 6:1971–1980, 2010.

IOP Roadmap: Density functional methods - implementation, challenges, successes

Luigi Genovese
Univ. Grenoble Alpes, MEM, L_Sim, F-38000 Grenoble, France

Paolo Giannozzi
Department of Mathematics, Computer Science, and Physics, University of Udine, I-33100 Udine, Italy

David B. Williams-Young
Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States of America

Status

Beginning in the 1980's, the introduction and development of reliable computational methods for the solution of Kohn-Sham equations, within both plane waves and localized basis sets, has made of Density Functional Theory (DFT) an exceedingly useful tool for the simulation of molecular and materials properties [1, 2]. More recently, further theoretical and computational advances, in conjunction with commensurate increases in computational power, have considerably extended the scope of what can be computed using DFT methods. We mention in particular: first-principle molecular dynamics, the modern theory of polarizability, density functional perturbation theory and improvements in functionals (hybrid, meta-GGA, Hubbard-corrected, to name a few). DFT simulations are by now routinely performed for systems containing $O(100-1000)$ atoms and for time scales of tens of ps, very often using advanced functionals beyond the simple GGA. In addition to a wealth of useful results, these methods provide an excellent starting point for studying excited states within many-body perturbation theory. These successes and the availability of efficient software implementations capable of leveraging the latest advances in modern high-performance computing have made DFT the *de facto* standard tool for studying large scale quantum systems and a "must-have" capability in the modern electronic structure ecosystem.

We are now undergoing a scientific and technological revolution in computer simulation that will enable the discovery of new materials and the understanding of complex processes in condensed matter, chemistry, and biology. This simulation revolution stands on three

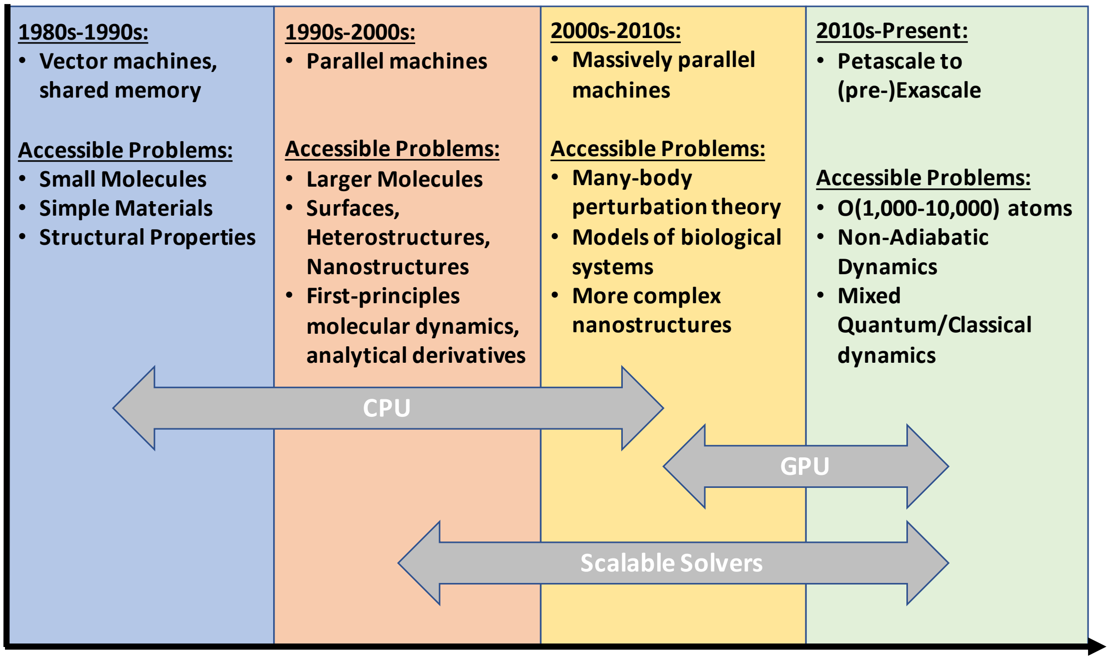


Figure 1. Brief graphical depiction of the history and status of DFT implementation, capabilities and scope.

pillars: (1) The design of clever algorithms which permit the computation of properties and the simulation of processes of ever growing complexity in ever more realistic operating conditions; (2) The availability of larger computing resources, exploiting increasingly powerful and complex processors and accelerators (e.g. GPUs), and (3) the development of increasingly flexible and portable software implementations capable of implementing (1), while optimising the performance for (2). The concurrent evolution of algorithms and codes has proceeded steadily, having and continuing to require a continuous and extensive re-engineering of the code-base for DFT methods. In this report, we briefly examine a number of pressing implementation and software challenges associated with the development of the next generation of DFT software and highlight a number of recent successes.

Current and Future Challenges

DFT simulations are still limited in terms of size of systems and scales of time that can be realistically accessed and simulated. A correct treatment of strongly correlated and of open-shell systems is still problematic, with results whose quality and reliability is difficult to assess. Advanced functionals often carry a sizable computational overhead and suffer from numerical stability problems. Finally, the ability to perform calculations including non-adiabatic effects and excited-state molecular dynamics is strongly desired for dealing e.g. with photo-physics.

GPU accelerators have revolutionized DFT software and science efforts while introducing a number of software development challenges [3]. The starkest departure has been in modern hardware diversity which has largely precluded single-source development models in favor of vendor-specific software solutions, introducing a sustainability crisis for performance portability in development efforts. This increased computational power has also exposed previously less-apparent computational bottlenecks, such as data movement and communication, which has led to significant challenges in the development of scalable solvers for (post-)Kohn-Sham methods. These challenges are likely to be exacerbated as we endeavour into (post-GPU) energy-efficient computing technologies. We are approaching a threshold for DFT development efforts which will require significant advances in software, algorithms and hardware to overcome.

As summarized in a past review coauthored by one of us [2], a fundamental challenge in atomistic simulations, especially with DFT codes, is their demand for expertise, encompassing both scientific and technical knowledge, for effective utilization. Originally developed by and for scientists within the same community, these codes often prioritize scientific functionality over user-friendliness, graphical interfaces, or collaborative tools. Addressing these aspects is crucial for fostering collaborations across diverse communities. Moreover, conducting atomistic simulations entails the utilization of substantial high-performance computing (HPC) resources. While such resources are commonplace in academia, their availability is not guaranteed in the industrial sector, where best practices can vary among academic communities. “By taking the most advanced codes for atomistic simulations and lowering their adoption barriers, it will be possible to ’democratize’ atomistic simulations and to open them up to a much broader community.” [2]

Advances in Science and Technology to Meet Challenges

Most advances in electronic structure theory will likely take place in ”beyond-DFT” calculations. Incremental improvements are however to be expected also for ”pure” DFT calculations, in particular in the field of advanced functionals: faster and more robust implementations, providing better and more reliable results. The usefulness and wide adoption of those improvements will depend upon the availability of a portable and maintainable implementation. It will also be crucial to extend the work on validation and verification, performed in recent years for simple properties and GGA [4], to more advanced property calculations.

While DFT simulations involving thousands of atoms are routinely performed, aspirations involving millions atoms remain out-of-reach, and will require improvements in software and algorithmic scalability to achieve. Recently, efforts to address performance portability and solver scalability through modular development models have been explored [5, 6, 7]. Although often requiring significant developer effort in targeting new architectures, these efforts have represented a departure from the monolithic development models of old, and offer clear pathways for future extensibility. Recent years have also seen explorations into low-precision computing, leading to the surpassing of the exaflop barrier for a DFT

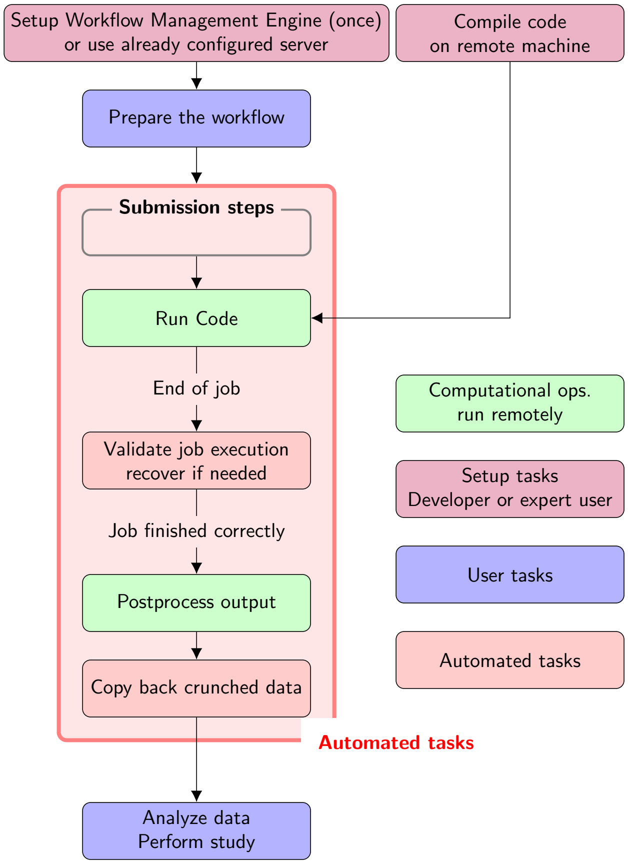


Figure 2. Depiction of how a workflow-management engine can be employed for the scheduling of the operations on a remote HPC machine.

application [8]. These developments must be extended as they will play a critical role in the scalable, energy-efficient computing future.

As the capacity for simulating larger systems grows, insights generated by communities accustomed to dealing with smaller systems, such as the electronic structure community, can be extended to other scientific domains. In this context, solutions like Software-as-a-Service (SaaS) platforms emerge as promising avenues to expedite research in nanoscale systems. SaaS integrates cutting-edge simulation codes, predefined workflows, user-friendly high-level tools, an intuitive collaborative interface, and adaptable cloud-based computing resources, streamlining the utilization of atomistic simulations.

Among the tools gaining popularity in recent years, Jupyter notebooks stand out, enabling the creation of reproducible scientific workflows. These notebooks consolidate pre-processing, calculation execution, and post-processing/analysis of results, offering a unified platform. Numerous codes from diverse scientific communities contribute tools that seamlessly interact with these notebooks. The concept of a ”separation of concerns” facilitates viewing codes as a service, wherein a client part—potentially Python-based—and a server part collectively execute computations, utilizing local or high-performance computing

(HPC) resources (refer to Fig. 2). As experiments scale up in complexity, such high-level tools become imperative. They serve as essential aids for end users, enabling them to effectively manage all the elements required to run, comprehend, and replicate experiments.

Concluding Remarks

Over the years, DFT methods development has reached a level of maturity which enables it to be employed as the basis for novel investigation directions that, thanks to synergistic progress in high-performance computing and software development strategies, were unfeasible even some years ago. In light of the many research pathways which have been enabled by these efforts, there remain a number of pressing software, algorithmic, and methodological challenges which must be addressed in the years to come. Such a future will also certainly include an increasing amount of interdisciplinary collaborations, necessitating the continuance of dissemination activities to bring DFT calculations to communities of non-specialists. Although we still have a long road ahead to achieve our aspirations, recent progress and successes in DFT methodology, implementation, and outreach efforts, together with growing participation and collaboration within the DFT software community, indicate a promising future.

Acknowledgements

PG and LG acknowledge support by the EU through the MaX Centre of Excellence for HPC applications (Project No. 824143). DWY acknowledges support from the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

References

[1] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, 87:897–923, Aug 2015.

[2] William Dawson, Augustin Degomme, Martina Stella, Takahito Nakajima, Laura E. Ratcliff, and Luigi Genovese. Density functional theory calculations of large systems: Interplay between fragments, observables, and computational complexity. *WIREs Computational Molecular Science*, 12(3):e1574, 2022.

[3] Mark S. Gordon and Theresa L. Windus. Editorial: Modern architectures and their impact on electronic structure theory. *Chemical Reviews*, 120(17):9015–9020, 2020. PMID: 32900196.

[4] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E. Castelli, Stewart J. Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, Marcin Dułak, Olle Eriksson, José A. Flores-Livas, Kevin F. Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, Oscar Grånäs, E. K. U. Gross, Andris Gulans, François Gygi, D. R. Hamann, Phil J. Hasnip, N. A. W. Holzwarth, Diana Iuşan, Dominik B. Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, Emine Küçükbenli, Yaroslav O. Kvashnin, Inka L. M.

Locht, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, Lars Nordström, Taisuke Ozaki, Lorenzo Paulatto, Chris J. Pickard, Ward Poelmans, Matt I. J. Probert, Keith Refson, Manuel Richter, Gian-Marco Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, Patrik Thunström, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J. van Setten, Veronique Van Speybroeck, John M. Wills, Jonathan R. Yates, Guo-Xu Zhang, and Stefaan Cottenier. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280):10.1126/science.aad3000, 2016.

[5] David B. Williams-Young, Abhishek Bagusetty, Wibe A. de Jong, Douglas Doerfler, Hubertus J.J. van Dam, Álvaro Vázquez-Mayagoitia, Theresa L. Windus, and Chao Yang. Achieving performance portability in gaussian basis set density functional theory on accelerator based architectures in nwchemex. *Parallel Computing*, 108:102829, 2021.

[6] Victor Wen zhe Yu, Carmen Campos, William Dawson, Alberto García, Ville Havu, Ben Hourahine, William P. Huhn, Mathias Jacquelin, Weile Jia, Murat Keçeli, Raul Laasner, Yingzhou Li, Lin Lin, Jianfeng Lu, Jonathan Moussa, Jose E. Roman, Álvaro Vázquez-Mayagoitia, Chao Yang, and Volker Blum. Elsi — an open infrastructure for electronic structure solvers. *Computer Physics Communications*, 256:107459, 2020.

[7] Carlo Cavazzoni, Fabio Affinito, Uliana Alekseeva, Claudia Cardoso, Augustin Degomme, Pietro Delugas, Andrea Ferretti, Alberto Garcia, Luigi Genovese, Paolo Giannozzi, Anton Kozhevnikov, Ivan Marri, Stephan Mohr, and Daniel Wortmann. First report on code profiling and bottleneck identification, structured plan of forward activities. Deliverable D4.2 of the H2020 CoE MAX. EC grant agreement no: 824143, CINECA, Bologna, Italy (2019). Available at `max-centre.eu/project-repository`.

[8] Robert Schade, Tobias Kenter, Hossam Elgabarty, Michael Lass, Thomas D Kühne, and Christian Plessl. Breaking the exascale barrier for the electronic structure problem in ab-initio molecular dynamics. *arXiv preprint arXiv:2205.12182*, 2022.

Green's function based many-body perturbation theory

Vojtěch Vlček[1], **Xinguo Ren**[2], **and Lucia Reining**[3]

[1]Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, USA and Department of Materials, University of California, Santa Barbara, California 93106, USA

[2]Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

[3]LSI, CNRS, CEA/DRF/IRAMIS, École Polytechnique, Institut Polytechnique de Paris, F-91120 Palaiseau, France

Status

Ground state properties of materials, such as the charge density or total energy, are today routinely described for many materials using Density Functional Theory (DFT). However, properties involving electronic excitations are more easily expressed as functionals of interacting Green's functions (GFs) that play a major role in the description and analysis of the response of materials to electromagnetic radiation or beams of charged particles. This response is the focus of our interest here: it governs absorption or electron energy loss spectroscopy (EELS), direct and inverse photoemission, inelastic x-ray scattering, and scanning tunneling spectroscopy, to name a few experimental techniques, and it is responsible for things such as stopping power, radiation damage, or light-induced phase transitions. In essence, the aforementioned measurements can be formally expressed in terms of one- and two-body GFs [1, 2, 3] capturing the propaagation of individual quasiparticles (QPs). Calculating those GFs is, however, still a challenge. In principle, the one-body GF can be formulated in terms of self-energy: an effective energy-dependent potential governing the QP propagation and representing all many-body interactions. The self-energy plays a role analogous to the exchange-correlation (xc) potential in the Kohn-Sham (KS) equations.

For weakly and moderately correlated systems, many-body perturbation theory (MBPT) is one popular way to approximate the self-energy and hence, GFs. The most frequent choice is to expand the self-energy in terms of the screened Coulomb interaction W and stop at first order: this leads to the widely used GW approximation [1]. It was initially employed to describe QP excitations in extended systems; indeed, the GW approximation has been extremely successful in describing the band gaps and band structures [2] including complex and composite materials. This flexibility is largely due to the fact that screening is explicitly taken into account, at variance with more approximate approaches, such as, e.g. hybrid functionals. In recent years, there has also been considerable interest in applying GW to finite systems where its performance can be benchmarked against quantum chemistry methods [4, 5]. For small molecules with a large energy gap, one instead typically expands the self-energy or the GF in terms of the bare Coulomb interaction. For example, truncating such an expansion at the second order leads to the second Born approximation. Further, to go beyond GW, vertex corrections can be added to the self-energy, or one can express the GF in terms of a cumulant expansion, efficiently describing excitations beyond the QPs, in particular, plasmon satellites [6].

The one-body GF describes electron addition and removal in materials. Analogously, two-body GFs can describe rich phenomena involving two-particle excitations. In particular, the particle-hole sector of the two-body GF encodes information about optical absorption and EELS. A key equation that captures the two-particle interactions is the Bethe-Salpeter equation (BSE), analogous to the Dyson equation for the one-body GF [2, 3]. The most common approximation for the electron-hole BSE builds upon GW and uses a QP approximation [7]. Today, GW+BSE is the state-of-the-art approach for determining optical spectra and exciton binding energies in extended materials.

Finally, one should mention that MBPT can also accommodate the coupling of the electronic charge degrees of freedom to other collective modes, in particular, phonons [8] and magnons [9].

Current and Future Challenges

GF-based MBPT is today one of the most widely used approaches for first principles electronic structure and spectroscopy calculations. Despite their success, perturbative GF methods are also encountering significant challenges that pose limits to their applications in condensed matter physics and materials science, both computationally and conceptually. Some of these are briefly discussed below as points on a roadmap to future research directions.

Choice of ingredients or self-consistency. MBPT calculations are often used in a non self-consistent way on top of a mean-field (KS of Hartree-Fock) calculation. The reason is twofold: the relatively high computational cost and the fact that the fully self-consistent evaluation of a low-order expression may lead to unsatisfactory results. On the other hand, the absence of self-consistency makes the results starting-point dependent and of low quality when the starting charge density and density matrix are poor.

For example, for many years, GW was supposed to fail for materials with localized d- or f-electrons, but this was often due to a poor description of charge localization. It was later shown that GW performed on top of a suitable starting point or in a partially self-consistent way allows one to describe many of those materials [10]. Self-consistency is also important for the simulation of the real-time dynamics of externally perturbed systems, where the fulfillment of sum rules may be particularly critical. On the other hand, self-consistency is not always possible computationally, and it does not always improve the results. Therefore, GW and beyond calculations still face the issue of the appropriate choice of GF and the screened interaction to be used to build the self-energy, and more theoretical arguments, model results, and computational benchmarks are needed to settle this question.

Computational challenge Besides the conceptual hurdles, the methodology is also limited by practical considerations. GF calculations come with a relatively high computational cost. Indeed, the bottleneck of a straightforward implementation of GW is the calculation of the full dynamical screened interaction W, which scales typically as N_{el}^4 with the number of electrons (or higher for methods beyond GW). More generally, contrary to DFT, the GF approaches deal with electronic transitions rather than individual states and, as such, also involve the spectrum of empty states. This calls for the use of clever basis sets. Moreover, when one is interested in spectra instead of numbers such as total energy, details do not integrate out, resulting in a need for a dense sampling of the Brillouin zone. Further, memory problems arise, particularly when two or more particle correlator problems are solved, e.g., in the BSE. Iterative inversion schemes help to avoid such bottlenecks, but often, there is a price to pay concerning the analysis of the results.

Precision and reproducibility of results. GW and related approaches have met great success for the qualitative correction of the band gap, starting from the KS eigenvalue gap. Today, quantitatively reliable results are required, which necessitates well-established standards. Work is needed along several lines, including:

(i) the treatment of time or frequency. The GW self-energy is a product in time or convolution in frequency space. While in principle equivalent, evaluation of Green's function and self-energy in frequency and time domains require numerical treatments that address distinct forms of instabilities (e.g., low-rank approximation of operators based on energy/frequency cutoffs or finite propagation time in operator/correlator evolution). A different problem is encountered in a finite temperature formalism that operates on the imaginary frequency axis: extracting observables of interest requires transformation to the real-frequency axis, leading to numerical difficulties.

(ii) Another source of approximation is related to the *solution of the QP problem that constitutes a non-linear eigenvalue problem*, which has been commonly either approximated (e.g., by linear extrapolation of the self-energy) or neglected (e.g., using the static limit in the electron-hole interaction kernel in BSE which avoids a two-frequency equation).

(iii) To address the *time evolution* of highly excited systems, non-equilibrium techniques, and numerically stable time-evolution techniques for both one and two-body GF are required.

(iv) *Different implementations* of GF methods address the above-mentioned problems distinctly. Moreover, GW calculations contain several *convergence parameters* that are partially interlinked. This is the case, for example, for the very slow convergence with the number of empty states that is also related to the size of the plane wave basis and the size of the response matrix. This makes it difficult to obtain reliable results, especially for newcomers. In practice, this impedes the reproducibility of computational results, and validation/verification has typically focused merely on a few single QP levels (e.g., fundamental gaps or individual quasiparticle energies). The development of novel functionalities, e.g., formulation and implementation of total energy and its gradients or the detailed analysis of multi-quasiparticle signatures in theoretical spectra, however, requires a set of robust and transferable numerical methods. Only recently, thorough comparisons of distinct implementations and various convergence parameters have been performed [4] across a wide variety of codes, underlying the effort in validation, verification, and reproducibility of tools employing MBPT. Clearly, more collective initiatives are needed to compare different approximations and implementations, and also to establish detailed and reliable workflows.

Extension of the MBPT formalism to encompass additional aspects. Calculations capturing relativistic effects, in particular spin-orbit coupling (SOC), are gaining interest. Relativistic effects are sometimes included *via* pseudopotentials, and SOC is often incorporated at the mean-field (DFT) level, to which QP energy corrections are added. More advanced calculations include relativistic effects also in the evaluation of MBPT [11, 12], but much still remains to be explored. Another important topic is temperature effects. Most calculations are done at vanishing lattice temperature, whereas at room temperature, for example, absorption spectra may have significantly different peak positions and spectral shapes.

A significantly more complex task is the description of an interacting system out-of-equilibrium. In the treatment of explicitly time-dependent problems, besides the computational cost associated with the self-energy evaluation, the scaling further increases with the overall simulation time (typically as $O(N_t^3)$, where N_t is the number of timesteps). [13]. This is because such simulations also require explicitly evaluating GFs and the self-energy as functions of two times and including memory effects. Additional problems appear due to the numerical instability of the time evolution. As a result, the non-equilibrium GF simulations are mostly employed in studies of model systems, and only recently, simulations of realistic systems became possible using approximations, in particular for memory effects [14, 15].

Diagrams beyond, or alternative to, GW. In GW, electronic correlations are limited to charge linear response coupled to electron addition or removal and treated in an approximate way that is correct to first order in W. There is no unique and well-established way yet to go beyond GW, and of course, the straightforward addition of higher-order terms leads to a strong increase of the computational cost [16, 17, 18, 19]. On the other hand, some important physical effects are clearly identified and can be linked to appropriate corrections. These are, in particular, the correction of a self-polarization error in GW that can be mitigated using second-order terms and/or approximations derived from time-dependent DFT [20], and the coupling of QPs to bosonic excitations, which leads to the emergence of satellite peaks and represents multi-particle excitations. This requires, in principle, vertex corrections that are of first and higher orders in W. In practice, satellites are most efficiently incorporated using cumulant GFs, both for electron addition and removal spectra and for electron-hole excitations [21, 22]. This also includes coupling to phonons.

Many questions remain to be explored, though: these include the convergence of MBPT, the choice of classes of diagrams and their resummation that is needed for a given problem, the rigorous combination of ingredients from MBPT and DFT, the fulfillment of exact constraints such as positivity of the spectral function, and the appropriate level of self-consistency in the vertex corrected approximations. The design of vertex corrections is based on the idea that in Hedin's formalism, the self-energy is expressed as GW plus additional (correcting) terms [23]. However, GW is not the most appropriate starting point for systems with strong particle-particle interaction effects. Alternatively, one may choose another starting point, e.g., a T-matrix expression, to which, e.g., screening corrections are then added [24]. Some works using T-matrix self-energies in first-principles calculations have been carried out for molecules and solids, but there is still a choice of classes of diagrams to be made and the combination with screening is not obvious [25]. Moreover, the computational cost of a full implementation is higher than that of GW calculations. Nevertheless, including such an alternative to GW in the toolbox of *ab initio* MBPT calculations may open the way to describe materials and properties that were considered to be out of reach of MBPT beforehand. Finally, it may be appropriate to explicitly address three-body and higher-order correlation functions to gain access to complex QPs such as trions and coupled exciton-electron excitations.

Combination with other approaches. The low-order approximations in MBPT are often not sufficient for situations close to degeneracy, where correlation determines the physics. Even in situations of more modest correlation strength, the accuracy of GW may not be sufficient to meet the practical needs. This happens, for example, concerning band offsets, effective masses, or total energies. In this case, combinations with other approaches may overcome the problem. One historically explored route is to simulate vertex corrections by using the exchange-correlation kernel of time-dependent DFT, which leads to a screened interaction that is more appropriate for interacting fermions than the screened interaction W of the GW approximation and, in particular, reduces the self-screening problem of GW. For strongly correlated materials, a combination with the dynamical mean-field theory (DMFT) can be envisaged. This is also a GF approach but implicitly includes all site-local skeleton diagrams in the self-energy [26]. Especially concerning this combination with DMFT, new computational problems arise.

Despite these outstanding challenges, MBPT-based methods have become established even outside of their traditional field of computational condensed matter physics, and they are nowadays widely applied in fields such as materials science or chemistry, for questions of astrophysics or biological processes, matter under extreme conditions, disordered systems such as liquids, or complex processes such as in (photo-driven) catalysis, and also become explored in the context of quantum computing. This brings new hurdles and opportunities, including taking into account parameters of the environment such as temperature or experimental setups, and in any case, leads to a drastically increased complexity.

Advances in Science and Technology to Meet Challenges

The application of GF techniques has become more widespread and has found its way to becoming the *de facto* workhorse in materials and computational physics communities. On one hand, this is thanks to the increasing power and availability of large-scale HPC devices. Simultaneously, numerous algorithmic advances made the calculations computationally less expensive and numerically more reliable.

Multiple low-scaling algorithms emerged in the past decade, in particular in the context of equilibrium GW. In spirit, the methods aim to reduce the complexity via some form of sparse linear algebra, decreasing the prefactor and reducing the overall scaling to be quadratic with the number of electrons (in combination with localized bases and/or pair bases for the polarizability). For instance, one of the largest bottlenecks is associated with the calculations of the screened interactions, which, in turn, require computing the response functions. These steps have been optimized by numerical compression techniques (e.g., of response functions in a matrix form) and exploiting the low-rank operator structure. Such techniques have retained relatively high scaling $O(N^3 - N^4)$ for GW (or higher for the vertex-corrected methods), but significantly reduced the scaling prefactor, allowing simulations of large-scale systems [27]. Steady progress comes from improvements in aspects such as time/frequency Fourier transforms [28, 29] or the use of Resolution of Identity approximations [30]. Alternatively, random algorithms exploit the information redundancy via sampling of the single-particle states and lead to linear scaling for large-scale systems in GW [31] and beyond and for the BSE.

A separate problem arises for finite temperature formulations, in which the self-consistency is typically implemented in the imaginary time/frequency domain. However, many observables, such as the single-particle spectra, require real-frequency information. Analytic continuation techniques are constantly developed further [32] and succeed in determining the energies of well-defined QP states with sufficient accuracy, whereas it is still difficult to access the rest of the spectral function.

For non-equilibrium problems, the progress is largely lagging behind, but recent time-linear scaling methods pave the way for realistic simulations in this area. The explicit evolution of one and two-body GFs [33] and the application of model order reduction techniques have recently emerged as a powerful scheme for performing non-equilibrium simulations.

The combination of these techniques and their applications in a broader context represents a promising research direction. Furthermore, the next steps critically hinge upon the development of a transferable computational implementation that leverages new computational hardware, exhibits scalable parallelization and is GPU-ready for the most advanced HPC architectures.

Concluding Remarks

GF-based MBPT has yielded an important class of first-principles approaches for predicting materials' excited state and, sometimes, ground state properties. In particular, for extended weakly to moderately correlated systems, GF-based MBPT methods often provide predictive accuracy for QP band structures and excitation spectra, including exciton binding energies, for a large variety of materials. Such a performance is hardly achievable otherwise. Still, further developments of these methods in terms of their theoretical robustness and practical capability face significant conceptual and computational challenges. In this roadmap, we have briefly discussed several prominent research directions, as well as the necessary advances in science and technology to address these challenges. We would like to stress that on top of ongoing important initiatives, a joint effort of the community is needed to make the theoretical and computational choices unambiguous and to improve the reproducibility of the results. In spite of all the difficulties, thanks to the rapid theoretical and algorithmic developments, as well as more efficient and numerically stable implementations across different hardware platforms, we expect GF-based MBPT methods to become significantly more powerful and play an indispensable role in future first-principles computational studies of real materials.

Acknowledgements

We would like to thank Volker Blum, Theresa Windus, Dominika Zgid, and Steven Louie for the helpful discussions.

References

[1] L. Hedin. Phys. Rev., 139:A796, 1965.

[2] Lars Hedin. On correlation effects in electron spectroscopies and the GW approximation. Journal of Physics: Condensed Matter, 11(42):R489, 1999.

[3] Giovanni Onida, Lucia Reining, and Angel Rubio. Electronic excitations: density-functional versus many-body Green's function approaches. Rev. Mod. Phys., 74:601, 2002.

[4] Michiel J van Setten, Fabio Caruso, Sahar Sharifzadeh, Xinguo Ren, Matthias Scheffler, Fang Liu, Johannes Lischner, Lin Lin, Jack R Deslippe, Steven G Louie, et al. GW100: Benchmarking G_0W_0 for molecular systems. J. Chem. Theory Comput., 11(12):5665–5687, 2015.

[5] Dorothea Golze, Marc Dvorak, and Patrick Rinke. The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy. Frontiers in Chemistry, 7:377, 2019.

[6] Matteo Guzzo, Giovanna Lani, Francesco Sottile, Pina Romaniello, Matteo Gatti, Joshua J. Kas, John J. Rehr, Mathieu G. Silly, Fausto Sirotti, and Lucia Reining. Valence Electron Photoemission Spectrum of Semiconductors: Ab Initio Description of Multiple Satellites. Phys. Rev. Lett., 107:166401, Oct 2011.

[7] Michael Rohlfing and Steven G. Louie. Electron-hole excitations and optical spectra from first principles. Phys. Rev. B, 62:4927–4944, Aug 2000.

[8] Feliciano Giustino. Electron-phonon interactions from first principles. Rev. Mod. Phys., 89:015003, Feb 2017.

[9] Mathias C. T. D. Müller, Stefan Blügel, and Christoph Friedrich. Electron-magnon scattering in elementary ferromagnets from first principles: Lifetime broadening and band anomalies. Phys. Rev. B, 100:045130, Jul 2019.

[10] Mark van Schilfgaarde, Takao Kotani, and Sergey Faleev. Quasiparticle self-consistent GW theory. Phys. Rev. Lett., 96(22):226402, 2006.

[11] R Sakuma, C Friedrich, T Miyake, S Blügel, and F Aryasetiawan. GW calculations including spin-orbit coupling: Application to Hg chalcogenides. Phys. Rev. B, 84(8):085144, 2011.

[12] Chia-Nan Yeh, Avijit Shee, Qiming Sun, Emanuel Gull, and Dominika Zgid. Relativistic self-consistent GW: Exact two-component formalism with one-electron approximation for solids. Phys. Rev. B, 106(8):085121, 2022.

[13] Karsten Balzer and Michael Bonitz. Nonequilibrium Green's Functions Approach to Inhomogeneous Systems. Springer, 2012.

[14] Davide Sangalli, Stefano Dal Conte, Cristian Manzoni, Giulio Cerullo, and Andrea Marini. Nonequilibrium optical properties in semiconductors from first principles: A combined theoretical and experimental study of bulk silicon. Phys. Rev. B, 93(19):195205, 2016.

[15] Enrico Perfetto, Yaroslav Pavlyukh, and Gianluca Stefanucci. Real-Time GW: Toward an Ab Initio Description of the Ultrafast Carrier and Exciton Dynamics in Two-Dimensional Materials. Phys. Rev. Lett., 128(1):016801, 2022.

[16] Andreas Grüneis, Georg Kresse, Yoyo Hinuma, and Fumiyasu Oba. Ionization potentials of solids: the importance of vertex corrections. Phys. Rev. Lett., 112(9):096401, 2014.

[17] Andrey L Kutepov. Electronic structure of Na, K, Si, and LiF from self-consistent solution of Hedin's equations including vertex corrections. Phys. Rev. B, 94(15):155101, 2016.

[18] Vojtech Vlcek. Stochastic vertex corrections: Linear scaling methods for accurate quasiparticle energies. J. Chem. Theory Comput., 15(11):6254–6266, 2019.

[19] Yanyong Wang, Patrick Rinke, and Xinguo Ren. Assessing the $G_0W_0\Gamma_0^{(1)}$ Approach: Beyond G_0W_0 with Hedin's Full Second-Order Self-Energy Contribution. J. Chem. Theory Comput., 17:5140–5154, 2021.

[20] Wei Chen and Alfredo Pasquarello. Accurate band gaps of extended systems via efficient vertex corrections in GW. Phys. Rev. B, 92(4):041115, 2015.

[21] Jianqiang Sky Zhou, Lucia Reining, Alessandro Nicolaou, Azzedine Bendounan, Kari Ruotsalainen, Marco Vanzini, JJ Kas, JJ Rehr, Matthias Muntwiler, Vladimir N Strocov, et al. Unraveling intrinsic correlation effects with angle-resolved photoemission spectroscopy. Proc. Natl. Acad. Sci, 117(46):28596–28602, 2020.

[22] Pierluigi Cudazzo and Lucia Reining. Correlation satellites in optical and loss spectra. Phys. Rev. Res., 2(1):012032, 2020.

[23] Carlos Mejuto-Zaera and Vojtěch Vlček. Self-consistency in $GW\Gamma$ formalism leading to quasiparticle-quasiparticle couplings. Phys. Rev. B, 106(16):165129, 2022.

[24] Robert van Leeuwen, Nils Erik Dahlen, and Adrian Stan. Total energies from variational functionals of the Green function and the renormalized four-point vertex. Phys. Rev. B, 74:195105, Nov 2006.

[25] M Springer, F Aryasetiawan, and Krister Karlsson. First-principles T-matrix theory with application to the 6 eV satellite in Ni. Phys. Rev. Lett., 80(11):2389, 1998.

[26] Gabriel Kotliar, Sergej Y Savrasov, Kristjan Haule, Viktor S Oudovenko, O Parcollet, and CA Marianetti. Electronic structure calculations with dynamical mean-field theory. Rev. Mod. Phys., 78(3):865, 2006.

[27] Marco Govoni and Giulia Galli. Large scale gw calculations. Journal of Chemical Theory and Computation, 11(6):2680–2696, 2015. PMID: 26575564.

[28] Merzuk Kaltak, Jiří Klimeš, and Georg Kresse. Low scaling algorithms for the random phase approximation: Imaginary time and laplace transformations. J. Chem. Theory Comput., 10(6):2498–2507, 2014.

[29] Peitao Liu, Merzuk Kaltak, Ji ří Klimeš, and Georg Kresse. Cubic scaling gw: Towards fast quasiparticle calculations. Phys. Rev. B, 94:165109, Oct 2016.

[30] Jan Wilhelm, Dorothea Golze, Leopold Talirz, Jürg Hutter, and Carlo A. Pignedoli. Toward gw calculations on thousands of atoms. Phys. Rev. B, 9:306, 2018.

[31] Daniel Neuhauser, Yi Gao, Christopher Arntsen, Cyrus Karshenas, Eran Rabani, and Roi Baer. Breaking the theoretical scaling limit for predicting quasiparticle energies: The stochastic gw approach. Phys. Rev. Lett., 113:076402, Aug 2014.

[32] Jiani Fei, Chia-Nan Yeh, and Emanuel Gull. Nevanlinna analytical continuation. Phys. Rev. Lett.,

126:056402, Feb 2021.

[33] Niclas Schlünzen, Jan-Philip Joost, and Michael Bonitz. Achieving the scaling limit for nonequilibrium green functions simulations. Phys. Rev. Lett., 124:076601, Feb 2020.

IOP Roadmap: Wave-function theory approaches – explicit approaches to electron correlation

Roland Lindh
Department of Chemistry – BMC, Uppsala University, Uppsala, Sweden

Markus Reiher
Lab. f. Phys. Chem., ETH Zurich, Zurich, Switzerland

Edward F. Valeev
Department of Chemistry, Virginia Tech, Blacksburg, USA

Hans-Joachim Werner
Institute for Theoretical Chemistry, University of Stuttgart, Stuttgart, Germany

Status

Wave-function-based electronic structure theory (WFT) is concerned with the development of approximate quantum mechanical descriptions of electrons in molecules and materials that can be systematically improved towards the exact solution. Key physico-chemical quantities obtained in such approaches are, first and foremost, the energies of ground and excited stationary electronic states, the energy response to field and other perturbations, and state-to-state transition probabilities. Altogether these allow for first-principles prediction of molecular structures, chemical reactivity, and spectroscopy.

The key concept underlying all standard approaches in this field is the expansion of the many-electron wave function in terms of electronic configurations which are anti-symmetric with respect to the exchange of any two electrons (the spin–statistics theorem). The configurations can be either anti-symmetrized products (Slater determinants) of one-electron functions (molecular spin-orbitals, or spinors) or spin-adapted linear combinations of them (configuration state functions). The molecular orbitals are usually expanded in a basis of atom-centered Gaussian basis functions. The simplest approximation is to represent the wave function by a single Slater determinant and to optimize the orbitals by minimizing the energy (Hartree-Fock method, independent particle model). This forms the basis of molecular orbital theory. The Hartree-Fock approximation recovers the vast majority (99% or more) of the exact electronic energy. The remainder is denoted as the electron correlation energy. The magnitude of the correlation energy is (at least) of the same order as chemical energies (e.g.

reaction energies). It is therefore essential to recover 99% or more of the correlation energy in order to make quantitative predictions of chemical reactivity and molecular properties. Since the correlation energy is an extensive quantity, i.e. it is proportional to the molecular size, this is particularly difficult in calculations for large molecules.

The wave function expansion becomes exact for a complete determinantal expansion – full configuration interaction, FCI – in a complete basis of orbitals (a.k.a complete CI). Due to the factorial scaling of the number of determinants with the number of electrons FCI becomes impractical for more than ~ 20 electrons. Therefore, practical applications must deploy approximations by (a) truncating the expansion adaptively (e.g., selected CI) and/or systematically (e.g., truncated CI[1], many-body perturbation theory[2]), (b) parametrizing the coefficients in the complete expansion nonlinearly (e.g., coupled-cluster[2], tensor network methods[3]), or (c) using stochastic expansions (e.g., determinantal quantum Monte-Carlo[4, 5]). The rich phenomenology of the basic approximations techniques and their combinations is further compounded by algorithmic and computational innovations, thereby precluding even a brief enumeration of recent research directions.

Several well-established classes of methods of WFT have been turned into tools for the accurate determination of properties and energetics for small to medium sized molecular systems that are robust enough for use by nonspecialists. These tools are usually either based on a single determinant or a linear combination of determinants as a reference function, i.e., single- and multi-configuration reference methods, respectively. The single-reference coupled-cluster methods are the primary workhorse of WFT, capable of predicting chemical energy differences with kJ/mol accuracy for small systems[6]. However, they may fail when the wave function is dominated by more than one Slater determinant. This is for example the case when molecular bonds are stretched (such as at the transition states) or dissociated, for open-shell ground states (such as radicals or most transition metal compounds) and for almost all excited states. In such cases multi-reference methods are usually needed.

Practical application of even the simplest WFT methods still face two fundamental problems, the steep polynomial scaling of the computational cost with the number of electrons and the large and slowly decaying basis set errors. The slow basis set convergence results from the poor description of the wave function at short interelectronic distances by truncated Slater determinant expansions. This is due to the electron-electron cusps for $r_{ij} \to 0$, which cannot be described by products of spin-orbitals. For small molecules the basis set limit can be estimated using extrapolation approaches, but this is limited by the steep increase of the computational cost with basis set size. Another more satisfying approach is to include terms in the wavefunction that depend explicitly on the inter-electronic distances,[7] but these methods are significantly more complicated to implement and also need additional auxiliary basis sets.

These scaling problems with system size can be overcome by local correlation or fragmentation treatments, combined with explicit correlation approaches. With modern explicitly correlated local correlation methods it is currently possible to compute accurate energies for molecules with 100-200 atoms, and to reach chemical accuracy (below 1 kcal/mol) for relative energies (e.g. reaction energies, isomerization energies, conformational energy changes, or intermolecular interactions).[8] Still, the accuracy of such methods for large systems is difficult to assess. Local approximations and basis set incompleteness errors can be tested for medium sized molecules (up to about 30 atoms) by comparison with canonical methods, but it is uncertain how well the results can be extrapolated to much larger systems, in which other effects such as dispersion interactions become increasingly important. The errors of the energy are extensive, i.e. they increase with molecular size, and high accuracy of relative energies can only be achieved if large parts of the errors cancel in reactants and products. Fortunately, many chemical processes involve local changes in electronic structure, and such error cancellations seem to work well in most applications. The atomistic surroundings of such molecular transformations can then be modelled efficiently by environment embedding schemes.[9]

Current and Future Challenges

A severe limitation of wave function approaches, especially for large molecules, is the complexity of their mathematical formalisms and approximations. This complexity makes extension to higher excitation ranks and evaluation of energy derivatives with respect to nuclear coordinates or other perturbations technically challenging.

While single reference coupled-cluster methods can be used in a black-box manner and achieve high accuracy for ground state properties around the equilibrium structure qualitative failures are regularly observed in the simulation of deformative processes which involve the breaking and formation of bonds. Multi-configurational methods have been developed to address these and other problems of single-reference methods. However, these methods typically do not yet reach the accuracy of single-reference coupled-cluster methods.

Yet another challenge is that multi-reference WFT approaches typically require a high degree of expertise on behalf of the user. However, there are no fundamental limitations that would prevent a high degree of automatism in actual calculations, which would also make them less error prone. Moreover, computations that can be used in automated workflows (especially for high throughput virtual screening campaigns) require a high degree of robustness, which is particularly hard to achieve for multi-configurational schemes or composite methods that require different approaches for different electron-correlation regimes.

Despite the fact that a hierarchy of WFT approaches exists that allows to reduce errors systematically, this is limited by the extremely high cost of higher-order calculations. Therefore, the actual error in a specific calculation is mostly not known. Even if it is supposed to be small, it depends on the molecular system and the target application whether the error can be tolerated or not. So far, error assessment has been based on benchmarking, but intrinsic uncertainty quantification and error control for a specific calculation at hand will be a challenge and a key for predictive work in the future.[10] Essential to this task will be overcoming the known limitations of the traditional atomic-orbital-based numerical representations that support modern WFT; the use of alternative numerical representations (e.g., real space grids, finite elements, etc.) could greatly improve the ability to quantify and control the discretization errors of the WFT methods.

The modern set of wave function ansätze are more or less all exclusively based on determinant-based expansions. Alternative approaches could, for instance, utilize geminals (which describe two explicitly correlated electrons) as the building blocks; the unique advantages of geminal approaches range from compact description of certain types of strong electron correlation to supremely accurate calculations of few-body systems.[11]

So far, our focus has been on the electronic energies, on the associated Born-Oppenheimer surfaces, and couplings between them. Naturally, there are further challenges for electronic structure models such as (1) multi-component approaches that also consider quantum nuclei, photons, or polarons,[12] (2) response properties of large molecules with accurate wave functions, and (3) magnetic resonance parameters for relativistic heavy-atom molecules[13].

Finally, on the hardware side, we face severe challenges that range from compatibility and reproducibility issues due to software evolution (e.g., brought about by programming language and compiler development) to hardware constraints (such as disjoint memory spaces, shrinking memory and bandwidth budget per FLOP, increasing specialization of computing units, etc.).

Advances in Science and Technology to Meet Challenges

To address these challenges and realize the full potential of WFT for predictive molecular and materials simulation new scientific advances – concepts, algorithms, and computational infrastructure – will be needed. While it is futile to try to identify where the next great advances will occur, it is possible to bring such advances closer to reality by sustaining and accelerating the rate of scientific innovation in this field. We identify the following technological factors crucial for that.

- All fields of computational science have benefited crucially from the exponential increase of the classical computing power over the past 50 years. Continuation of the technological progress that sustained the evolution of classical computation platforms is needed to make wave function simulations even more affordable, both by reducing the time to solution and by shrinking the required electrical power budget.
- Recent trends in classical computing all make WFT ansätze far more difficult to program, and the hardware roadmaps suggest that these trends will continue. Thus, the adoption of GPUs in our field, even for established but especially for emerging methods, is poor. New tools/programming models are needed to make programming modern massively-parallel classical computers, with heterogeneous execution units (GPUs) and address spaces (clusters, distributed file systems), easier.

- In designing wave function methods we should take greater account for their fitness for hardware. A prominent example of such field-wide co-design is the recent developments of "quantum computing" algorithms suitable for execution on the universal quantum computing devices.[14, 15, 16] It is inevitable that such co-design will continue to be necessary, not only for the case of rapidly evolving quantum hardware but also for the modern and emerging classical devices.
- Greater use of domain-specific automation to make scientific innovation easier (e.g., automated derivation and implementation of complex wave function models and their responses correctly and efficiently), reduce the amount of code to maintain, and improve performance portability.

Concluding Remarks

Wave-function-based electronic structure theory is at the core of numerous endeavors in theoretical chemistry – such as quantum dynamics, classical dynamics, mechanistic explorations, property prediction, data generation for machine learning and physico-chemical modeling, reference data production, and so forth. By construction, its approaches contain very little bias and therefore allow for an universal applicability with an option for error assessment by comparison to results of increasing accuracy. It is for these reasons that the importance of their further development can hardly be overestimated. The successes of the past decades have clearly demonstrated their value and also uncovered the avenues to follow in the future, as outlined in this roadmap.

Acknowledgements

R.L. acknowledges the Swedish Research Council (VR, Grant No. 2020-03182). M.R. acknowledges continuous and very generous support by ETH Zurich over the past decades. E.F.V. acknowledges support of the U.S. Department of Energy via award DE-SC0022327.

References

[1] B. Huron, J. P. Malrieu, and P. Rancurel. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. J. Chem. Phys., 58(12):5745–5759, June 1973.

[2] Isaiah Shavitt and Rodney J. Bartlett. Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory. Cambridge University Press, first edition, August 2009.

[3] Alberto Baiardi and Markus Reiher. The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges. J. Chem. Phys., 152(4):040903, 2020.

[4] George H. Booth, Alex J. W. Thom, and Ali Alavi. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. J. Chem. Phys., 131(5):054106, 2009.

[5] Mario Motta and Shiwei Zhang. Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method. WIREs Comput. Mol. Sci., 8(5):e1364, 2018.

[6] Attila Tajti, Peter G Szalay, Attila G Császár, Mihály Kállay, Jürgen Gauss, Edward F Valeev, Bradley A Flowers, Juana Vázquez, and John F Stanton. HEAT: High accuracy extrapolated ab initio thermochemistry. J. Chem. Phys., 121(23):11599–11613, December 2004.

[7] W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev. R12 methods in explicitly correlated molecular electronic structure theory. Int. Rev. Phys. Chem., 25:427, 2006.

[8] Qianli Ma and Hans-Joachim Werner. Explicitly correlated local coupled-cluster methods using pair natural orbitals. WIREs Comput Mol Sci., 2018.

[9] Leighton O. Jones, Martín A. Mosquera, George C. Schatz, and Mark A. Ratner. Embedding Methods for Quantum Chemistry: Applications from Materials to Life Sciences. J. Am. Chem. Soc., 142(7):3281–3295, 2020.

[10] Markus Reiher. Molecule-specific Uncertainty Quantification in Quantum Chemical Studies. Isr. J. Chem., 62:e202100101, 2021.

[11] Sergiy Bubin, Michele Pavanello, Wei-Cheng Tung, Keeper L. Sharkey, and Ludwik Adamowicz. Born–Oppenheimer and Non-Born–Oppenheimer, Atomic and Molecular Calculations with Explicitly Correlated Gaussians. Chem. Rev., 113(1):36–79, 2013.

[12] Fabijan Pavošević, Tanner Culpitt, and Sharon Hammes-Schiffer. Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method. Chemical Reviews, 120(9):4222–4253, 2020.

[13] Jochen Autschbach. Relativistic calculations of magnetic resonance parameters: background and some recent developments. Phil. Trans. Roy. Soc. A, 372(2011):20120489, 2014.

[14] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. Rev. Mod. Phys., 92:015003, 2020.

[15] Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin-Lic Chan. Quantum Algorithms for Quantum Chemistry and Quantum Materials Science. Chem. Rev., 120(22):12685–12717, 2020.

[16] Hongbin Liu, Guang Hao Low, Damian S. Steiger, Thomas Häner, Markus Reiher, and Matthias Troyer. Prospects of quantum computing for molecular sciences. Mater. Theory, 6:11, 2022.

IOP Roadmap: Quantum Monte Carlo and stochastic electronic structure methods

Paul R. C. Kent

Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States of America

Claudia Filippi

MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Eran Rabani

Department of Chemistry, University of California and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States of America
The Raymond and Beverly Sackler Center of Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel

Alex J. W. Thom

Yusuf Hamied Department of Chemistry, Lensfield Road, CB2 1EW Cambridge, United Kingdom

Status

‡

By reformulating direct numerical approaches, stochastic methods greatly extend the complexity, accuracies, and scale that can be reached with many-body electronic structure and quantum chemical approaches, in exchange for introducing a controlled statistical error. Stochastic sampling changes both the power law scaling and the computational prefactors of methods as compared to using conventional numerical integration. This transformation can allow application of high-accuracy approaches to system sizes and phenomena that would otherwise be out of reach. The reformulations

---

‡ This manuscript has been authored in part by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

also permit more parallelizable implementations and easier use of supercomputers, reducing the time to solution.

In the case of Quantum Monte Carlo (QMC) algorithms,[1, 2] solutions of the full many-body Schrödinger equation are obtained statistically. An input trial wavefunction is typically constructed using the best available mean-field or many-body approach and then usually decorated with additional physics-motivated terms to describe additional electron correlation. It is then used either directly, as in variational Monte Carlo (VMC), or in a projection scheme to obtain a systematically better approximation, as in diffusion Monte Carlo (DMC), auxiliary field quantum Monte Carlo, and full configuration-interaction Monte Carlo[3] and stochastic coupled cluster theory[4]. The majority of these approaches are general and can be applied to metals, insulators, molecular systems, and to ground and select excited states within the same framework. Although exact treatments are possible in simple cases[5], in practice a fixed-node or phase approximation is employed to treat the Fermion sign-problem. Results are still highly accurate, but this approximation must be tested. While computationally expensive, system sizes similar to density functional theory (DFT) can be studied. e.g. already in 2016, TiO_2 phases with up to 1728 electron supercells were studied[6], and computational power has increased significantly since then. Therefore, a large range of scientific problems are within reach.

Stochastic orbital or vector techniques have recently been introduced to reduce the scaling of mean-field based approaches [7, 8, 9] as well as for many-body perturbation techniques,[10, 11, 12] and provide a framework for reducing algorithmic complexity and for facilitating efficient parallelization. While stochastic vector approaches share features with the aforementioned QMC techniques (as further discussed below), they differ by relying on approximate methods such as DFT and many-body perturbation theories, thereby targeting much larger system sizes.

Current and Future Challenges

We have identified four major challenges:

Reducing and controlling the statistical noise and biases: To extend the range of methods studied by these approaches and broaden their use, the computational costs must be reduced and the biases (approximations) reduced. This requires fundamental improvements in, e.g., the projection QMC methods, or improvements in the importance sampling used in stochastic vector approaches. And for any given method, the statistics and biases must be optimally and automatically controlled to minimize the overall computational cost for a desired accuracy. For QMC methods, a particular challenge is the consistent and reliably automatable determination of the trial wavefunction coefficients. Reduced biases – increased physical accuracy – generally result through use of more complex wavefunction forms with more coefficients, but their determination through stochastic optimization in turn becomes more difficult. Reducing the noise in stochastic vector techniques has so far relied on fragmentation and embedding[13], for open [14] and periodic boundary conditions [15, 16]. While the noise can be reduce by two orders of magnitude, automatization of the noise reduction schemes and the removal of bias remains a challenge.

Stochastic techniques for structural optimization and molecular dynamics: Forces computed with stochastic methods have an intrinsic statistical uncertainty in both direction and magnitude, and the magnitude may be non-zero with finite sampling even at structural minima. Therefore, conventional numerical approaches are not appropriate, and a naïve molecular dynamics will not conserve energy. Convergence must be handled delicately. While several approaches have recently been proposed for structural optimization [17, 18, 19] and for sampling the canonical distribution, they have yet to be widely used or demonstrate generality to large, low symmetry systems.

Stochastic embedding techniques and improved treatment of finite-size effects: Embedding techniques offer the accuracy of a fully many-body approach and improve scaling by partitioning the system to strong and weakly correlated regimes, but as-yet are little developed. For example, combining QMC and stochastic vector techniques, stochastic embedding techniques would offer improved scaling as well as circumvent approximations introduced in deterministic approaches.

Improved interoperability with other electronic structure methods: To-date the most important QMC results remain the exact calculations for the homogeneous electron gas[5]. The density-dependent energies were later parameterized in the local density approximation of DFT. Beside energies and densities, many-body methods have access numerous many-body quantities such as the exchange-correlation hole and two-body density matrices. In principle, these could be used to inform or validate the construction of computationally cheaper electronic structure methods, which can then be applied more widely.

Advances in Science and Technology to Meet Challenges

Reducing and controlling the statistical noise and biases: In QMC, improvements in the trial wavefunction leads to improved accuracy and reduced intrinsic variance/statistical cost. New forms of wavefunction and methodologies based on the developments seen in machine-learned force-fields provide an as-yet little explored route to achieve this. Modifications to the long-established QMC move generation algorithms could be derived to improve the importance sampling and overall statistical efficiencies. Crucially, improvements in the biases need to be achieved consistently between different systems so that energy and property differences are consistently improved. Similarly, in stochastic vector techniques, reference systems are used to reduce the statistical error but often introduce a bias. The most common reference system relies on fragmenting the system,[16] but the optimal choice of the fragments is still an open area of research and requires further developing more accurate schemes. Another notable hurdle is the automation of fragment identification, a task that could potentially benefit from the application of machine learning and neural networks. In all cases, a deeper understanding of the sources of error and origins in the statistical variance will aid the design of improved sampling schemes.

Stochastic techniques for structural optimization and molecular dynamics: Reliable and efficient structural optimization requires the development and deployment of algorithms that factor all statistical uncertainties to efficiently converge to the optimized structures. For dynamics, the requirement are stricter and must ensure conservation of the desired observables, such as energy, etc. One promising approach is based on the recent development of highly training-data efficient, “second generation” machine learned interatomic potentials [20]. Such approaches rely on training data (forces on the nuclei and energies) generated by first principle techniques. To date, DFT has been the main framework used to generate the training data, often restricted to a small sub-system due to the computational complexity.

Quantum Monte Carlo and stochastic vector techniques present more precise frameworks with lower computational complexity, making them well-suited for generating training data. Despite the statistical nature of the training procedures, the investigation of the impact of noise arising from force fluctuations calculated using QMC or stochastic vector techniques has been limited [21]. Therefore, the generalization of training steps must consider these statistical fluctuations, and the development of noise reduction schemes specifically tailored for training neural network force fields is imperative.

Stochastic embedding techniques and improved treatment of finite-size effects: Quantum Monte Carlo (QMC) provides a formally exact many-body framework, especially well-suited for strongly correlated systems, whereas stochastic vector techniques depend on approximations that prove effective for weakly correlated systems. Describing systems with mixed strong and weak correlations, especially in extended systems, poses a significant challenge. A promising avenue involves integrating QMC with stochastic vector techniques through quantum embedding methods.

Embedding methods often rely on many-body Green’s function approaches, but in principle should also be applicable to wavefunction based techniques. The fusion of stochastic vector techniques with QMC approaches promises a balance between high accuracy and low computational complexity, broadening the scope of problems amenable to first-principles analysis. Achieving this entails crafting an embedding framework, devising novel algorithms for integrating stochastic realms, and evaluating the accuracy and computational efficiency of these hybrid techniques.

Improved interoperability with other electronic structure methods: Here we believe that many of the necessary methods are in place for a bidirectional exchange between different classes of methods, primarily through observables other than the total energy, e.g. density matrices. However, computing many of the desired observables for a great many systems is computationally infeasible. A dialog with the broader electronic structure and quantum chemical communities is required on the preferred systems and quantities to ensure that the efforts are well targeted, making full use of the trends and uncertainties identified through computational materials and chemical databases.

Making use of new technology: While stochastic methods are often embarrassingly parallel, the inherent branching can lead to them not taking full advantage of parallel pipelines such as GPUs. While advances in compilation and appropriate languages are making such approaches easier to code, the relatively small user and developer bases have fewer resources to develop on these architectures, though significant increases in computational efficiency could result.

Concluding Remarks

Stochastic methods extend the reach of high-accuracy and many-body approaches, and are well suited to take advantage of the ongoing increases in available computational power. For greater scientific reach and wider adoption, technical improvements are desired to reduce both the statistical costs and the remaining biases and approximations in the algorithms. As the methods become more affordable and in some cases less artisanal to run, opportunities for both direct application and for validation or improvement of more scalable approaches are poised to greatly increase.

Acknowledgments

PRCK was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials.

References

[1] Federico Becca and Sandro Sorella. Quantum Monte Carlo Approaches for Correlated Systems, November 2017.

[2] Lucas K. Wagner and David M. Ceperley. Discovering correlated fermions using quantum Monte Carlo. *Reports on Progress in Physics*, 79(9):094501, 2016.

[3] George H. Booth, Alex J. W. Thom, and Ali Alavi. Fermion monte carlo without fixed nodes: A game of life, death and annihilation in slater determinant space. *J. Chem. Phys.*, 131:054106, 2009.

[4] Alex J. W. Thom. Stochastic coupled cluster theory. *Phys. Rev. Lett.*, 105:263004, 2010.

[5] D. M. Ceperley and B. J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Physical Review Letters*, 45(7):566, 1980.

[6] Ye Luo, Anouar Benali, Luke Shulenburger, Jaron T. Krogel, Olle Heinonen, and Paul R. C. Kent. Phase stability of TiO_2 polymorphs from diffusion Quantum Monte Carlo. *New Journal of Physics*, 18(11):113049, 2016.

[7] Roi Baer, Daniel Neuhauser, and Eran Rabani. Self-Averaging Stochastic Kohn-Sham Density-Functional Theory. *Physical Review Letters*, 111(10):106402, September 2013.

[8] Marcel D. Fabian, Ben Shpiro, Eran Rabani, Daniel Neuhauser, and Roi Baer. Stochastic density functional theory. *WIREs Computational Molecular Science*, 9(6):e1412, 2019.

[9] Roi Baer, Daniel Neuhauser, and Eran Rabani. Stochastic Vector Techniques in Ground-State Electronic Structure. *Annual Review of Physical Chemistry*, 73:255, 2022.

[10] Daniel Neuhauser, Yi Gao, Christopher Arntsen, Cyrus Karshenas, Eran Rabani, and Roi Baer. Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic *GW* Approach. *Physical Review Letters*, 113(7):076402, August 2014.

[11] Wenjie Dou, Tyler Y. Takeshita, Ming Chen, Roi Baer, Daniel Neuhauser, and Eran Rabani. Stochastic Resolution of Identity for Real-Time Second-Order Green's Function: Ionization Potential and Quasi-Particle Spectrum. *Journal of Chemical Theory and Computation*, 15(12):6703–6711, December 2019.

[12] Eran Rabani, Roi Baer, and Daniel Neuhauser. Time-dependent stochastic Bethe-Salpeter approach. *Physical Review B*, 91(23):235302, June 2015.

[13] D. Neuhauser, R. Baer, and E. Rabani. Embedded fragment stochastic density functional theory. *J. Chem. Phys.*, 141:041102, 2014.

[14] E. Arnon, E. Rabani, D. Neuhauser, and R. Baer. Equilibrium configurations of large nanostructures using embedded fragment stochastic density functional theory. *J. Chem. Phys.*, 146:224111, 2017.

[15] M. Chen, D. Neuhauser, R. Baer, and E. Rabani. Overlapped embedded fragment stochastic density functional theory for covalently bonded materials. *J. Chem. Phys.*, 150:034106, 2019.

[16] M. Chen, D. Neuhauser, R. Baer, and E. Rabani. Stochastic density functional theory: Real- and energy-space fragmentation for noise reduction. *J. Chem. Phys.*, 154:204108, 2021.

[17] Juha Tiihonen, Paul R. C. Kent, and Jaron T. Krogel. Surrogate Hessian accelerated structural optimization for stochastic electronic structure theories. *The Journal of Chemical Physics*, 156(5):054104, February 2022.

[18] Siyuan Chen and Shiwei Zhang. A structural optimization algorithm with stochastic forces and stresses. *arXiv:2204.12074 [cond-mat]*, April 2022. arXiv: 2204.12074.

[19] Jesse van Rhijn, Claudia Filippi, Stefania De Palo, and Saverio Moroni. Energy Derivatives in Real-Space Diffusion Monte Carlo. *Journal of Chemical Theory and Computation*, 18(1):118–123, January 2022.

[20] Oliver T. Unke, Stefan Chmiela, Huziel E. Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T. Schütt, Alexandre Tkatchenko, and Klaus-Robert Müller. Machine learning force fields. *Chem. Rev.*, 121:10142–10186, 2021.

[21] David M. Ceperley, Scott Jensen, Yubo Yang, Hongwei Niu, Carlo Pierleoni, and Markus Holzmann. Training models using forces computed by stochastic electronic structure methods, 2023.

IOP Roadmap: Heavy element relativity, spin-orbit physics, and magnetism

Gustav Bihlmayer, Stefan Blügel

Peter Grünberg Institute and Institute for Advanced Simulation, Forschungszentrum Jülich, Jülich, Germany

Lucas Visscher

Department of Chemistry and Pharmaceutical Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Status

In electronic structure theory it is quite common to ignore magnetism and consider only electric interactions. Likewise, one often dismisses the effects of the relativistic increase of mass at high electron velocities and the spin-orbit coupling (SOC). These three approximations simplify theory considerably as it suffices to solve the non-relativistic Schrödinger equation, rather than the more complete Dirac equation. For some important technological applications such a neglect can, however, lead to quantitatively or even qualitatively incorrect results. Three examples suffice to illustrate this point.

The first example is the accurate prediction of complexation and adsorption free energies of actinide species. These data are required to model transport properties of these species when evaluating the safety of long-term storage options for nuclear waste. Electronic structure calculations of such materials can only be done if relativistic effects are included from the outset as they dramatically change the relative energies of the s- and f-bands[1].

A second example concerns organic light emitting diodes (OLEDs). Due to use of phosphorescent emission[2], OLED technology has become one of the most energy efficient ways of creating colour displays. To further increase this efficiency, one needs to accurately model and mitigate all undesired energy quenching processes. Being able to model spin-orbit coupling is thereby essential.

A third example comes from the field of quantum materials. The energetically tiny spin-orbit coupling lifts degeneracies of electronic states and acts as emergent magnetic fields with important ramifications for the spin-polarization of the electronic structure in nonmagnetic solids (e.g., even in a relatively light compound like 3R-MoS_2 splittings induced by the Rashba effect reach 200 meV, see Fig. 1), the creation of topological matter [3] (e.g., topological insulators), spintronic functionalities, e.g., spin-orbit torque to manipulate the magnetization by electrical current, the emergence of orbital magnetic moments, or complex

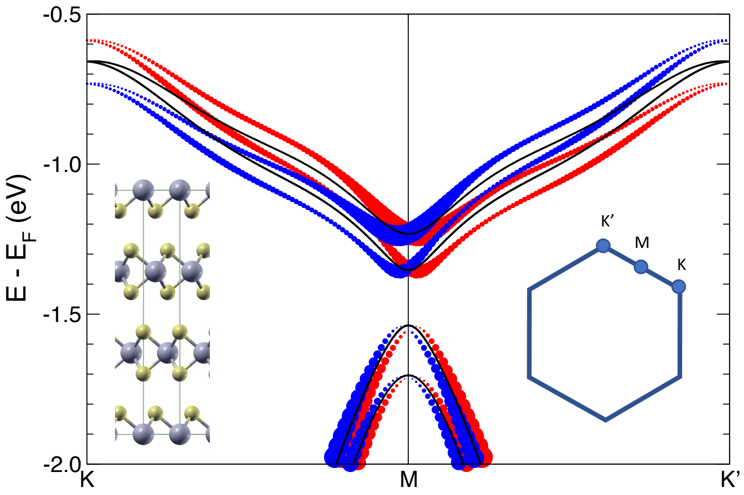


Figure 1. Spin-orbit splitting of the valence bands in 3R-MoS_2: Black lines show the band structure without spin-orbit coupling (SOC), red/blue dots indicate the spin polarization perpendicular to the layers (up/down) of the bands with SOC included. Insets show the structure (grey/yellow spheres: Mo/S) and the Brillouin zone.

magnetic interactions (e.g., Dzyaloshinskii-Moriya interaction) in magnets, which can lead to topologically protected noncollinear spin-textures and magnon excitations [4].

Today, scalar-relativistic approximations to the fully relativistic treatment are realised in many community electronic structure methods based on density functional theory (DFT) as well as in more advanced methods to treat electron correlation such as many-body perturbation theory with Hedin's *GW* approximation to the self-energy, coupled cluster, or density matrix renormalization group. The SOC is often included as perturbation, but also implementations treating the Dirac equation with magnetism in a mean field approximation are available, which goes back to early efforts in the mid sixties [5]. The treatment of non-collinear magnetism becomes increasingly available. While properties like the magnetic anisotropy or complex magnetic structures induced by relativistic interactions can often be well predicted, it should be noted that some tiny, but important effects like elemental bulk anisotropies or orbital polarizations still evade an accurate description [6].

Current and Future Challenges

Naturally, a relativistic DFT based on the Dirac equation rather than the Schrödinger equation seems ideal to address the topics mentioned above, but its formulation and application turns out to be quite challenging [7]. Some simplification can be gained by the reformulation of the Dirac equation as an equation for electrons (rather than for electrons and positrons) that is possible with the so-called exact two-component (X2C) approach [8]. In many cases relativistic corrections to electron-electron interactions are omitted, but they can relatively easily be included in many-body perturbation theory such as *GW* or in a mean-field theory such as DFT [9]. Nevertheless, it must be kept in mind that beyond mean-field relativistic theory is needed in many problems, e.g. for heavy transition metal compounds with strongly correlated electron systems [10] for which already a qualitatively correct description of the wave function calls for a multi-reference approach. We further note the study of time-dependent (TD) phenomena, where relativistic versions of TD-DFT have been developed [11], but where more advanced methods may be needed to fully describe finite temperature effects and magnetic disorder.

Another challenge arises from the complexity of the material models that need to be constructed. While heavy elements form the crucial and most difficult to model part of a material, they are typically surrounded by other materials that can be modelled sufficiently accurately with a cheaper approach. This calls for use of multilevel and multiscale approaches, but introduces a dependency on the adequacy of the partitioning of the system to be studied into essential and secondary regions and on the quality of treating the interface between regions that are described at different levels of theory.

Advances in Science and Technology to Meet Challenges

Methods that handle magnetism and other relativistic effects require much more data as the wave function models do not separate out the spin degrees of freedom. This is also unavoidable if the essence of the problem lies in strong coupling of spin and spatial parts of the wave function. In addition, many important quantities are tiny in size and require a high numerical resolution. New computer technology does help here, with the large memories and enormous processor counts available on modern GPU-based supercomputers it is possible to store and process the extra data needed to handle the more extended wave function models. The advent of quantum coprocessors will also help as they can be able to efficiently process the large active orbital spaces needed to model strongly correlated electrons.

Tackling the second problem mentioned above, the efficient construction of multiscale and multilevel methods, will require closer interaction of domain scientists with software engineers. The efficient storage and reuse of data is essential in such methods and while standards for simple data such as molecular and materials geometric structure have matured, this is not yet the case for electronic structure data such as (excitation) energies, electron (spin) densities, and molecular orbital coefficients for 2-component wave functions. Another important aspect is the parallelization of workflows, in many studies one may use conceptually trivial parallelization over studies of multiple materials or initial conditions that will benefit from automatically parallelizing workflow engines. Further developing these for the rather heterogenic compute systems (CPUs, GPUs, QPUs) that will emerge in the future can be viewed as important challenge.

Concluding Remarks

While much progress has been made in the past decades toward better understanding and modelling the implications of magnetic and other relativistic effects in chemistry and material sciences, more work is certainly needed. The available algorithms and their implementations are typically at least an order of magnitude more resource consuming than their non-relativistic counterparts and also the methods to visualize and analyze the resulting feature-rich data still need to mature.

Acknowledgements

G.B. and S.B. acknowledge financial support from the MaX Center of Excellence funded by the EU through the H2020- INFRAEDI-2018 (project: GA 824143).

References

[1] Pekka Pyykko. Relativistic effects in structural chemistry. *Chemical Reviews*, 88(3):563–594, 1988-05.

[2] Chihaya Adachi, Marc A. Baldo, Mark E. Thompson, and Stephen R. Forrest. Nearly 100% internal phosphorescence efficiency in an organic light-emitting device. *Journal of Applied Physics*, 90(10):5048–5051, 2001.

[3] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Rev. Mod. Phys.*, 83:1057–1110, Oct 2011.

[4] Anjan Soumyanarayanan, Nicolas Reyren, Albert Fert, and Christos Panagopoulos. Emergent phenomena induced by spin–orbit coupling at surfaces and interfaces. *Nature*, 539:509–517, 2016.

[5] T. L. Loucks. Relativistic electronic structure in crystals. I. Theory. *Phys. Rev.*, 139:A1333–A1337, Aug 1965.

[6] M. D. Stiles, S. V. Halilov, R. A. Hyman, and A. Zangwill. Spin-other-orbit interaction and magnetocrystalline anisotropy. *Phys. Rev. B*, 64:104430, Aug 2001.

[7] Reiner M. Dreizler and Eberhard Engel. Relativistic density functional theory. In Daniel Joubert, editor, *Density Functionals: Theory and Applications: Proceedings of the Tenth Chris Engelbrecht Summer School in Theoretical Physics Held at Meerensee, near Cape Town South Africa, 19–29 January 1997*, pages 147–189, Berlin, Heidelberg, 1998. Springer.

[8] W Kutzelnigg and WJ Liu. Quasirelativistic theory equivalent to fully relativistic theory. *The Journal of Chemical Physics*, 123(24):241102, 01 2005.

[9] Camilla Pellegrini, Tristan Müller, John K. Dewhurst, Sangeeta Sharma, Antonio Sanna, and Eberhard K. U. Gross. Density functional theory of magnetic dipolar interactions. *Phys. Rev. B*, 101:144401, Apr 2020.

[10] William Witczak-Krempa, Gang Chen, Yong Baek Kim, and Leon Balents. Correlated quantum phenomena in the strong spin-orbit regime. *Annual Review of Condensed Matter Physics*, 5(1):57–82, 2014.

[11] Michal Repisky, Lukas Konecny, Marius Kadek, Stanislav Komorovsky, Olga L. Malkin, Vladimir G. Malkin, and Kenneth Ruud. Excitation Energies from Real-Time Propagation of the Four-Component Dirac–Kohn–Sham Equation. *Journal of Chemical Theory and Computation*, 11(3):980–991, 2015.

IOP Roadmap: Semiempirical methods

Christoph Bannwarth
Institute of Physical Chemistry; RWTH Aachen University, Aachen, Germany

Ben Hourahine
SUPA, Department of Physics, University of Strathclyde, Glasgow, UK

Jonathan Moussa
Molecular Sciences Software Institute, Blacksburg, VA 24060, USA

Status

Semiempirical electronic structure methods reduce the cost of solving the many-body Schrödinger equation by simple models and approximate solutions and mitigate the resulting errors with parameters fitted to reference data, either from experiments or higher levels of theory. Typically, they use a minimal atomic orbital basis set, parameterized multi-center integral approximations, and mean-field calculations based on Hartree-Fock (HF) theory or density-functional theory (DFT). The semiempirical Hückel method for π electrons was proposed only a year after HF theory in 1931, and it inspired more general models based on the zero-differential overlap (ZDO) approximation in the 1950's. By the 1980's, this had been further refined into the neglect of diatomic differential overlap (NDDO) approximation and developed into popular thermochemistry models such as AM1 and PM3, which are implemented in the MOPAC program [1].

The popularity of DFT in the early 1990's shifted most semiempirical method development from minimal-basis models to semiempirical density functionals with fitted parameters, and the last few decades of development has produced hundreds of new semiempirical density functionals but relatively few new minimal-basis models. While large-basis DFT calculations are typically more accurate than minimal-basis models, this accuracy comes at a roughly three orders of magnitude increase in computational cost. Semiempirical methods were also able to reduce the cost of DFT, and extended Hückel theory (EHT) from the 1960's inspired the development of density functional tight binding (DFTB) in the late 1990's, as implemented in software such as DFTB+ [2].

Even with steady growth in computing power, scientists still have limited computational budgets and often seek lower-cost methods, particularly when the size or number of systems is large or the required time to solution is short. Currently, semiempirical models are mainly used for explorations of conformational and chemical spaces and interactive quantum

Model name	PM7	GFN2-xTB	DFTB3/3OB-D4
Model family	MNDO	GFN	DFTB3
Parent software	MOPAC	xTB	DFTB+
Primary output	heat of formation	total energy	total energy
Reference data	heats, geometries, dipole moments, ionization potentials	geometries, forces, vibrational frequencies, non-covalent energies	energies, geometries, vibrational frequencies, barrier heights
Elemental coverage	H-La, Lu-Bi	H-Rn	H, C-F, Na, Mg, Zn, P-Cl, K, Ca, Br, I
Orbital type	orthogonal	non-orthogonal	non-orthogonal
Hopping integrals	Wolfsberg-Helmholz approximation of Slater-type orbitals	generalized Wolfsberg-Helmholz approximation of STO-nG orbitals	tabulated Slater-Koster matrix elements from atomic and diatomic DFT calculations
Coulomb integrals	NDDO approximation	multipole approximation	monopole approximation
Exchange energy	Fock exchange	density functional	density functional
Dispersion energy	short-range DH+ model	self-consistent D4 model	self-consistent D4 model

Table 1. Basic features and approximations of several popular semiempirical models.

mechanical studies, which continue to drive semiempirical model development. As shown in Table 1, the GFN family of models in the recent xTB program [3] combines the DFTB formalism with some design elements from EHT and atomic multipole expansions up to quadrupoles. There is also progress towards more unified software, with SCINE Sparrow [4] providing implementations of both NDDO-based and DFTB-based methods.

Current and Future Challenges

The applicability of semiempirical methods remains constrained for the following reasons: limited availability of suitable reference data combined with the employed Hamiltonian simplifications hinders their accuracy and transferability. Linear scaling of parameters with the number of elements has been a very successful strategy for the PM6/PM7 and the GFN-xTB methods in their Wolfsberg-Helmholz-type expressions to cover 70 and 86 elements of the periodic table, respectively. In contrast, the original DFTB models use the pairwise parameterized Slater-Koster tight-binding formalism, which has limited its model coverage of the periodic table. Nowadays, a plethora of quantum chemistry packages and powerful computers are available, enabling the fast generation of theoretical reference data at large scale. With enough data, parameters for nearly arbitrary elements and, possibly, element combination can be generated.

Additionally, existing approximations in contemporary semiempirical methods may require revision for improved accuracy, transferability to more diverse chemical environments, or extended applicability to a broader set of physical properties. One direction is to better understand and systematically improve established concepts such as the NDDO approximation [5]. Another direction is to incorporate more information and concepts from first-principles calculations as in done in composite methods such as PBEh-3c [6] and avoid the approximation of multi-center integrals altogether. Furthermore, the inclusion of more basis functions or core electrons to minimal-basis models may enable new spectroscopic applications like NMR or XAS. However, increasing the number of basis functions in semiempirical methods also increases their cost and thus reduces their computational advantage over first-principles methods.

Lastly, the computational scaling and efficiency of semiempirical models needs to be improved for both existing and future models. For all semiempirical schemes, the linear algebra necessary to solve for the density matrix is the rate-determining step. To compete with existing force-field methods, this step needs to be accelerated. Different schemes relying on fragmentation, sparse linear algebra and highly parallel computing architectures have been suggested [7], but only a few of them have been successfully applied in a black-box fashion on commodity computers at large scale [8].

Advances in Science and Technology to Meet Challenges

At the core of any model improvement in semiempirical methods will be the availability of more reference data: well-balanced, in large amounts, and preferably of high quality. The development of semiempirical methods will greatly benefit from the ongoing efforts to generate large data for machine-learning (ML) models. The ML priorities will likely be different and the resulting data might not be ideally suited for fitting new semiempirical models. Particularly, semiempirical models are different from purely geometry-based ML models, especially when extrapolation beyond the reference data space is important, such as in chemical space exploration and photochemistry. For this, it will be important that semiempirical Hamiltonians can be applied with appropriate wavefunctions for both the ground and excited states. While some software implementations of semiempirical methods already include excited-state and multi-determinant functionality, semiempirical models are primarily fit to reproduce single-determinant calculations of electronic ground states because that is what the vast majority of reference data is available for.

Even with sufficient data available, it may be challenging to choose between different model ingredients. ML machinery is effective at high-dimensional interpolation, and it is possible to generate semiempirical model parameters as the output of ML models, which improves the interpretability of the overall model relative to black-box ML predictions of total electronic energies [9]. Semiempirical models may also benefit in other ways from ML developments, particularly in accelerating rate-determining steps: improved initial guesses for SCF calculations and case-specific semiempirical parameter adjustments can both be aided by ML schemes. Alternatively, the framework of statistical model selection and tools such as the Akaike Information Criterion might be useful for selecting between semiempirical models with differing numbers of parameters. An improved formal understanding of semiempirical methods can also make these choices easier.

Similar to classical force fields, semiempirical models are well-suited to benefit from heterogeneous computing architectures that can leverage mixed-precision such as commodity GPUs, which enable much faster calculations than standard computing architectures [10]. This will likely increase the relevance of GPUs in quantum chemistry, which correspondingly follows their growth in ML applications.

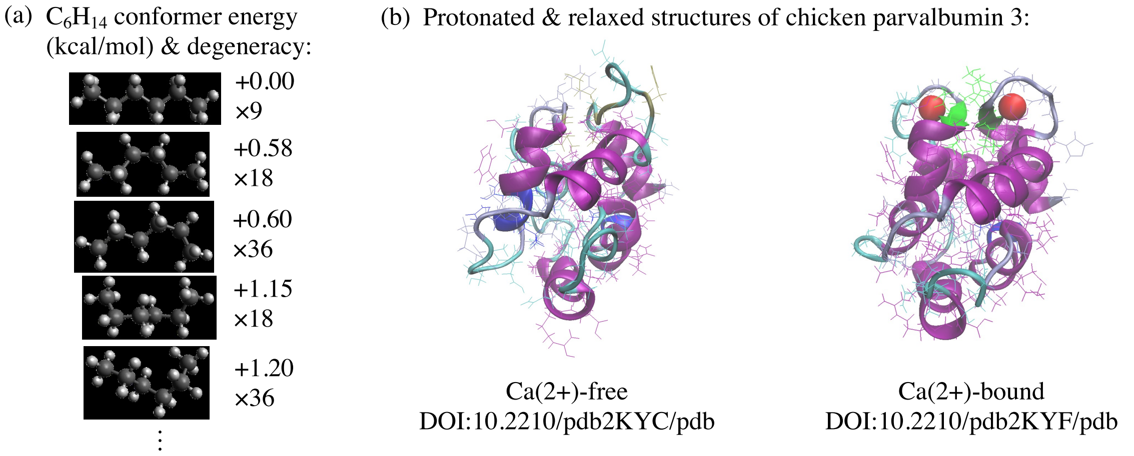


Figure 1. The low cost of semiempirical models enables novel functionality such as (a) conformer searches using GFN2-xTB and GBSA implicit water with CREST and (b) protein modeling using PM7 and COSMO implicit water with the MOZYME solver in MOPAC.

Concluding Remarks

Within electronic structure theory, semiempirical methods remained successful because of their unmatched computational efficiency. In recent years, models covering most of the periodic table have consolidated their role among computational chemists and materials scientists alike. Particularly, for chemical and conformational space exploration, examples of which are highlighted in Fig. 1, they are in frequent use. With plentiful reference data within reach, many remaining limitations might be remedied in the near future. Via modular software implementations, semiempirical Hamiltonians will become more generalizable than existing models or, alternatively, case-specific reparametrization will be highly simplified. Due to the generally low precision requirements, semiempirical models are well-suited to be combined with consumer-grade GPUs and linearly scaling algorithms. This will push the limits of routine applications that are possible with semiempirical models. Overall, semiempirical methods are as popular as ever and will remain so for the foreseeable future.

Acknowledgements

J.E.M. and the Molecular Sciences Software Institute are supported by NSF Grant No. CHE-2136142. C.B. acknowledges funding from the Ministry of Culture and Science of the German State of North Rhine-Westphalia (MKW) via the *NRW-Rückkehrprogramm.*

References

[1] James J. P. Stewart. MOPAC: A semiempirical molecular orbital program. *Journal of Computer-Aided Molecular Design*, 4(1):1–103, 1990.

[2] B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshaye, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W.-z. Yu, and T. Frauenheim. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *The Journal of Chemical Physics*, 152(12):124101, 2020.

[3] Christoph Bannwarth, Eike Caldeweyher, Sebastian Ehlert, Andreas Hansen, Philipp Pracht, Jakob Seibert, Sebastian Spicher, and Stefan Grimme. Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science*, 11(2):e1493, 2021.

[4] Francesco Bosia, Peikun Zheng, Alain Vaucher, Thomas Weymuth, Pavlo O. Dral, and Markus Reiher. Ultra-fast semi-empirical quantum chemistry for high-throughput computational campaigns with Sparrow. *The Journal of Chemical Physics*, 158(5):054118, 02 2023.

[5] Tamara Husch and Markus Reiher. Comprehensive analysis of the neglect of diatomic differential overlap approximation. *Journal of Chemical Theory and Computation*, 14(10):5169–5179, 10 2018.

[6] Stefan Grimme, Jan Gerit Brandenburg, Christoph Bannwarth, and Andreas Hansen. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *The Journal of Chemical Physics*, 143(5):054107, 2015.

[7] D R Bowler and T Miyazaki. $O(N)$ methods in electronic structure calculations. *Reports on Progress in Physics*, 75(3):036503, 2012.

[8] Robert Schade, Tobias Kenter, Hossam Elgabarty, Michael Lass, Ole Schütt, Alfio Lazzaro, Hans Pabst, Stephan Mohr, Jürg Hutter, Thomas D. Kühne, and Christian Plessl. Towards electronic structure-based ab-initio molecular dynamics simulations with hundreds of millions of atoms. *Parallel Computing*, 111:102920, 2022.

[9] Frank Hu, Francis He, and David J. Yaron. Treating semiempirical hamiltonians as flexible machine learning models yields accurate and interpretable results. *Journal of Chemical Theory and Computation*, 19(18):6185–6196, 09 2023.

[10] Xin Wu, Axel Koslowski, and Walter Thiel. Semiempirical Quantum Chemical Calculations Accelerated on a Hybrid Multicore CPU–GPU Computing Platform. *Journal of Chemical Theory and Computation*, 8(7):2272–2281, 2012.

Simulating Nuclear Dynamics with Quantum Effects

Sharon Hammes-Schiffer[1], Nancy Makri[2] and Mariana Rossi[3]

[1]Department of Chemistry, Princeton University, Princeton, New Jersey 08544
[2]Department of Chemistry, Department of Physics, and Illinois Quantum Information Science and Technology Center, University of Illinois, Urbana, Illinois 61801
[3]Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

Status

The structure and dynamics of molecules and materials, in all thermodynamic states, are determined by the laws of quantum mechanics. Solving various problems in this area requires a sufficiently accurate solution of the time-dependent or time-independent Schrödinger (or Dirac) equation for a system composed of many interacting electrons and nuclei. Under the much-celebrated Born-Oppenheimer (BO) approximation, the electronic problem has been addressed by a variety of approaches. Applications of these techniques have been central in the area of computational electronic structure theory.

Solving the equivalent equations for a system of interacting nuclei, as well as going beyond the BO approximation and obtaining the coupled dynamics of electrons and nuclei, is significantly more challenging. Accounting for nuclear quantum and non-BO effects can be far from just a small correction to a conventional calculation that considers nuclei as clamped point particles or as classical objects. Quantum nuclei have quantized energy levels, can tunnel through barriers, are delocalized, and can exhibit wave interference. Such quantum effects can dramatically change thermodynamic phase transitions, stabilize different crystal structures, influence the response of matter to stimuli, impact rates and equilibrium constants for chemical reactions, and cause isotope-dependent changes to the thermodynamics and kinetics. Thus, developing theoretical methods that incorporate quantum effects in nuclear dynamics is critical for answering many open questions in biology, physics, chemistry, and materials science.

A series of algorithmic developments, along with the increase of computer power has allowed quantum dynamics simulations of complex systems, triggering, in turn, the discovery of new situations where nuclear quantum dynamics are essential.[1] Available approaches can be broadly classified as those based on nuclear or nuclear-electronic wavefunctions, mixed quantum-classical approximations, and path integral (PI) methods. Each of these methods has advantages and limitations, as well as software implementations with varying degrees of accessibility. When choosing a method, one must balance accuracy and feasibility for the particular process of interest.

Current and Future Challenges

The ultimate goal of simulation methods is to treat all nuclei and all electrons quantum mechanically. For a wide range of important processes in chemistry and biology, this means accounting for zero-point energy (ZPE), nuclear tunneling, coherence, decoherence and quantum dissipation, treating the nuclear motion with full anharmonicity and accounting for changes in the electronic states (non-BO effects) when the nuclei rearrange.

An obvious difficulty in accounting for nuclear quantum effects in the dynamics of large molecular, biological, and condensed phase processes is the vast computational resources required to store and manipulate the quantum mechanical wavefunction. Finite-temperature effects pose an additional challenge to wavefunction-based methods when there are several low-frequency vibrational modes with many thermally populated states. Although a fully classical treatment of the nuclei cannot describe quantum effects such as hydrogen tunneling, in some cases treating (in addition to the coupled BO states) only one or a few nuclei (usually protons) by quantum mechanics is sufficient. The proper feedback among electronic states, quantum nuclei, and classical nuclei is important.

Treating the classical nuclei in terms of classical trajectories, which are local, while retaining a quantum treatment of electronic and/or some nuclear degrees of freedom, is possible through Ehrenfest's approximation, where the force on the classical particles is averaged with respect to the quantum wavefunction. Such a treatment can lead to unphysical results (for example, incorrect branching ratios). A significant improvement over Ehrenfest's approximation is achieved through surface hopping[2-3] (SH), which allows trajectories to hop between quantum states in a probabilistic fashion. Feynman's path integral (PI) formulation of quantum mechanics eliminates the need for delocalized wavefunctions, eliminating storage and allowing a consistent combination of quantum and classical treatments, but numerical integration of the resulting high-dimensional oscillatory function generally encounters serious convergence issues. When (as with normal mode vibrations, or through the validity of linear response) the nuclei can be treated as a harmonic bath coupled to the quantum system, the PI formulation offers a unique advantage, allowing a fully quantum mechanical treatment of all harmonic degrees of freedom, at zero or finite temperature, which can be evaluated using stable, numerically exact algorithms. The PI formulation in imaginary time offers an exact description of equilibrium processes with arbitrary potential functions, and efficient Monte Carlo and molecular dynamics methods are available for such calculations. This approach cannot describe time evolution but provides the basis for dynamical approximations.

Advances in Science and Technology to Meet Challenges

Fully quantum mechanical wavefunction propagation with many coupled degrees of freedom is often possible using the multiconfiguration time-dependent Hartree (MCTDH) methodology[4-5] This method converges to fully quantum mechanical results and has found many molecular applications. However, inclusion of a large number of relevant degrees of freedom and accounting for finite-temperature effects are generally not practical.

In hybrid approaches, specified nuclei are treated quantum mechanically, and the other nuclei are propagated on vibrational or vibronic surfaces with a nonadiabatic method such as SH. These approaches are useful for quantizing protons in simulations of proton transfer and proton-coupled electron transfer.[6] The nuclear-electronic orbital (NEO) approach treats specified nuclei, typically protons, quantum mechanically on the same level as the electrons with wave function or density functional theory methods.[7] The nuclear delocalization, ZPE, and tunneling of the quantum nuclei, as well as the anharmonic effects of the entire system, are inherently included. The nonadiabatic effects between the electrons and quantum nuclei are included without any BO separation, and the nonadiabatic effects of the classical nuclei with respect to the quantum subsystem can be included with Ehrenfest or SH dynamics. This approach enables real-time quantum dynamical simulations of thermal and photoexcited processes but neglects the quantum effects of the heavy nuclei, other than nonadiabatic effects. The NEO methods are multicomponent extensions of their conventional electronic structure counterparts and have been implemented in a wide range of software packages, including but not limited to Q-Chem,[8] Chronus Quantum,[9] and FHI-aims.[10-11] These software packages enable both expert and non-expert users to perform NEO calculations with ease at relatively low computational expense.

For system-bath Hamiltonians, the quasi-adiabatic propagator path integral[12] (QuAPI) removes the instabilities arising from the oscillatory quantum phase, allowing numerically exact propagation. Various developments, including the use of time-evolving matrix product operators[13] (TEMPO) to compress the QuAPI tensors, can be used to increase the efficiency in various regimes. An analytically derived small matrix decomposition[14] (SMatPI) completely eliminates tensor storage, allowing calculations with many quantum states. The modular path integral[15] (MPI) extends these methods to large molecular aggregates, where each unit includes electronic states coupled to intramolecular vibrations. These real-time PI methods, which account for all interference and decoherence effects without approximation, have been used in many simulations of proton, electron and energy transfer and are implemented in the software package PATHSUM.[16] The restriction to harmonic bath degrees of freedom is removed in the quantum-classical path integral[17] (QCPI), which captures the motion of the nuclei through classical trajectories that interact rigorously and consistently with the quantum subsystem.

The imaginary-time PI formalism for quantum statistical mechanics leads to useful and efficient (but mostly *ad hoc*) quantum dynamical approximations that can be applied to general anharmonic potentials, with a large number of quantum atoms at given thermodynamic conditions, and can be combined with electronic structure methods.[18] These methods, based on path-integral molecular dynamics (PIMD), combine quantum statistics with different types of classical time propagation, and can thus capture ZPE and incoherent tunneling effects, but completely miss quantum coherence. The recent development of Matsubara dynamics has exposed the relationship of centroid molecular dynamics[19] and (thermostatted[20]) ring polymer molecular dynamics[21] to quantum dynamics,[22] leading to new developments that improve these schemes. While it is straightforward to use these methods within the BO approximation and at equilibrium, there are many open challenges related to their extension to nonadiabatic and nonequilibrium situations.[23-24] Performing PIMD-based simulations efficiently in high-performance computing architectures requires the parallel evaluation of several replicas of the system, clever algorithms for the reduction of the number of these replicas and integrators that allow using large timesteps, among other acceleration techniques. A vast variety of such techniques are available in the open-source i-Pi code[25], which is interfaced to around 10 electronic structure codes and several other machine-learned-potential packages, allowing these methods to be broadly applied to many relevant problems in physical chemistry.

Concluding Remarks

This brief description of the challenges and advances for simulating quantum effects in nuclear dynamics cannot cover the rich history and diversity of this field, but instead focuses on a few successful approaches. Each approach has advantages and limitations, and methodological developments are underway to address the specific challenges. Importantly, these methods are currently implemented in code packages broadly adopted in the community, making them available to users addressing a wide range of problems where nuclear quantum dynamics plays a key role. Nevertheless, the goal of treating all nuclei and all electrons on equal footing beyond the BO approximation for realistic systems in a computationally practical way continues to be one of the most important frontiers in theoretical chemistry. Despite their limitations, however, the existing approaches enable simulations that provide useful insights into the physical mechanisms behind chemical and biological processes.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Awards CHE-1954348 (to S.H.S.) and CHE-1955302 (to N.M.). M.R. acknowledges funding from the Max Planck Society and the Deutsche Forschungsgemeinschaft (DFG) - Projektnummer 182087777 - SFB 951.

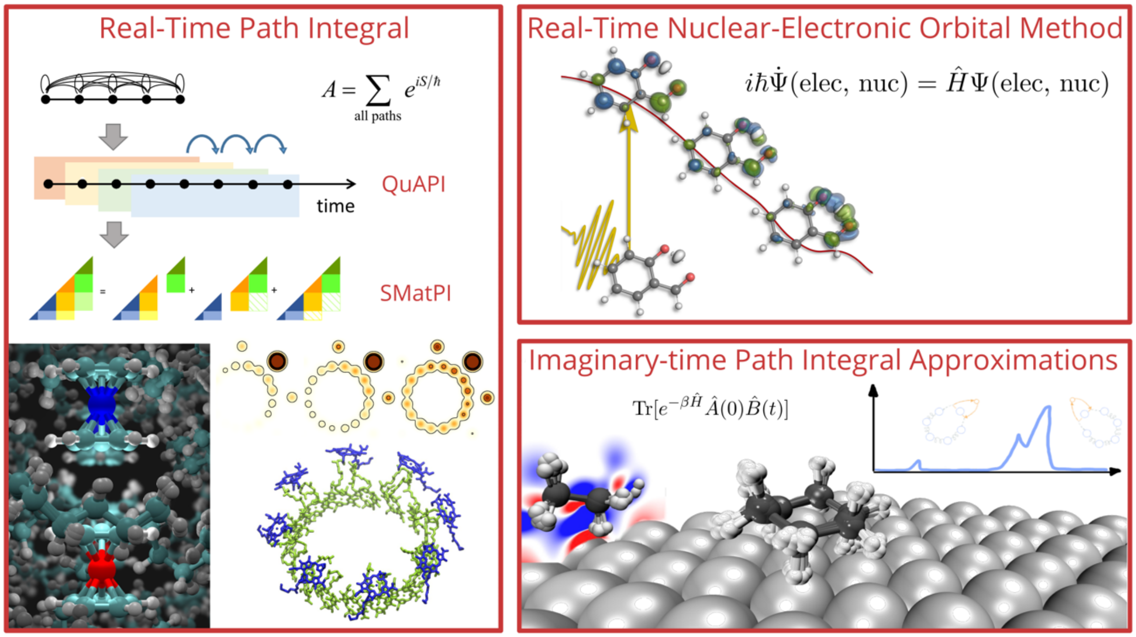


Figure Caption

Left panel: Schematic illustration of the path integral with an influence functional, the iterative QuAPI algorithm and the SMatPI decomposition. The image in the left bottom corner shows a snapshot of a QCPI simulation of electron transfer in the ferrocene-ferrocenium pair in liquid hexane, showing the solvent delocalization resulting from the superposition of three quantum-classical paths (adapted from Walters, P. L.; Makri, N., *J. Phys. Chem. Lett.* **2015,** *6*, 4959-4965). The yellow-brown contours in the right bottom corner of this panel are snapshots of the electronic density on the excited states of the 24 bacteriochlorophyll molecules in the B800-B850 LH2 complex of *Rhodopseudomonas molischianum* (with the two-ring structure shown in blue and green), following excitation of a pigment on the B800 ring (adapted from Kundu, S.; Dani, R.; Makri, N., *Science Advances* **2022,** *8*, eadd0023). **Upper right panel:** Real-time NEO-TDDFT trajectory of excited state intramolecular proton transfer following photoexcitation to the S_1 electronic state. The time-dependent electron density difference relative to the ground state is shown as green (positive) and blue (negative) isosurfaces, and the time-dependent proton density is shown as a light gray isosurface. Details are given in Ref. Zhao, L.; Tao, Z.; Pavošević, F.; Wildman, A.; Hammes-Schiffer, S.; Li, X., *J. Phys. Chem. Lett.* **2020,** *11*, 4052-4058, and the figure was adapted from the associated journal cover. **Bottom right panel:** Snapshots of *ab initio* path-integral molecular dynamics simulations of cyclohexane on Rh(111), which captures electron-density rearrangements (blue and red regions). Details in Ref. K. Fidanyan, I. Hamada, and M. Rossi, *Adv Theory Simulations* **2021**, *4*, 2000241. Such simulations can be used to approximate real-time quantum correlation functions and calculate vibrational spectra, as sketched in the upper-right corner.

References

1. Markland, T. E.; Ceriotti, M., Nuclear quantum effects enter the mainstream. *Nature Reviews Chemistry* **2018,** *2* (3), 0109.

2. Tully, J. C., Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990,** *93*, 1061-1071.

3. Hammes-Schiffer, S.; Tully, J. C., Proton transfer in solution: molecular dynamics with quantum transitions. *J. Chem. Phys.* **1994,** *101*, 4657-4667.

4. Wang, H.; Thoss, M., Multilayer formulation of the multiconfiguration time-dependent Hartree theory. *The Journal of Chemical Physics* **2003,** *119* (3), 1289-1299.

5. Wang, H., Multilayer Multiconfiguration Time-Dependent Hartree Theory. *The Journal of Physical Chemistry A* **2015,** *119* (29), 7951-7965.

6. Goyal, P.; Hammes-Schiffer, S., Tuning the Ultrafast Dynamics of Photoinduced Proton-Coupled Electron Transfer in Energy Conversion Processes. *ACS Energy Letters* **2017,** *2* (2), 512-519.

7. Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method. *Chemical Reviews* **2020,** *120* (9), 4222-4253.

8. Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergifosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, Á.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.; Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, Á.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrince, A. E., III; DiStasio, R. A., Jr.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Goddard, W. A., III; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, H. L., III; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I., Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *The Journal of Chemical Physics* **2021,** *155* (8).

9. Williams-Young, D. B.; Petrone, A.; Sun, S.; Stetina, T. F.; Lestrange, P.; Hoyer, C. E.; Nascimento, D. R.; Koulias, L.; Wildman, A.; Kasper, J.; Goings, J. J.; Ding, F.; DePrince III, A. E.; Valeev, E. F.; Li, X., The Chronus Quantum software package. *WIREs Computational Molecular Science* **2020,** *10* (2), e1436.

10. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M., Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications* **2009,** *180* (11), 2175-2196.

11. Xu, J.; Zhou, R.; Tao, Z.; Malbon, C.; Blum, V.; Hammes-Schiffer, S.; Kanai, Y., Nuclear–electronic orbital approach to quantization of protons in periodic electronic structure calculations. *The Journal of Chemical Physics* **2022,** *156* (22).

12. Makri, N., Quantum dissipative systems: a numerically exact methodology. *J. Phys. Chem.* **1998,** *102*, 4414-4427.

13. Strathearn, A.; Kirton, P.; Kilda, D.; Keeling, J.; Lovett, B. W., Efficient non-Markovian quantum dynamics using time-evolving matrix product operators. *Nature Communications* **2018,** *9*, 3322.

14. Makri, N., Small matrix disentanglement of the path integral: overcoming the exponential tensor scaling with memory length. *J. Chem. Phys.* **2020,** *152*, 041104.

15. Makri, N., Modular path integral methodology for real-time quantum dynamics. *J. Chem. Phys.* **2018,** *149*, 214108.

16. Kundu, S.; Makri, N., PathSum: A C++ and Fortran suite of fully quantum mechanical real-time path integral methods for (multi-)system + bath dynamics. *The Journal of Chemical Physics* **2023,** *158* (22).

17. Makri, N., Quantum-classical path integral: A rigorous approach to condensed phase dynamics. *International Journal of Quantum Chemistry* **2015,** *115*, 1209-1214.

18. Rossi, M., Progress and challenges in ab initio simulations of quantum nuclei in weakly bonded systems. *The Journal of Chemical Physics* **2021,** *154* (17), 170902.

19. Voth, G. A., Path integral centroid methods in quantum statistical mechanics and dynamics. *Adv. Chem. Phys.* **1996,** *93*, 135.

20. Rossi, M.; Ceriotti, M.; Manolopoulos, D. E., How to remove the spurious resonances from ring polymer molecular dynamics. *The Journal of Chemical Physics* **2014,** *140* (23), 234116.

21. Habershon, S.; Manolopoulos, D. E.; Markland, T. E.; III, T. F. M., Ring-Polymer Molecular Dynamics: Quantum Effects in Chemical Dynamics from Classical Trajectories in an Extended Phase Space. *Annual Review of Physical Chemistry* **2013,** *64* (1), 387-413.

22. Althorpe, S., Path-integral approximations to quantum dynamics. *The European Physical Journal B* **2021,** *94*.

23. Saller, M. A. C.; Runeson, J. E.; Richardson, J. O., Path-Integral Approaches to Non-Adiabatic Dynamics. In *Quantum Chemistry and Dynamics of Excited States*, 2020; pp 629-653.

24. Ananth, N., Path Integrals for Nonadiabatic Dynamics: Multistate Ring Polymer Molecular Dynamics. *Annual Review of Physical Chemistry* **2022,** *73* (1), 299-322.

25. Kapil, V.; Rossi, M.; Marsalek, O.; Petraglia, R.; Litman, Y.; Spura, T.; Cheng, B.; Cuzzocrea, A.; Meißner, R. H.; Wilkins, D. M.; Helfrecht, B. A.; Juda, P.; Bienvenue, S. P.; Fang, W.; Kessler, J.; Poltavsky, I.; Vandenbrande, S.; Wieme, J.; Corminboeuf, C.; Kühne, T. D.; Manolopoulos, D. E.; Markland, T. E.; Richardson, J. O.; Tkatchenko, A.; Tribello, G. A.; Van Speybroeck, V.; Ceriotti, M., i-PI 2.0: A universal force engine for advanced molecular simulations. *Computer Physics Communications* **2019,** *236*, 214-223.

IOP Roadmap: Real-Time Propagation in Electronic Structure Theory

Niranjan Govind

Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99354

E-mail: `niri.govind@pnnl.gov`

Yosuke Kanai

Department of Chemistry, Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3290

E-mail: `ykanai@unc.edu`

Xiaosong Li

Department of Chemistry, University of Washington, Seattle, WA, 98195-1700

E-mail: `xsli@uw.edu`

Status

Real-time electronic structure methods provide an unparalleled view of electron dynamics and ultrafast spectroscopy on the atto- and femto-second timescales, with vast potential to yield new insights into the complex electronic behavior of molecules and materials. In parallel, owing to foundational developments in experimental ultrafast science over the last three decades, culminating in the Nobel Prize in Physics in 2023 [1], the study of fundamental processes involving the dynamics of electrons on their natural timescales is now possible at X-ray free electron laser (XFEL) facilities [2] and access to data from these novel experiments offers new opportunities to validate and improve theoretical descriptions.

In a nutshell, real-time propagation in electronic structure theory explicitly considers the time-dependence of a quantum electronic system by evolving the time-dependent Schrödinger or Dirac equation in the time domain,

$$i\frac{\partial \Psi(\mathbf{r},t)}{\partial t} = \hat{H}(\mathbf{r},t)\Psi(\mathbf{r},t). \tag{1}$$

The Hamiltonian under the influence of an external perturbation results in the time-evolution of the wave function or the electron density, which forms the basis of all response properties (linear and non-linear) of a quantum electronic system, which is different from the traditional approach of casting the problem into an eigenvalue equation, in some form, that describes the system.

Beginning with pioneering developments in the late seventies on the time-dependent Hartree-Fock (TDHF) approximation to time-dependent correlated wave function methods in the early nineties to real-time time-dependent density functional theory (RT-TDDFT) since the mid-nineties and more recent scalable implementations of RT-TDDFT for both molecular and condensed phase systems have led to a broad range of studies including complex relativistic effects. In recent years, there has been renewed interest in explicit time-propagation of correlated methods such as multi-configurational self-consistent-field (MCSCF),configuration interaction (CI), algebraic diagrammatic construction, and coupled cluster (CC) theories. Alternatively, correlated electron dynamics can be modeled through the time evolution of the one-electron reduced density matrix (RDM) or the two-electron RDM, as opposed to the wave function, but such methods are plagued by N-representability problems resulting from the truncation of the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy of equations of motion for the RDMs.

Applications of real-time electronic structure methods span the field of spectroscopy, including valence-electron UV/Vis and photoelectron, circular dichroism, core-electron X-ray absorption, nonlinear optical response, photoionization, multidimensional nonlinear spectroscopies, and magnetization dynamics. These methods have also found utility in studies of molecular electronics, optimal control, coherence, charge-transfer dynamics, and non-equilibrium dynamics such as electronic stopping and electron transport in condensed phase systems [3, 4]. To probe chemical processes in complex environments, real-time electronic dynamics have been coupled to polarizable and non-polarizable molecular mechanical layers, implicit solvation models, quantum subsystems, and thermal baths within open quantum system formulations. RT-TDDFT has been also coupled with classical Maxwell equation for propagating electromagnetic fields within extended systems to simulate nonlinear light-matter interactions and particularly important for simulating intense ultrashort laser pulses [5]. For an exhaustive overview of real-time electronic structure approaches, we refer the interested reader to the recent review Ref.6.

Recent efforts in real-time electronic structure theory have also focused on extensions to multi-component 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 the intersystem crossing events, spin-couplings, and relativistic effects, variational treatments within the two- or four-component Dirac framework are needed [7]. The coupling of a molecule to a quantized electromagnetic field, real-time quantum electrodynamics (QED) [8], has led to studies of photon absorption and emission and simulations of cavity QED experiments. For many light-driven dynamic processes in chemical systems, quantum mechanical representations of proton dynamics has been demonstrated with the nuclear–electronic orbital (NEO) approach in the context of multi-component RT-TDDFT for molecular systems [9].

Current and Future Challenges

Despite advances in real-time methodologies and the broad range of natural applications, a key challenge lies in the time propagation of the wave function or the density matrix. Studies on linear and nonlinear spectroscopies and dynamical electronic processes on timescales ranging from atto to femto to even picoseconds requires a large number of steps as part of the time integration. This is because the typical electronic time step is several orders of magnitude smaller than that required for integrating ion/nuclear degrees of freedom. Mathematically, the essence of the problem lies in efficiently and accurately solving nonlinear differential equations. In addition, the correspondence between quantum Hamiltonians and unitary time propagators also imposes strict requirements on time-propagation algorithms. Another key component in real-time approaches is the construction of the Hamiltonian. While reduced scaling approaches, [10] fragment-, and embedding-based methods, [11] together with hardware advancements like GPU, have led to faster Hamiltonian construction approaches, all real-time propagation schemes still necessarily rely on sequential time-propagation. Time-acceleration approaches are still a major obstacle that greatly limits the applications of quantum electronic dynamics in practice across all real-time methodologies to date.

For investigations of condensed phase systems, RT-TDDFT largely remains the method of choice. An accurate description of excitonic effects, especially those of charge-transfer type, is a particularly important scientific challenge. While hybrid exchange-correlation (XC) approximations are promising and have been extensively used in molecular applications, the computational cost associated with evaluation of the exact exchange presents a computational bottleneck in extended systems due to the itinerant nature of orbitals. Gauge transformations techniques are increasingly pursued to reduce the large computational cost [12, 13]. Alternatively, modeling long-range screening of the electron-hole interaction using the XC vector potential has also become a promising avenue [14]. Extending the RT-NEO-TDDFT approach [9] for condensed phase systems using the periodic boundary conditions and the Brillouin zone integration enables simulation of the coupled quantum dynamics of protons and electrons in complex heterogeneous systems, opening up exciting frontiers for exploration [15].

Although QED-enabled real-time methods have emerged as a useful tool to study novel photon-driven chemical processes [16, 17], a complete theory requires a full first-principles QED electronic structure theory treatment with photon-mediated coupling between electrons and positrons and treatment of retardation with the frequency-dependent Breit Hamiltonian.

Advances in Science and Technology to Meet Challenges

GPUs are increasingly incorporated into the next generation of high performance computers in recent years. Real-time electronic structure codes would do well to take advantage of this recent advance. The INQ code [18], which is a new RT-TDDFT implementation based on the plane-wave pseudopotential formalism, is an example of such an effort. Molecular dynamics simulation has tremendously benefited from recent advances in machine-learning (ML) techniques. ML approaches like artificial neural networks have been demonstrated for efficiently estimating the quantum dynamics propagator for some simple model systems [19], and such new advances might translate also to first-principles electronic structure theories in the future. Recent quantum dynamics developments with tensor-train/matrix product state representations have also been encouraging [20]. Looking further ahead, the emerging area of quantum computing may also play an important for advancing real-time propagation approaches. For simple model systems like the spin-boson Hamiltonian, quantum algorithms for performing quantum dynamics simulation have been demonstrated [21].

Concluding Remarks

Over the last few decades, methodological developments and computing hardware advancements have greatly contributed to the increasing popularity of real-time propagation approaches in electronic structure theory. These recent developments have made it possible for researchers to investigate non-equilibrium electron dynamics beyond the usual linear response theory formalism. In addition to continued efforts in achieving greater accuracy for increasingly complex systems, our view is that describing the quantum-mechanical coupling of electron dynamics with other degrees of freedom like quantum nuclei and photons presents an important challenge and also an opportunity for the community.

Acknowledgements

NG is supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences through FWP 72684 at the Pacific Northwest National Laboratory (PNNL). YK is supported by the National Science Foundation under Award Nos. CHE-1954894 and OAC-2209858. XL is supported by the National Science Foundation under Award No. CHE-2154346 for the real-time electronic structure theory development, and by the Department of Energy, Office of Science, Basic Energy Sciences, in the Heavy-Element Chemistry program for the development of the four-component Dirac framework (Grant No. DE-SC0021100) and the development of quantum field method (Grant No. DE-SC0006863). Open-source software development is supported by the National Science Foundation (OAC-2103717 to XL) and the Offices of Advanced Scientific Computing Research (ASCR) and Basic Energy Sciences (BES) of the U.S. Department of Energy (DE-SC0022263 to XL). PNNL is operated by Battelle Memorial Institute for the United States Department of Energy under DOE Contract No. DE-AC05-76RL1830.

References

[1] The Nobel Prize in Physics 2023. https://www.nobelprize.org/uploads/2023/10/advanced-physicsprize2023-2.pdf. Accessed: 11-30-2023.

[2] XFELs WorldWide. https://xfel-hub.es/xfels-worldwide-2/. Accessed: 11-30-2023.

[3] Christopher Shepard, Dillon C. Yost, and Yosuke Kanai. Electronic excitation response of dna to high-energy proton radiation in water. *Phys. Rev. Lett.*, 130:118401, Mar 2023.

[4] Ruiyi Zhou, Dillon C. Yost, and Yosuke Kanai. First-principles demonstration of nonadiabatic thouless pumping of electrons in a molecular system. *The Journal of Physical Chemistry Letters*, 12(19):4496–4503, 2021. PMID: 33956458.

[5] Shunsuke A. Sato and Kazuhiro Yabana. Maxwell + tddft multi-scale simulation for laser-matter interactions. , 1(1):98–110, 2014.

[6] Xiaosong Li, Niranjan Govind, Christine Isborn, A Eugene DePrince III, and Kenneth Lopata. Real-time time-dependent electronic structure theory. *Chemical Reviews*, 120(18):9951–9993, 2020.

[7] Joshua J. Goings, Joseph M. Kasper, Franco Egidi, Shichao Sun, and Xiaosong Li. Real time propagation of the exact two component time-dependent density functional theory. *Journal of Chemical Physics*, 145(10):104107, 2016.

[8] Michael Ruggenthaler, F Mackenroth, and Dieter Bauer. Time-dependent kohn-sham approach to quantum electrodynamics. *Physical Review A*, 84(4):042107, 2011.

[9] Luning Zhao, Zhen Tao, Fabijan Pavošević, Andrew Wildman, Sharon Hammes-Schiffer, and Xiaosong Li. Real-time time-dependent nuclearelectronic orbital approach: Dynamics beyond the born–oppenheimer approximation. *The Journal of Physical Chemistry Letters*, 11(10):4052–4058, 2020. PMID: 32251589.

[10] Henryk Laqua, Johannes CB Dietschreit, Jorg Kussmann, and Christian Ochsenfeld. Accelerating hybrid density functional theory molecular dynamics simulations by seminumerical integration, resolution-of-the-identity approximation, and graphics processing units. *Journal of Chemical Theory and Computation*, 2022.

[11] Qiming Sun and Garnet Kin-Lic Chan. Quantum embedding theories. *Accounts of chemical research*, 49(12):2705–2712, 2016.

[12] Weile Jia and Lin Lin. Fast real-time time-dependent hybrid functional calculations with the parallel transport gauge and the adaptively compressed exchange formulation. *Computer Physics Communications*, 240:21–29, 2019.

[13] Christopher Shepard, Ruiyi Zhou, Dillon C. Yost, Yi Yao, and Yosuke Kanai. Simulating electronic excitation and dynamics with real-time propagation approach to tddft within plane-wave pseudopotential formulation. *The Journal of Chemical Physics*, 155(10):100901, 2021.

[14] Jiuyu Sun, Cheng-Wei Lee, Alina Kononov, André Schleife, and Carsten A. Ullrich. Real-time exciton dynamics with time-dependent density-functional theory. *Phys. Rev. Lett.*, 127:077401, Aug 2021.

[15] Jianhang Xu, Ruiyi Zhou, Tao E. Li, Volker Blum, Sharon Hammes-Schiffer, and Yosuke Kanai. First-principles approach for coupled quantum dynamics of electrons and protons in heterogeneous systems. 2023.

[16] Justin Malave, Alexander Ahrens, Daniel Pitagora, Cody Covington, and Kálmán Varga. Real-space, real-time approach to quantum-electrodynamical time-dependent density functional theory. *The Journal of chemical physics*, 157(19), 2022.

[17] Michael Ruggenthaler, Dominik Sidler, and Angel Rubio. Understanding polaritonic chemistry from ab initio quantum electrodynamics. *Chemical Reviews*, 123(19):11191–11229, 2023.

[18] Xavier Andrade, Chaitanya Das Pemmaraju, Alexey Kartsev, Jun Xiao, Aaron Lindenberg, Sangeeta Rajpurohit, Liang Z. Tan, Tadashi Ogitsu, and Alfredo A. Correa. Inq, a modern gpu-accelerated computational framework for (time-dependent) density functional theory. *Journal of Chemical Theory and Computation*, 17(12):7447–7467, 2021. PMID: 34726888.

[19] Maxim Secor, Alexander V. Soudackov, and Sharon Hammes-Schiffer. Artificial neural networks as

propagators in quantum dynamics. *The Journal of Physical Chemistry Letters*, 12(43):10654–10662, 2021. PMID: 34704767.

[20] Ningyi Lyu, Micheline B Soley, and Victor S Batista. Tensor-train split-operator ksl (tt-soksl) method for quantum dynamics simulations. *Journal of Chemical Theory and Computation*, 2022.

[21] Pauline J. Ollitrault, Alexander Miessen, and Ivano Tavernelli. Molecular quantum dynamics: A quantum computing perspective. *Accounts of Chemical Research*, 54(23):4229–4238, 2021. PMID: 34787398.

IOP Roadmap: Spectroscopy

Jochen Autschbach,(a) T. Daniel Crawford,(b,c) Claudia Draxl(d,e)

(a) Department of Chemistry, University at Buffalo, State University of New York, Buffalo, NY 14260-3000, USA
(b) Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA
(c) Molecular Sciences Software Institute, Blacksburg, VA 24060, USA
(d) Physics Department and IRIS Adlershof, Humboldt-Universität zu Berlin, Zum Großen Windkanal 2, 12489 Berlin, Germany
(e) European Theoretical Spectrocopy Facility

Status

Electronic structure calculations are an essential complement to experimental spectroscopy in its many forms because they can provide detailed understanding of the links between observed spectroscopic parameters/features and underlying chemical structure, bonding, environmental perturbations, and dynamics. *Ab initio* electronic structure theory can provide such links for a broad array of spectroscopic probes for both molecules and materials. Beyond the energy spectrum itself, most spectroscopic properties can be defined via perturbation theory as derivatives of a stationary energy or the time-averaged quasi-energy [1, 2]— a molecular or material *response*, in other words, which is accessible via derivative techniques or via suitable Fourier transforms in time-dependent simulations [3].

In the molecular domain, which includes gas-phase as well as solvated or encapsulated species, properties that feature prominently in current research include electric and magnetic multipole moments and transition moments (electric and magnetic field derivatives), field gradients, NMR magnetic shielding and spin-spin coupling (derivatives with respect to external and nuclear spin magnetic field amplitudes), vibrational frequencies and corresponding IR and Raman intensities (involving electric field and nuclear position derivatives), or chiroptical properties (mixed electric/magnetic/nuclear position derivatives). In higher orders, a 'zoo of properties' [4] is accessible, including nonlinear susceptibilities, multi-photon transition moments, magnetic field-induced optical activity, etc. Calculations tend to be based on density functional theory (DFT), time-dependent DFT (TD-DFT), or some flavor of wavefunction theory (WFT), although semi-empirical methods remain in use in some sub-fields.

For condensed-matter systems, there is a large variety of spectroscopic techniques, including (angle-resolved) photoemission (ARPES), optical absorption, second-harmonic generation, x-ray absorption (XANES), photoluminescence, resonant inelastic x-rays scatterin (RIXS), resonant Raman scattering, energy electron loss spectrocsopy (EELS), and more. They probe the various interactions taking place in a material on the same energy scale, i.e., electron-electron interaction, electron-hole correlation, electron-phonon coupling [5] as well as magnetic effects and spin-orbit coupling. The light-matter interaction is very often treated in the linear-response (LR) regime, where the methods of choice [6] are Green-function based approaches as realized in many-body perturbation theory (MBPT) or TD-DFT. The latter also allows for going beyond LR, propagating the excited system in time. What method is most appropriate also depends on the nature of the material.

Current and Future Challenges

The challenges the field of computational spectroscopy faces include both *formulating* and *predicting* spectroscopic responses—getting the right answer for the right reason. However, in practice, researchers must often settle for 'decent results for good reasons', although even this can be elusive for more challenging systems or properties that are difficult to calculate. In the molecular domain, predictions of spectroscopic properties are often sensitive to (i) the treatment of electron correlation, (ii) basis set quality, (iii) environmental interactions, and (iv) dynamic effects. The polynomial computational scaling of the most accurate and robust theoretical methods, such as coupled cluster (CC) theory[7], present a substantial obstacle for convergent simulations, and, as a result, conventional implementations of such methods are currently limited to fewer than 50 atoms (and even smaller systems for response properties). DFT-based methods are able to treat larger systems, but the development of suitable functional approximations remains a formidable challenge. There are additional complexities in TD-DFT [8], such as memory effects in the exchange-correlation potential and the associated response kernels. At present, the vast majority calculations rely on the adiabatic approximation. For systems with heavy elements, or in very high-accuracy calculations, the level at which Einstein relativity is treated is another dimension in which calculations must converge [9]. For open-shell species, electron-spin angular momentum presents special challenges for magnetic-field responses, a problem that is further exacerbated by spin contamination/ill-defined spin-states for many high-accuracy WFT methods. Multi-configurational states in general pose challenges for the calculation of (response) properties, be it in DFT, WFT, or some combination of the two approaches, because of the need to both treat static and dynamic correlation.

Related problems exist for solid-state systems. While DFT calculations can nowadays be carried out for 1000 atoms and more, depending on the employed basis set and functional, for excited-state properties, it is typically an order of magnitude less. Just to name one example, the GW method of MBPT suffers, besides the formally quartic scaling with system size, from slow convergence with the number of empty states and the starting-point dependence, i.e., the underlying DFT functional, if carried out in a perturbative ("single-shot") manner. On the methodology side, there are several ways of tackling self-consistency; other issues may come from the various implementations and algorithms used in different codes. All this hampers fair comparison and the assessment of what *the* result of a given method for a certain material is. Similar arguments hold for the Bethe-Salpeter equation or TDDFT. Most forward-looking approaches concern the description of time-dependent phenomena to tackle, for instance, the evolution of charge excitations, the build-up and decay of electron-hole pairs — also considering exciton-phonon coupling. These are particularly challenging as methodological and algorithmic complexity comes also with tremendous computational costs.

Advances in Science and Technology to Meet Challenges

While advances in computing hardware have significantly extended the reach of computational methods to larger and more complex systems, only substantial improvements in the formulation of accurate models and their algorithmic implementations will ultimately overcome the polynomial scaling wall of electronic structure theory. To that end, for WFT approaches such as coupled cluster theory, continued progress in localization, fragmentation, and other reduced-scaling techniques will prove to be vital for modeling the spectroscopic responses of molecular systems containing hundreds to thousands of atoms, including explicit simulations of dynamic solvent effects. There is also a pressing need for practical approaches to deal with the combination of static and dynamic correlation in the wavefunctions as well as in their response. On the DFT/TD-DFT front, practical approaches with widespread adoption that go beyond the adiabatic approximation have yet to emerge, and the treatment of multi-configurational states and their response remains a challenge.

In order for theoretical approaches to meet the rapidly increasing resolution and capabilities on the experimental side and to going substantially beyond currently accessible system sizes, progress is required on all levels: (i) novel methodology, (ii) adequate approximations, (iii) highly-performant algorithms, (iv) exascale compute power, as well as (v) collaborative efforts by the community. Point (i), concerns processes involving — speaking in the language of Green's functions, going beyond 2- and 4-point functions that are currently state of the art for condensed matter — and ways for efficiently describing non-equilibrium dynamics. Here, point (ii) comes into play where clever strategies need to be found that may be very much tailored to a specific excitation processes of interest. On the computational side (iii), only proper algorithms that scale on hundred thousands of processors will allow the community to make use of exascale computers (iv), the first of which are currently being launched. Regarding (v), first steps towards reaching the ambitious goals, are for instance the EU centers of excellence (see, e.g. NOMAD, MAterials design at the eXascale, and Targeting Real Chemical accuracy at the EXascale) on exascale computing [10] and related initiatives in the US, such as the Molecular Sciences Software Institute. Thereby, the NOMAD CoE is particularly dedicated to advanced methods, including CC theory and excited states.

Concluding Remarks

As experimental techniques advance at a rapid pace, theory and computation must continue to evolve. Different spectroscopic properties of a system probe the electronic structure in different ways, which means that they likely expose the approximations in a calculation to different degrees. There is no single computational approach available, emerging, or even conceivable at present, that would be able to treat a large variety of spectroscopic parameters at the same level of accuracy, applicable to fairly large molecular systems or complex materials, and be sufficiently accurate for most intended applications. In other words, there is exciting and important research to be done.

Acknowledgements

TDC is supported by the U.S. National Science Foundation via grants CHE-2154753 and CHE-2136142. CD appreciates support from the European Union's Horizon 2020 research and innovation program under the grant agreement N$^{\underline{o}}$ 951786 (NOMAD CoE) and from the Deutsche Forschungsgemeinschaft (DFG), project 182087777 (SFB 951). JA acknowledges support from the National Science Foundation, grant CHE-2152633.

References

[1] Trygve Helgaker, Sonia Coriani, Poul Jørgensen, Kasper Kristensen, Jeppe Olsen, and Kenneth Ruud. Recent advances in wave function-based methods of molecular-property calculations. *Chem. Rev.*, 112:543–631, 2012.

[2] Patrick Norman, Kenneth Ruud, and Trond Saue. *Principles and Practices of Molecular Properties: Theory, Modeling, and Simulations.* Wiley, UK, 2018.

[3] Xiaosong Li, Niranjan Govind, Christine Isborn, A. Eugene DePrince, and Kenneth Lopata. Real-time time-dependent electronic structure theory. *Chem. Rev.*, 120:9951–9993, 2020.

[4] Jochen Autschbach. *Quantum Theory for Chemical Applications.* Oxford University Press, New York (NY), 2021.

[5] Feliciano Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89:015003, Feb 2017.

[6] Giovanni Onida, Lucia Reining, and Angel Rubio. Electronic excitations: density-functional versus many-body green's-function approaches. *Rev. Mod. Phys.*, 74:601–659, Jun 2002.

[7] T. Daniel Crawford, Ashutosh Kumar, Alexandre P. Bazanté, and Roberto Di Remigio. Reduced-scaling coupled cluster response theory: Challenges and opportunities. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, e1406:1–25, 2019.

[8] Jan Gerit Brandenburg, Kieron Burke, Emmanuel Fromager, Matteo Gatti, Sara Giarrusso, Nikitas I. Gidopoulos, Paola Gori-Giorgi, Duncan Gowland, Trygve Helgaker, Matthew J. P. Hodgson, Lionel Lacombe, Gianluca Levi, Pierre-François Loos, Neepa T. Maitra, Eduardo Maurina Morais, Nisha Mehta, Filippo Monti, Manasi R. Mulay, Katarzyna Pernal, Lucia Reining, Pina Romaniello, Matthew R. Ryder, Andreas Savin, Dumitru Sirbu, Andrew M. Teale, Alex J. W. Thom, Donald G. Truhlar, Jack Wetherell, and Weitao Yang. New approaches to study excited states in density functional theory: general discussion. *Faraday Discuss.*, 224:483–508, 2020.

[9] M. Reiher and A. Wolf. *Relativistic quantum chemistry. The fundamental theory of molecular science.* Wiley-VCH, Weinheim, second edition, 2015.

[10] H J Kulik, T Hammerschmidt, J Schmidt, S Botti, M A L Marques, M Boley, M Scheffler, M Todorović, P Rinke, C Oses, A Smolyanyuk, S Curtarolo, A Tkatchenko, A P Bartók, S Manzhos, M Ihara, T Carrington, J Behler, O Isayev, M Veit, A Grisafi, J Nigam, M Ceriotti, K T Schütt, J Westermayr, M Gastegger, R J Maurer, B Kalita, K Burke, R Nagai, R Akashi, O Sugino, J Hermann, F Noé, S Pilati, C Draxl, M Kuban, S Rigamonti, M Scheidgen, M Esters, D Hicks, C Toher, P V Balachandran, I Tamblyn, S Whitelam, C Bellinger, and L M Ghiringhelli. Roadmap on machine learning in electronic structure. *Electronic Structure*, 4(2):Kulik2022, jun 2022.

Tools for exploring potential energy surfaces

Satoshi Maeda, Department of Chemistry and Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University
H. Bernhard Schlegel, Department of Chemistry, Wayne State University

Status

Structures and reactivities of molecules and materials are governed by potential energy surfaces (PESs) [1,2]. Tools for exploring PESs thus are vital for analysing and predicting the behaviour of these systems. The PES is a function of electronic energy with structural parameters as variables. Figure 1(a) shows a schematic of a model PES. On a PES, there are local minima corresponding to stable structures. They are connected by reaction paths passing over first-order saddles representing transition states (TSs).

In many theoretical studies, identifying stable structures is the first step. This can be done routinely for a wide range of system sizes thanks to efficient and robust geometry optimization techniques such as quasi-Newton algorithms [1,2]. In complex systems, an experimentally observed compound can be an ensemble of many conformers. Structural sampling methods such as molecular dynamics (MD) simulation, genetic algorithms and Monte Carlo Basin Hopping are useful for finding relevant conformers and isomers systematically [1].

Exploring PES to characterise molecular reactivity involves finding TSs and reaction paths connecting stable structures. This provides an energy profile as illustrated in Figure 1(b). TSs are often difficult to optimize and require a good initial guess. Nudged elastic band (NEB) and related string methods, coordinate driving and the artificial force induced reaction (AFIR) method are robust techniques for getting close to TSs [3,4]. The mechanisms of reactions of many types, such as organic reaction, organometallic catalysis, organocatalysis and heterogeneous catalysis, have been elucidated based on energy profiles. MD is also useful in simulating ultrafast processes such as those with low barriers or triggered by photoabsorption, but running MD simulations longer than a microsecond is often impractical.

As systems increase in size, numerous conformations of the TSs need to be considered. Structural sampling of TSs is therefore necessary for probing stereoselectivity in organo and organometallic catalysis. Construction of global reaction path networks as in Figure 1(c) provides *ab initio* predictions of chemical reactivity including kinetics of competing paths and formation of by-products [1,3,4,5]. Therefore, further developments in this area will contribute to improving the accuracy, efficiency, and robustness of chemical structure and reactivity predictions.

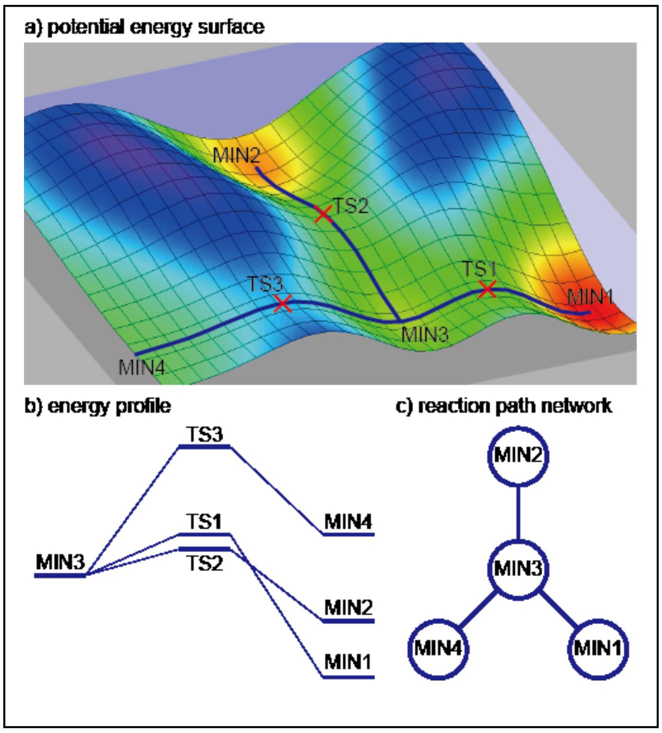


Figure 1. (a) A schematic of a two-dimensional PES, (b) an energy profile representation of (a), and (c) a reaction path network representation of (a). MINs and TSs correspond to local minima and transition states, respectively.

Current and Future Challenges

While current methods are practical for individual reaction steps, the ultimate goal is to predict the reactivity for an entire system such as the Strecker synthesis illustrated in Figure 2. Typical reaction networks involve hundreds of intermediate and thousands of individual reaction steps [3,4,5]. There are three major difficulties in exploring reaction networks: the combinatorial explosion of the structural space, the cost of electronic structure calculations in large systems, and the reliability of the methods for characterizing individual reaction steps.

The example in Figure 2 contains many paths leading to energetically unstable compounds. To establish the physically relevant paths connecting the reactants to the major product, one must show that all other paths are less reactive and determine all potential by-products. In principle, it is necessary to systematically explore all intermediates and reactions connecting them. This leads to a combinatorial explosion as the size of the reactive system increases.

To provide chemically useful predictions of reaction networks, suitably accurate electronic structure calculations are required. Semi-empirical methods are usually not sufficient for calculating barrier heights. Density functional theory (DFT) has a much wider range of applicability. Highly correlated methods like coupled cluster (CC) and complete active space self-consistent field (CASSCF) methods may be required for more accurate calculations of rates. While individual calculations may be affordable, computing thousands of elementary reactions that comprise a reaction network can be very costly. Furthermore, the electronic structure calculations must account for the environment – solvent for reactions in solution, protein embedding for enzymatic reactions and surfaces for heterogeneous catalysis.

Clearly, with thousands of individual reactions to explore for a reaction network, automated workflows are needed. In manual exploration of an individual reaction step, if one calculational fails, there are numerous methods to try until one achieves success. By contrast, in an automated workflow, the calculations for exploring a PES to characterize individual reaction step (electronic structure computations, optimization techniques, etc.) need to be nearly 100% successful. In addition, graphical user interfaces are needed to make the exploration of reaction networks accessible to the general chemistry community and not just to the computational chemistry specialist.

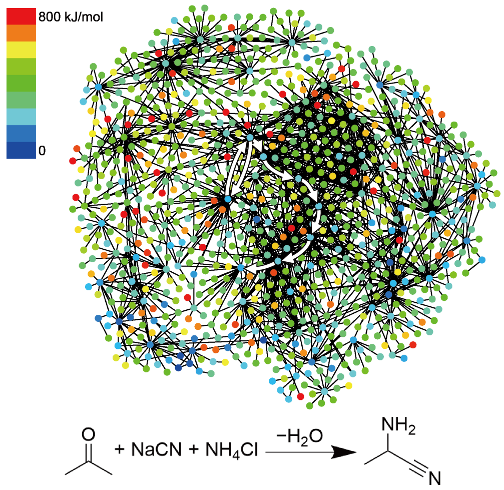


Figure 2. An example illustrating how complex a network describing a chemical reaction can be, even for a simple organic reaction, in this case the condensation step of the Strecker amino acid synthesis indicated in the chemical equation. Nodes and edges represent different chemical species and elementary steps, respectively. The network comprehensively includes not only the most feasible route to the major product, α-aminonitrile, highlighted by the white arrows, but also minor paths that cross the boundary between kinetically accessible and inaccessible regions from reactant's node, thus proving that the reaction yields α-aminonitrile. The figure was reproduced from [*AsiaChem Magazine* **2021**, *2*, 56-63 (DOI: 10.51167/acm00024)] (**CC BY 4.0**).

Advances in Science and Technology to Meet Challenges

Computer codes to meet some of the challenges are already being developed [3,4,5] but more advances in software are needed. Automated workflows are essential since it is not practical to explore the thousands of reaction steps in a network by manual methods. One example is the artificial force induced reaction (AFIR) method that systematically explores paths and minima by inducing various chemical transformations by applying an artificial force to different fragment pairs of various local minima [3]. Alternatively, double ended and growing string methods can be used to obtain reaction paths followed by TS optimization [4]. Another approach is to use molecular dynamics at a high temperature to explore the PES [6], but this is typically limited to modest levels of electronic structure theory.

The biggest challenge is the combinatorial explosion associated with an exhaustive search of a reaction network. Limiting the search of reaction networks to desired pathways can be done based on kinetic criteria, thermodynamic criteria, structural criteria, and/or computational cost-related criteria [3,5]. When the truncation is done based on a kinetics simulation under reaction time and temperature, the exploration could be regarded as an on-the-fly kinetics simulation. Heuristic rules (like ‘arrow pushing’) can be used to limit the network to more feasible reactions [7]. With a large enough database, artificial intelligence and machine learning can be trained to identify the most probable reaction paths [8].

Reaction path databases will also help reduce the cost of exploring PESs. Machine learning can improve the accuracy of electronic structure calculations thereby producing higher quality PESs at lower cost [9]. Databases and machine learning can provide better starting estimates of TSs and reaction paths [8]. Gaussian process regression can improve the efficiency of optimizing intermediates, TSs and reaction paths [10]. Since the many individual reaction steps in a large network are independent, they can be explored in parallel. Libraries of algorithms for the many different tasks involved in exploring potential energy surfaces will speed the development of more robust, reliable and efficient codes for calculating reaction path networks.

Concluding Remarks

Geometry optimization has grown into a tool that anyone in the chemistry community can use practically to investigate a wide range of chemical systems. In recent years, various techniques have been developed to explore and analyse global reaction pathway networks, enabling the prediction of the whole picture of a chemical reaction, taking into account not only the energy profile for the major product but also the paths for side reactions. Further development, however, is needed to apply these techniques to complex systems. Taming combinatorial explosions in the chemical space to be explored is one of the major difficulties. For accurate and effective exploration of reaction networks, it is also essential to employ state-of-the-art electronic structure calculations, environment modelling techniques, and informatics methods involving databases and machine learning. Solving these problems and enabling the prediction of chemical reaction networks routinely and reliably is one of the grand challenges of computational chemistry.

Acknowledgements

Please include any acknowledgements and funding information as appropriate.
SM thanks the Japanese Science and Technology Agency for financial support (grant JST-ERATO JPMJER1903) and HBS thanks the National Science Foundation US for funding (grant CHE1856437).

References

[1] Wales, D. J. Energy Landscapes: with Applications to Clusters, Biomolecules and Glasses (Cambridge University, Cambridge, England, 2003).
[2] Schlegel, H. B. Geometry optimization. *WIREs Comput. Mol. Sci.* **2011**, *1*, 790-809.
[3] Maeda, S.; Harabuchi, Y. Exploring Paths of Chemical Transformations in Molecular and Periodic Systems: An Approach Utilizing Force. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1538.
[4] Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1354.
[5] Unsleber, J. P.; Reiher, M. The exploration of chemical reaction networks. *Annu. Rev. Phys. Chem.* **2020**, *71*, 121-142.
[6] Wang, L. P.; Titov, A.; McGibbon. R.; Liu. F.; Pande, V. S.; Martinez, T. J. Discovering chemistry with an ab initio nanoreactor. *Nat. Chem.* **2014**, *6*, 1044–1048.
[7] Rappoport, D; Galvin, C. J.; Zubarev, D. Y.; Aspuru-Guzik, A. Complex chemical reaction networks from heuristics-aided quantum chemistry. *J. Chem. Theory Comput.* **2014**, *10*, 897–907.
[8] Meuwly, M. Machine learning for chemical reactions. *Chem. Rev.* **2021**, *121*, 10218−10239.
[9] Nandi, A.; Qu, C.; Houston, P. L.; Conte, R.; Bowman, J. M. **Δ**-machine learning for potential energy surfaces: A PIP approach to bring a DFT-based PES to CCSD(T) level of theory. *J. Chem. Phys.* **2021**, *154*, 051102.

[10] Denzel, A.; Kästner, J. Gaussian Process Regression for Transition State Search*.* *J. Chem. Theory Comput.* **2018,** *14*, 5777−5786.

IOP Roadmap: Managing complex computational workflows

Anubhav Jain

Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley CA USA

Ask Hjorth Larsen

CAMD, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Giovanni Pizzi

Theory and Simulation of Materials (THEOS), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
Laboratory for Materials Simulations (LMS), Paul Scherrer Institut (PSI), CH-5232 Villigen PSI, Switzerland

Status

The development and use of formal workflows in computational materials science, and in particular density functional theory (DFT) calculations, is a relatively recent trend. Workflow frameworks, initially rudimentary, became necessary about 15 years ago to perform high-throughput calculations. Their further development resulted in the ecosystem of frameworks nowadays available to the community and their applications towards creating large online databases of materials properties. Today, workflows enable researchers to perform large numbers of calculations (with a single workflow potentially defining hundreds of individual simulations), standardize calculation procedures, reduce errors, make it simpler to perform multiple simulations, increase reproducibility, and make calculation techniques available to non-experts.

At its core, a *workflow* is a specification of multiple computational processing steps as a dependency graph, see Fig. 1. A *workflow framework* helps coordinate and execute *simulation codes* over computing resources, potentially interacting with a *queue manager* to obtain and distribute those computing resources. Additionally, workflow frameworks for computational material science must define and manage complex workflows, execute jobs on various high performance computing (HPC) platforms, support long-running jobs and facilitate rerunning and fixing calculation errors, record provenance, and assist in data management. The

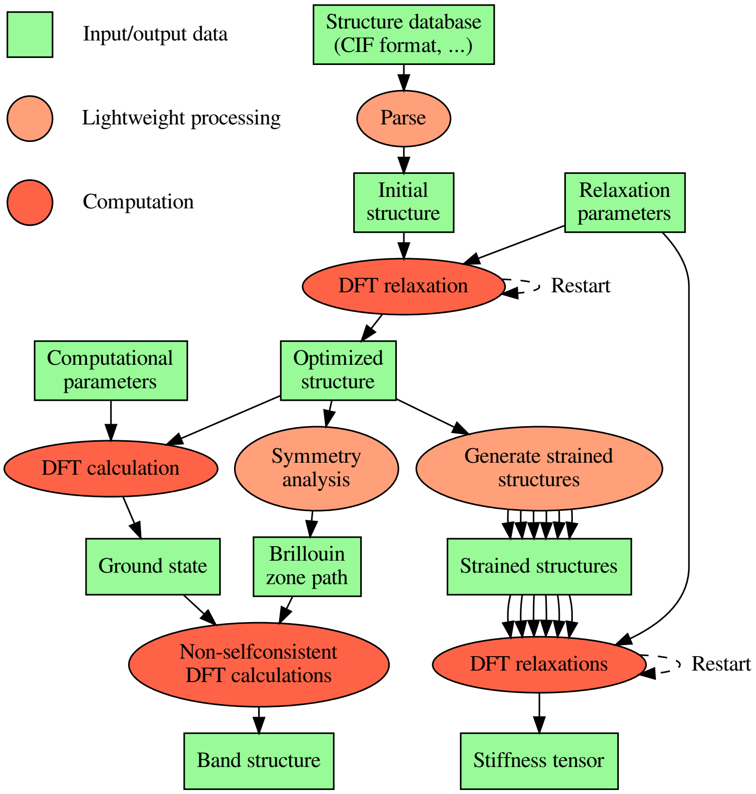


Figure 1. Typical elements of a materials workflow. Processing steps are represented by ellipses while input and output data is represented by boxes. Some inputs are chosen directly by the user whereas others may be taken from the output of previous processing steps. Longer multi-step computations like relaxations may be automatically continued or *restarted* if a run fails due to walltime limit. Another common feature is the ability to recover from failures by switching to more stable algorithms as a fallback. A workflow framework takes care of storing and organizing input and output data in files or databases.

currently dominant programming language for such frameworks is Python, which facilitates integration with common materials analysis libraries such as ASE [1] and pymatgen [2]. Available workflow frameworks include AFLOW [3], used to create the AFLOWlib database [4], AiiDA [5], used to create the databases on Materials Cloud [6], ASR+ASE+myqueue [7, 8], used to create the C2DB database [9], atomate+FireWorks [10, 11], used to create the Materials Project database [12], httk [13], used to create the Open Materials Database [14], MISPR [15], pyiron [16] and qmpy [17], used to develop the OQMD database [18], among others. Commercial offerings such as Materials Studio [19], MedeA [20], and SimStack [21] are also available; these generally emphasize and facilitate graphical user interface-based interaction. With the many options available, workflow frameworks are now standard tools for tens of thousands of materials researchers worldwide. More details on some of these frameworks and how they are used, particularly in the context of the Battery2030+ initiative, can be found in Ref. [22].

Current and Future Challenges and Advances to Meet Them

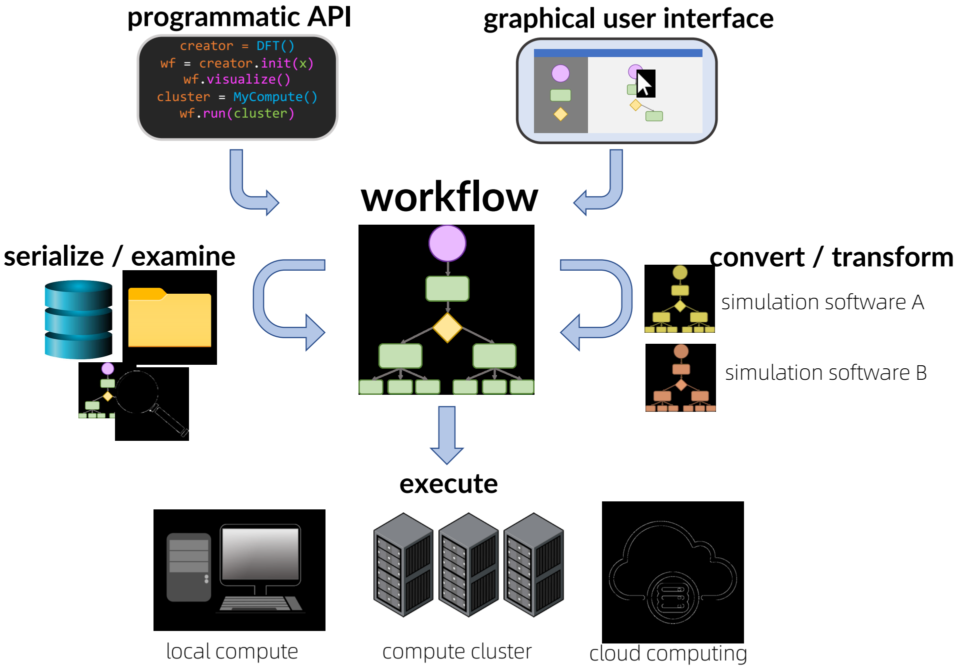
FAIR and Reproducible Data

In the past few years, the field has seen a strong increase of awareness for the need of ensuring FAIR (Findable, Accessible, Interoperable and Reusable) [23] access to research data. In addition, a fifth principle should also be considered: *Reproducibility*, a cornerstone of the scientific method. Due to the extremely large number of simulations (and the huge amount of resulting data) that can be managed automatically by current complex computational workflows, the need for generating FAIR and reproducible data is particularly relevant, and workflow frameworks will need to embed FAIR concepts natively into their design to make the process simple and straightforward for researchers. This can be achieved by ensuring a detailed tracking of the history of data and simulations, to guarantee reproducibility not only of individual simulations but of the full computational workflow; and by providing native functionalities to export data adopting ontologies and interoperable data and metadata formats recognised by the scientific domain (e.g., the OPTIMADE API for sharing molecules and crystal structures [24]).

Beyond FAIR Data

FAIR concepts should be pushed beyond just data, towards providing also FAIR access to simulations, workflows, and analysis tools. It will be crucial to democratize simulations, making current and future advanced workflows accessible also to non-experts: e.g., experimentalists who might want to use them to interpret experimental results or guide the design of new experiments. This goal can be achieved by a concerted combination of the following ingredients: 1) automatic selection of numerical (non-physical) parameters of simulation codes; 2) robust workflows able to recover from failures of the simulation codes (such as non-convergence) or, even better, by implementing more robust algorithms directly into the codes (e.g., switching to slower but more robust variational minimization methods, rather than iterative ones, when the latter do not converge); 3) making workflows interoperable and code-agnostic by using a common language for workflow inputs and outputs, only specific to a given simulation task [25, 26, 27, 28], thus enabling transparent swapping of simulation codes; 4) graphical user interfaces (GUIs) to assist non-experts in preparing the inputs, submitting and monitoring the workflows, and analyzing the results.

A further outstanding challenge is making workflows independent of the managing workflow framework (and not only of the DFT code). This goal is extremely relevant for workflow implementers, who can thus develop high-level workflows for advanced materials properties without having to limit themselves to one specific workflow framework. This remains a challenging task because of the different design approaches and concepts of each workflow framework, even when these are implemented in the same programming language (e.g., Python). These challenges are summarized in Figure 2.



```
creator = DFT()
wf = creator.init(x)
wf.visualize()
cluster = MyCompute()
wf.run(cluster)
```

Figure 2. Core challenges and desired functionalities in realizing flexible workflows for materials science. Ideally, creation of the the central workflow object could be performed using either a programmatic application programming interface (API) or a graphical user interface. Such workflow objects could be serialized, shared, and examined and modified by others. Furthermore, workflow specifications should ideally be inter-convertible so that they are independent of the particular simulation software being used to perform the core calculations. Finally, execution of the workflows should be flexible to computing hardware, including cloud computing options.

Software Ecosystem and Platforms

Additional challenges include developing an ecosystem where both commercial and open-source packages can coexist and interoperate, leveraging the strengths of each of them (e.g., selecting them based on available simulation features, their FAIR-sharing capabilities, or their GUI and accessibility). On the technical side, it will be critical for workflow frameworks to become fully independent of the computing infrastructure, e.g., to support running simulations directly on the cloud, rapidly gaining attractiveness as an effective alternative to standard HPC supercomputers. A further technical challenge is to ensure that the HPC-center access models (currently designed for humans directly interacting with systems via input and output files) can support seamless integration with database-driven workflows. These are indeed often difficult to integrate due to security issues (e.g., multi-factor authentication in the case of workflow frameworks pushing jobs) or network and firewall issues (e.g., HPC compute nodes restricted from accessing or sending data to a database).

Sustainable software: Funding and careers

The challenges discussed here relate more to software development than physics. Data formats, compatibility, and portability are central, as is the long-term maintenance of the software necessary to solve the tasks. By contrast, funding opportunities are overwhelmingly centered on science and publication metrics rather than the development and maintenance of tools and infrastructure. Since long-term options to fund qualified software developers are scarce, research groups have to deal with large amounts of code and data developed during short-term projects, and are forced to maintain these by borrowing time from researchers who are neither funded nor particularly specialized in that line of work. This model is not particularly sustainable; to address it, there must exist realistic career paths to attract qualified software engineers in the long term. Universities and funding agencies should reassess funding priorities and increase focus on sustainable software infrastructure.

Concluding Remarks

The development of materials-science workflows is still relatively young, but has already enabled research and discovery in the past 15 years in ways that we would not have imagined. Workflows have helped formalize code that researchers have written over the past decades, making research more reproducible and allowing scientists to focus less on technicalities and more on scientific discovery. In the same way as the development of DFT codes has moved from self-written, in-house codes towards today's large-scale, accessible and efficiently hardware-accelerated codes, we expect that also the existing challenges for managing computational workflows will be addressed effectively in the near future.

Acknowledgements

AJ acknowledges funding by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract no. DE-AC02-05-CH11231: Materials Project program KC23MP. AHL acknowledges funding from the European Union's Horizon 2020 research and innovation program Grant agreement No. 951786 (NOMAD CoE). GP acknowledges the NCCR MARVEL (a National Centre of Competence in Research, funded by the Swiss National Science Foundation, grant No. 205602) and the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957189 (BIG-MAP), also part of the BATTERY 2030+ initiative under grant agreement No. 957213.

References

[1] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph

Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017.

[2] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

[3] Stefano Curtarolo, Wahyu Setyawan, Gus L.W. Hart, Michal Jahnatek, Roman V. Chepulskii, Richard H. Taylor, Shidong Wang, Junkai Xue, Kesong Yang, Ohad Levy, Michael J. Mehl, Harold T. Stokes, Denis O. Demchenko, and Dane Morgan. Aflow: An automatic framework for high-throughput materials discovery. *Computational Materials Science*, 58:218–226, 2012.

[4] Stefano Curtarolo, Wahyu Setyawan, Shidong Wang, Junkai Xue, Kesong Yang, Richard H. Taylor, Lance J. Nelson, Gus L.W. Hart, Stefano Sanvito, Marco Buongiorno-Nardelli, Natalio Mingo, and Ohad Levy. Aflowlib.org: A distributed materials properties repository from high-throughput ab initio calculations. *Computational Materials Science*, 58:227–235, 2012.

[5] Sebastiaan P. Huber, Spyros Zoupanos, Martin Uhrin, Leopold Talirz, Leonid Kahle, Rico Häuselmann, Dominik Gresch, Tiziano Müller, Aliaksandr V. Yakutovich, Casper W. Andersen, Francisco F. Ramirez, Carl S. Adorf, Fernando Gargiulo, Snehal Kumbhar, Elsa Passaro, Conrad Johnston, Andrius Merkys, Andrea Cepellotti, Nicolas Mounet, Nicola Marzari, Boris Kozinsky, and Giovanni Pizzi. Aiida 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance. *Scientific Data*, 7(1):300, Sep 2020.

[6] Leopold Talirz, Snehal Kumbhar, Elsa Passaro, Aliaksandr V. Yakutovich, Valeria Granata, Fernando Gargiulo, Marco Borelli, Martin Uhrin, Sebastiaan P. Huber, Spyros Zoupanos, Carl S. Adorf, Casper Welzel Andersen, Ole Schütt, Carlo A. Pignedoli, Daniele Passerone, Joost VandeVondele, Thomas C. Schulthess, Berend Smit, Giovanni Pizzi, and Nicola Marzari. Materials cloud, a platform for open computational science. *Scientific Data*, 7(1):299, Sep 2020.

[7] Morten Gjerding, Thorbjørn Skovhus, Asbjørn Rasmussen, Fabian Bertoldo, Ask Hjorth Larsen, Jens Jørgen Mortensen, and Kristian Sommer Thygesen. Atomic simulation recipes: A python framework and library for automated workflows. *Computational Materials Science*, 199:110731, 2021.

[8] Jens Jørgen Mortensen, Morten Gjerding, and Kristian Sommer Thygesen. Myqueue: Task and workflow scheduling system. *Journal of Open Source Software*, 5(45):1844, 2020.

[9] Morten Niklas Gjerding, Alireza Taghizadeh, Asbjørn Rasmussen, Sajid Ali, Fabian Bertoldo, Thorsten Deilmann, Nikolaj Rørbæk Knøsgaard, Mads Kruse, Ask Hjorth Larsen, Simone Manti, Thomas Garm Pedersen, Urko Petralanda, Thorbjørn Skovhus, Mark Kamper Svendsen, Jens Jørgen Mortensen, Thomas Olsen, and Kristian Sommer Thygesen. Recent progress of the computational 2d materials database (c2db). *2D Materials*, 8(4):044002, jul 2021.

[10] Kiran Mathew, Joseph H. Montoya, Alireza Faghaninia, Shyam Dwarakanath, Muratahan Aykol, Hanmei Tang, Iek heng Chu, Tess Smidt, Brandon Bocklund, Matthew Horton, John Dagdelen, Brandon Wood, Zi Kui Liu, Jeffrey Neaton, Shyue Ping Ong, Kristin Persson, and Anubhav Jain. Atomate: A high-level interface to generate, execute, and analyze computational materials science workflows. *Computational Materials Science*, 139:140–152, November 2017.

[11] Anubhav Jain, Shyue Ping Ong, Wei Chen, Bharat Medasani, Xiaohui Qu, Michael Kocher, Miriam Brafman, Guido Petretto, Gian-Marco Rignanese, Geoffroy Hautier, Daniel Gunter, and Kristin A. Persson. Fireworks: a dynamic workflow system designed for high-throughput applications. *Concurrency and Computation: Practice and Experience*, 27(17):5037–5059, 2015.

[12] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials

innovation. *APL Materials*, 1(1):011002, 07 2013.

[13] Rickard Armiento. *Database-Driven High-Throughput Calculations and Machine Learning Models for Materials Design*, pages 377–395. Springer International Publishing, Cham, 2020.

[14] `https://openmaterialsdb.se/`.

[15] Rasha Atwi, Matthew Bliss, Maxim Makeev, and Nav Nidhi Rajput. Mispr: an open-source package for high-throughput multiscale molecular simulations. *Scientific Reports*, 12(1):15760, Sep 2022.

[16] Jan Janssen, Sudarsan Surendralal, Yury Lysogorskiy, Mira Todorova, Tilmann Hickel, Ralf Drautz, and Jörg Neugebauer. pyiron: An integrated development environment for computational materials science. *Computational Materials Science*, 163:24–36, 2019.

[17] `https://github.com/wolverton-research-group/qmpy`.

[18] James E. Saal, Scott Kirklin, Muratahan Aykol, Bryce Meredig, and C. Wolverton. Materials design and discovery with high-throughput density functional theory: The open quantum materials database (oqmd). *JOM*, 65(11):1501–1509, Nov 2013.

[19] `https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-materials-studio/`.

[20] `https://www.materialsdesign.com/medea-software`.

[21] Celso R. C. Rêgo, Jörg Schaarschmidt, Tobias Schlöder, Montserrat Penaloza-Amion, Saientan Bag, Tobias Neumann, Timo Strunk, and Wolfgang Wenzel. SimStack: An intuitive workflow framework. *Frontiers in Materials*, 9, 2022.

[22] Joerg Schaarschmidt, Jie Yuan, Timo Strunk, Ivan Kondov, Sebastiaan P. Huber, Giovanni Pizzi, Leonid Kahle, Felix T. Bölle, Ivano E. Castelli, Tejs Vegge, Felix Hanke, Tilmann Hickel, Jörg Neugebauer, Celso R. C. Rêgo, and Wolfgang Wenzel. Workflow engineering in materials design within the BATTERY 2030+ project. *Advanced Energy Materials*, page 2102638, December 2021.

[23] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J. G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A. C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. The fair guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, 2016.

[24] Casper W. Andersen, Rickard Armiento, Evgeny Blokhin, Gareth J. Conduit, Shyam Dwaraknath, Matthew L. Evans, Ádám Fekete, Abhijith Gopakumar, Saulius Gražulis, Andrius Merkys, Fawzi Mohamed, Corey Oses, Giovanni Pizzi, Gian-Marco Rignanese, Markus Scheidgen, Leopold Talirz, Cormac Toher, Donald Winston, Rossella Aversa, Kamal Choudhary, Pauline Colinet, Stefano Curtarolo, Davide Di Stefano, Claudia Draxl, Suleyman Er, Marco Esters, Marco Fornari, Matteo Giantomassi, Marco Govoni, Geoffroy Hautier, Vinay Hegde, Matthew K. Horton, Patrick Huck, Georg Huhs, Jens Hummelshøj, Ankit Kariryaa, Boris Kozinsky, Snehal Kumbhar, Mohan Liu, Nicola Marzari, Andrew J. Morris, Arash A. Mostofi, Kristin A. Persson, Guido Petretto, Thomas Purcell, Francesco Ricci, Frisco Rose, Matthias Scheffler, Daniel Speckhard, Martin Uhrin, Antanas Vaitkus, Pierre Villars, David Waroquiers, Chris Wolverton, Michael Wu, and Xiaoyu Yang. OPTIMADE, an API for exchanging materials data. *Scientific Data*, 8(1), aug 2021.

[25] Kiran Mathew, Joseph H. Montoya, Alireza Faghaninia, Shyam Dwarakanath, Muratahan Aykol, Hanmei Tang, Iek heng Chu, Tess Smidt, Brandon Bocklund, Matthew Horton, John Dagdelen, Brandon Wood, Zi Kui Liu, Jeffrey Neaton, Shyue Ping Ong, Kristin Persson, and Anubhav Jain. Atomate: A high-level interface to generate, execute, and analyze computational materials science workflows. *Computational Materials Science*, 139:140–152, November 2017.

[26] Morten Gjerding, Thorbjørn Skovhus, Asbjørn Rasmussen, Fabian Bertoldo, Ask Hjorth Larsen, Jens Jørgen Mortensen, and Kristian Sommer Thygesen. Atomic simulation recipes: A python framework and library for automated workflows. *Computational Materials Science*, 199:110731, 2021.

[27] Sebastiaan P. Huber, Emanuele Bosoni, Marnik Bercx, Jens Bröder, Augustin Degomme, Vladimir Dikan, Kristjan Eimre, Espen Flage-Larsen, Alberto Garcia, Luigi Genovese, Dominik Gresch, Conrad Johnston, Guido Petretto, Samuel Poncé, Gian-Marco Rignanese, Christopher J. Sewell, Berend Smit, Vasily Tseplyaev, Martin Uhrin, Daniel Wortmann, Aliaksandr V. Yakutovich, Austin Zadoks, Pezhman Zarabadi-Poor, Bonan Zhu, Nicola Marzari, and Giovanni Pizzi. Common workflows for computing material properties using different quantum engines. *npj Computational Materials*, 7(1):136, 2021.

[28] Daniel G. A. Smith, Annabelle T. Lolinco, Zachary L. Glick, Jiyoung Lee, Asem Alenaizan, Taylor A. Barnes, Carlos H. Borca, Roberto Di Remigio, David L. Dotson, Sebastian Ehlert, Alexander G. Heide, Michael F. Herbst, Jan Hermann, Colton B. Hicks, Joshua T. Horton, Adrian G. Hurtado, Peter Kraus, Holger Kruse, Sebastian J. R. Lee, Jonathon P. Misiewicz, Levi N. Naden, Farhad Ramezanghorbani, Maximilian Scheurer, Jeffrey B. Schriber, Andrew C. Simmonett, Johannes Steinmetzer, Jeffrey R. Wagner, Logan Ward, Matthew Welborn, Doaa Altarawy, Jamshed Anwar, John D. Chodera, Andreas Dreuw, Heather J. Kulik, Fang Liu, Todd J. Martínez, Devin A. Matthews, Henry F. Schaefer, Jiří Šponer, Justin M. Turney, Lee-Ping Wang, Nuwan De Silva, Rollin A. King, John F. Stanton, Mark S. Gordon, Theresa L. Windus, C. David Sherrill, and Lori A. Burns. Quantum chemistry common driver and databases (qcdb) and quantum chemistry engine (qcengine): Automation and interoperability among computational chemistry programs. *The Journal of Chemical Physics*, 155(20):204801, 2021.

IOP Roadmap: Current and Future Computer Architectures

Jeff R. Hammond
NVIDIA Helsinki Oy, Helsinki, Finland

William Dawson
RIKEN Center for Computational Science, Kobe, Japan

Takahito Nakajima
RIKEN Center for Computational Science, Kobe, Japan

Status

Computational chemistry applications have been heavy users of the most advanced computer architectures for decades now, with algorithms and software adapting to technological shifts at every level of High Performance Computing (HPC) systems. For example, when in early machines compute capability outpaced storage it became expedient to develop “direct” algorithms that reduced storage requirements by recomputing atomic integrals repeatedly [1]. When large-scale computing resources transitioned from shared-memory to interconnected distributed-memory architectures, codes like NWChem [2] were designed from scratch to take advantage of these systems and, in due course, new software techniques for managing their complexity and diversity were developed.

As distributed computing became the norm and essentially all software began to assume MPI and POSIX-compatible operating systems as a common foundation, stagnation in the growth of processor frequencies led to significant changes in processor architecture (Figure 1). The number of independent processing units (cores) grew rapidly, and were often combined with fine-grained parallelism in the form of vector, or SIMD, instructions. Massively parallel processors, especially graphics processing units (GPUs), carried this trend even further. Around the year 2000, a processor was a single core running at $\sim$ 1GHz, and could do 1-2 arithmetic operations per cycle. In the year 2022, server processors with as many as 128 cores running at $\sim$ 3GHz, each capable of 64 arithmetic operations per cycle, are common. The most powerful HPC systems have multiple GPUs per node, each capable of more than 10^{14} arithmetic operations per second. As with prior paradigm shifts, novel algorithms and software were developed to address new technologies, such as GPUs [3].

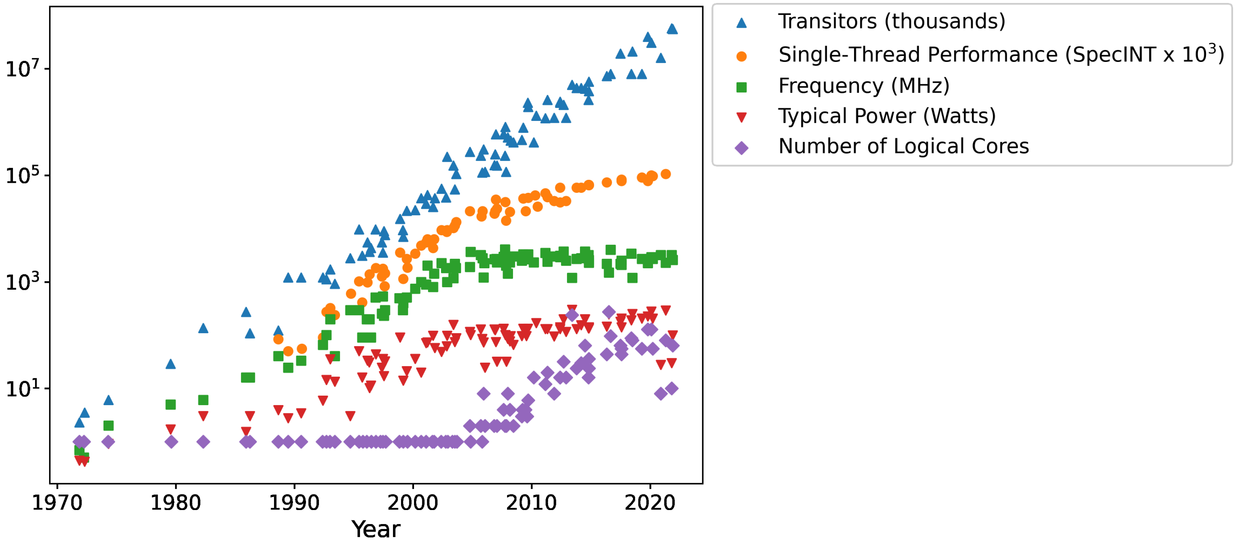


Figure 1: Computer processor capability trends over the past 50 years (taken from Ref. [4]). Since 2005, frequencies are flat and growth in sequential performance has slowed dramatically. Processor parallelism has increased to utilize still-growing transistor counts.

Current and Future Challenges

While increasingly powerful computer hardware can unlock larger system sizes or higher accuracy methods, more efficient algorithms can deliver increased scientific capability using the same or fewer resources. For example, the development of reduced-scaling methods for density-functional theory (DFT) has made possible orders of magnitude larger simulations than is possible with conventional $O(N^3)$ algorithms [5]. Unfortunately, algorithmic developments can not easily displace highly tuned implementations of baseline algorithms: systems are co-designed for performance on standard implementations, reduction in scaling leads to a lower FLOPs/byte ratio, and the software skills associated with the new algorithms are often quite different. This challenge is not unique to chemistry, and the pursuit of multiple Ps – performance, portability, productivity – is a major initiative in HPC [6].

There is no simple solution to this challenge. Some developers have chosen to adopt programming languages and environments more friendly to novice programmers (e.g. PySCF, Fermi.jl), which is a major shift from the long tradition of building large codes using Fortran without modern features for abstraction. While code that is easier to write may run slower than more laborious implementations, it is always faster than that which is never written at all. One generally applicable method for addressing the aforementioned tension in software development is to shift away from monolithic applications – each with its own atomic integral package, SCF solver, etc. – towards designs that can bring together the best components, which are themselves written by performance and numerical experts. This approach has been common in engineering disciplines for years, but is slowly gaining traction in chemistry. Additionally, recent developments in using Artificial Intelligence for code generation may provide wholly new kinds of solutions.

Advances in Science and Technology to Meet Challenges

The massive increase in node counts and the inadequacy of using *only* coarse-grained parallelism (e.g. MPI) for supporting these architectures requires a dramatic shift in how chemists develop software. Support for threading and vectorization requires redesigning and rewriting critical portions of applications such as atomic integrals, exchange correlation evaluations, and tensor computations. The composition of high-performance libraries requires chemists to think about memory management and synchronization, or to depend on a qualified programming framework. At the same time, most reduced-scaling algorithms require a shift away from large array-based data structures to sparse containers, which are both more complicated to reason about and lack a simple parallelization strategy.

Because these challenges are not unique to chemistry, there is substantial progress in support for parallelism in the general purpose programming environments. For example, since C++17, the ensemble of standard template library algorithms support parallelism where possible, and the ubiquitous Python NumPy framework now supports parallelism up to and including multi-node, multi-GPU systems [7]. A key difference between C++ standard parallelism and previous approaches is the availability of multiple product-grade implementations, which provide the greatest chance of long-term support over the multiple decades that chemistry applications are expected to live.

In Figure 2 we present an example of multiple levels of parallelism and algorithm optimization combined in the context of state-of-the-art hardware: RIKEN's Fugaku Supercomputer, which was the fastest machine in the world as of 2021. A new version of the NTChem code that uses reduced-scaling algorithms based on sparse data structures with adaptive precision [8] and an optimized atomic integral library [9] has been developed which can run on thousands of nodes. These calculations were driven from a Jupyter virtual notebook using the PyBigDFT library [10]; the ease of programming in Python enabled the creation of a workflow incorporating automatic generation of a fragment guess and composition of multiple levels of accuracy.

Concluding Remarks

Rapid changes in computer architecture and major shifts in mainstream programming methodologies present significant challenges to the production computational chemistry codes in use today. Any new code developed must be capable of parallel execution across a range of processors – either directly or via libraries – to be considered high performance. At the same time, scientific creativity in algorithms and simulations should not be limited by tedious programming models. The tension between performance and productivity will drive

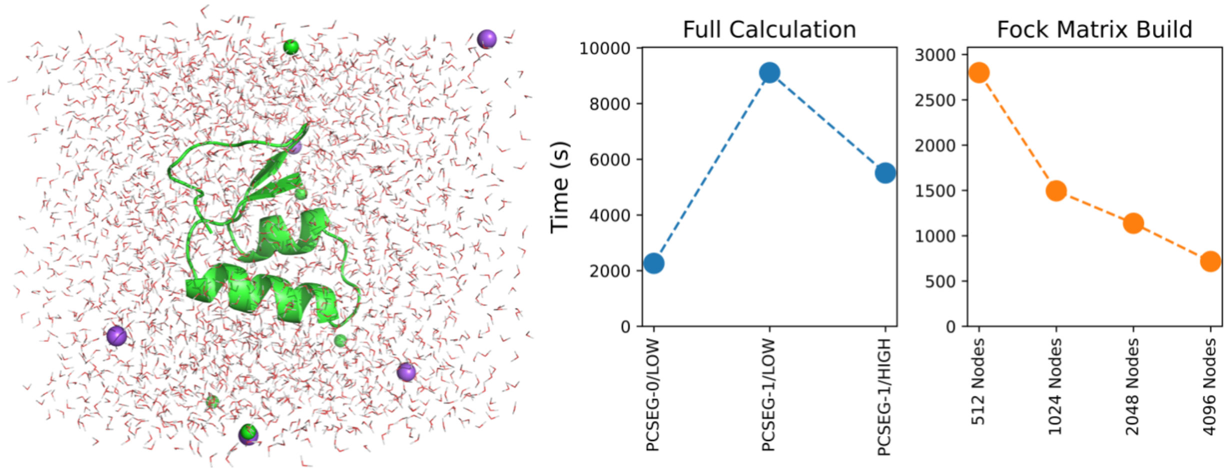


Figure 2: Calculation time required for a converged Hartree-Fock calculation of the 1CRN protein in a NaCl solution using 1024 nodes of Fugaku. With the largest basis set there are 8539 atoms and 69413 basis functions. Strong scaling is measured for a single Fock matrix build at the HIGH level of accuracy using the converged density.

rapid changes in the computational chemistry software ecosystem, and require significant investment in new ideas by developers.

Acknowledgements

This work used computational resources of the supercomputer Fugaku provided by RIKEN through the HPCI System Research Project (Project ID: hp200179), and was supported by MEXT as “Program for Promoting Research on the Supercomputer Fugaku” (Realization of innovative light energy conversion materials utilizing the supercomputer Fugaku, Grant Number JPMXP1020210317).

References

[1] J. Almlöf, K. Faegri Jr., and K. Korsell. Principles for a direct SCF approach to LCAO–MO ab-initio calculations. *Journal of Computational Chemistry*, 3(3):385–399, 1982.

[2] E. Aprà, E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. J. J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F. W. Aquino, R. Atta-Fynn, J. Autschbach, N. P. Bauman, J. C. Becca, D. E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G. N. Chuev, C. J. Cramer, J. Daily, M. J. O. Deegan, T. H. Dunning, M. Dupuis, K. G. Dyall, G. I. Fann, S. A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A. W. Götz, J. Hammond, V. Helms, E. D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B. G. Johnson, H. Jónsson, R. A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R. D. Lins, R. J. Littlefield, A. J. Logsdail, K. Lopata, W. Ma, A. V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J. E. Moore, J. M. Mullin, T. Nakajima, D. R. Nascimento, J. A. Nichols, P. J. Nichols, J. Nieplocha, A. Otero-de-

la Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati, J. Pittner, L. Pollack, R. M. Richard, P. Sadayappan, G. C. Schatz, W. A. Shelton, D. W. Silverstein, D. M. A. Smith, T. A. Soares, D. Song, M. Swart, H. L. Taylor, G. S. Thomas, V. Tipparaju, D. G. Truhlar, K. Tsemekhman, T. Van Voorhis, Á. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K. D. Vogiatzis, D. Wang, J. H. Weare, M. J. Williamson, T. L. Windus, K. Woliński, A. T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, and R. J. Harrison. NWChem: Past, present, and future. *The Journal of Chemical Physics*, 152(18):184102, 2020.

[3] Ivan S. Ufimtsev and Todd J. Martínez. Graphical processing units for quantum chemistry. *Computing in Science & Engineering*, 10(6):26–34, 2008.

[4] Karl Rupp. https://github.com/karlrupp/microprocessor-trend-data, 2022.

[5] Laura E. Ratcliff, Stephan Mohr, Georg Huhs, Thierry Deutsch, Michel Masella, and Luigi Genovese. Challenges in large scale quantum mechanical calculations. *WIREs Computational Molecular Science*, 7(1):e1290, Jan 2017.

[6] S. John Pennycook, Jason D. Sewall, Douglas W. Jacobsen, Tom Deakin, and Simon McIntosh-Smith. Navigating performance, portability, and productivity. *Computing in Science & Engineering*, 23(5):28–38, 2021.

[7] NVIDIA. https://developer.nvidia.com/cunumeric, 2022.

[8] William Dawson and Takahito Nakajima. Massively parallel sparse matrix function calculations with NTPoly. *Computer Physics Communications*, 225:154–165, 2018.

[9] Kazuya Ishimura and Shigeru Nagase. A new algorithm of two-electron repulsion integral calculations: a combination of Pople–Hehre and McMurchie–Davidson methods. *Theoretical Chemistry Accounts*, 120(1):185–189, May 2008.

[10] William Dawson, Eisuke Kawashima, Laura E. Ratcliff, Muneaki Kamiya, Luigi Genovese, and Takahito Nakajima. Complexity reduction in density functional theory: Locality in space and energy. *The Journal of Chemical Physics*, 158(16):164114, Apr 2023.

IOP Roadmap: Electronic structure software engineering

Lori A. Burns
Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, Atlanta, Georgia 30332, USA

Micael J. T. Oliveira
Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, D-22761 Hamburg, Germany

Benjamin P. Pritchard
Molecular Sciences Software Institute, Virginia Tech, Blacksburg, Virginia 24060, USA

Lyudmila V. Slipchenko
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA

Status

Electronic structure (ES) theory research groups have historically coalesced around one of the few tens of software suites that collected the large body of support code needed by developers to implement new methods and by users to serve as general-purpose research tools. The difficulties of distributing even one software project with high-performance computing (HPC) requirements, the broad competencies of developers, and the reassurance of correctness from method implementations in multiple suites led to high-performing code and science but served to reinforce the monolithic ES suite approach. Some background in hardware, software, and paradigms in quantum chemistry (QC) is reviewed at [1], and the state of scientific software in general is discussed at [2].

However, libraries as specialty tasked and even specially named collections of code have long been part of ES suites. Accompanying general growth over the past decade in open-source software and tooling (*e.g.*, GitHub and Cloud services), independent repository ES libraries (see citations in [3] and [4]) have been newly written (e.g., CheMPS2, LibEFP, ELSI, and Wannier90), extracted (e.g., LibXC and PSolver), and congregated around (e.g., LibXC and Libint). Among broadest use are integrals libraries (half-a-dozen major QC suite users), DFT libraries (>4 dozen), and, as human resources, the basis set library, BSE[5], and the norm-conserving pseudopotentials PseudoDojo library[6]. A compilation and discussion of open-source software in QC is at [3] and a survey of software libraries and their usage in physics packages can be found in [4]; some present and future roles for libraries are shown in Fig. 1. As domain experts become more specialized, there are advantages in letting them focus on independent library development that may be shared among ES suites to avoid excessive re-implementation and promote an ecosystem of loosely coupled but highly cohesive software projects. As niches fill with libraries, it is worth seeking *modularity*, the additional step of standardizing API or data interface to facilitate interchangability. Modularity has been tackled before through the Common Component Architecture (CCA) project[7, 8], which designed conventions still in use today. Successors assembling modular ecosystems include NWChemEX[9], CECAM[4], and MolSSI[10].

Current and Future Challenges

Software libraries bring many advantages to software developers and users, but their increasing use in ES codes over the last decade has revealed some issues and challenges. Foremost is designing the library's internal structure and application programming interface (API) so as to keep it cohesive and broadly useful. For example, APIs require some degree of standardization and stability, which partially conflicts with the flexibility necessary for open-ended research. Designating a clear and/or optimal division of tasks between the library and the caller is also often challenging, especially for parallelism and error handling.

Other challenges in the concrete interface and interfacing infrastructure are worth mentioning:

- As stand-alone packages, libraries require their own build systems, test suites, examples, and documentation, all of which present a shifting maintenance burden upon their developers.
- Many libraries provide poor or outdated documentation, making it hard to properly use their API or to build and install them.
- The proliferation of languages used for scientific computing places strain on providing maximal interfacing routes, not only for C or for Fortran and Python, but for Rust, Julia and other emerging languages.
- Additional tension arises from the desirability for the library itself to be written in low-level languages for broad call-ability and to minimize required dependencies, while still taking full advantage of flexible existing code like Python's SciPy for optimization.
- Most intimidating, the proliferation of hardware for cutting-edge speed places considerable stress on library developers whose expertise is likely in an ES domain, not computer science.

Finally, we would like to mention the challenge of packaging and distributing the libraries. For end users, the monolithic codes offered the advantage of a single software package to install, however difficult, with the assurance of implicit integration testing and version compatibility among its components. This is no longer the case when dealing with loose collections of libraries. Additionally, it is becoming more common that users want to deploy ES suites in workflows or install many in the same environment. In this scenario, strict requirements, like specific dependency versions and build options, present obstacles.

Advances in Science and Technology to Meet Challenges

Introduction of various "cross" tools assists in mitigating hardware and language proliferation problems. For example, build systems like modern-style CMake considerably smooth transitions between operating systems, compiler families, source languages, CPU/GPU targets, compile parameters, and developer versus user build environments. Hardware abstractions such as Intel OneAPI and NVIDIA CUDA allow cross-architecture and cross-accelerator (CPU, GPU, FPGA, etc.) access from a single implementation. Cross-platform and cross-language package managers like Conda, Spack or EasyBuild provide homogeneous specification for full heterogeneous software stacks, reducing variant hassle for users and packagers. Additionally, compiler options in Intel and projects like HPy can reduce the number of build variants while still allowing high-performance and lenient pins.

As libraries maintain their development and repository independence, affiliations with generic, scientific, and content management system (CMS) umbrella projects can help with software engineering standards and duties. For example, Linux packagers impose file layout, license, versioning, and library characteristics requirements. Similarly, conda-forge provides guidance and demands explicit dependency version management as well as nudges projects into compatibility with newer compilers and dependency versions and onto new hardware architectures. CECAM provides integration testing for their affiliated libraries in order to produce release bundles. Psi4 provides periodic CMake and Windows build help as well as integration testing. MolSSI provides a cookiecutter template with software engineering tools for Python projects and communication standardization and testing though QCArchive stack.

Overall, while affiliations, update bots, and the gradual polishing of tumbling through the open-source software ecosystem can aid in external management aspects, the lack of hardware expertise continues to be largely unaddressed. Whether GPU or parallelism challenges can be met by traveling experts or affiliations that lint projects for best practices remains to be seen.

Concluding Remarks

Over the years, there has been an increasing trend towards development of separate modules, libraries, and tools. These packages often implement seemingly small but important and often delicate pieces required for computation. In addition, adoption of industry best practices has also improved software stability and ease of distribution and installation. Use of these libraries allow researchers to focus on new science and capabilities while offloading other concerns to specialists. Overall, this should be viewed as a positive trend. However, with the

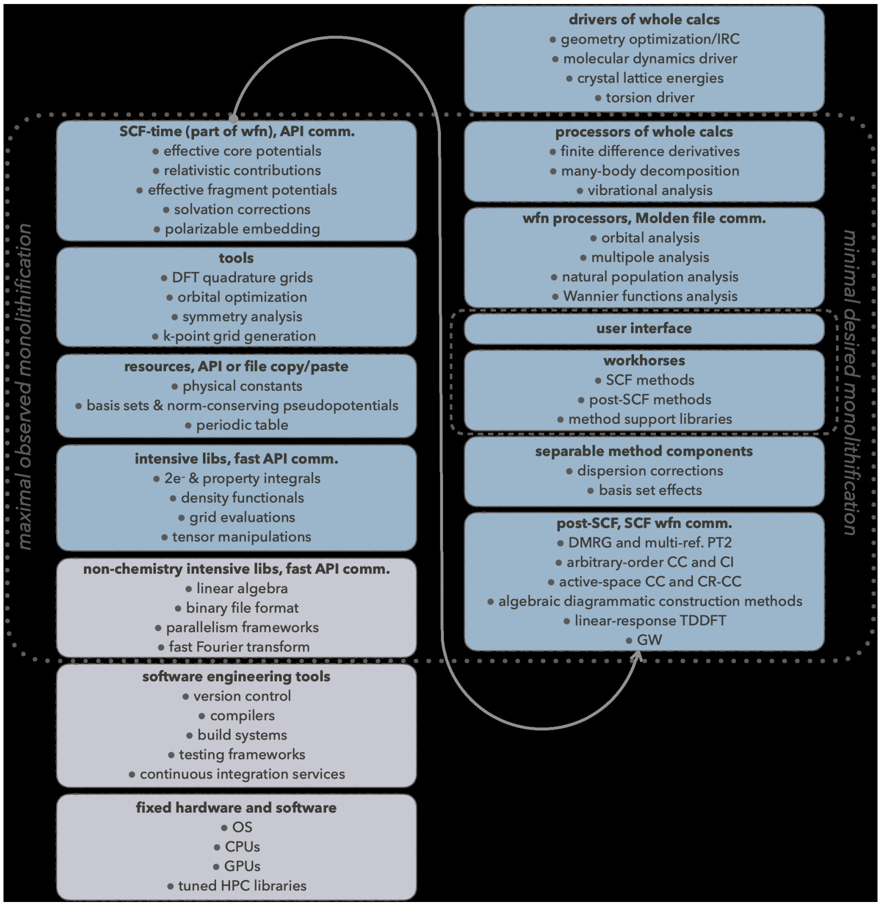


Figure 1. Existing and potential libraries and modules for electronic structure. See also Fig. 1 of [4].

explosion of packages, and an increasing diversity of programming languages and hardware, challenges emerge with respect to the interfaces between these packages. While these challenges are being tackled by the broader computational science and general programming communities, development and adoption of standards within the ES community should also help to alleviate these obstacles.

Acknowledgements

Please include any acknowledgements and funding information as appropriate.

L. A. B. was supported in part by the U.S. National Science Foundation through grant ACI-1449723. M. J. T. O wishes to thank CECAM for its continuing support of the Electronic Structure Library initiative and all the people who have contribute to the project since its inception. L. V. S. acknowledges support of the U.S. National Science Foundation (grants CHE-1800505, CHE-2102639).

References

[1] Mark S. Gordon, Giuseppe Barca, Sarom S. Leang, David Poole, Alistair P. Rendell, Jorge L. Galvez Vallejo, and Bryce Westheimer. Novel computer architectures and quantum chemistry. *The Journal of Physical Chemistry A*, 124(23):4557–4582, 2020. PMID: 32379450.

[2] Tim Storer. Bridging the chasm: A survey of software engineering practice in scientific programming. *ACM Comput. Surv.*, 50(4), aug 2017.

[3] Susi Lehtola and Antti J. Karttunen. Free and open source software for computational chemistry education. *WIREs Computational Molecular Science*, n/a(n/a):e1610, 2022.

[4] Micael J. T. Oliveira, Nick Papior, Yann Pouillon, Volker Blum, Emilio Artacho, Damien Caliste, Fabiano Corsetti, Stefano de Gironcoli, Alin M. Elena, Alberto García, Víctor M. García-Suárez, Luigi Genovese, William P. Huhn, Georg Huhs, Sebastian Kokott, Emine Küçükbenli, Ask H. Larsen, Alfio Lazzaro, Irina V. Lebedeva, Yingzhou Li, David López-Durán, Pablo López-Tarifa, Martin Lüders, Miguel A. L. Marques, Jan Minar, Stephan Mohr, Arash A. Mostofi, Alan O'Cais, Mike C. Payne, Thomas Ruh, Daniel G. A. Smith, José M. Soler, David A. Strubbe, Nicolas Tancogne-Dejean, Dominic Tildesley, Marc Torrent, and Victor Wen-zhe Yu. The CECAM electronic structure library and the modular software development paradigm. *The Journal of Chemical Physics*, 153(2):024117, 2020.

[5] Benjamin P. Pritchard, Doaa Altarawy, Brett Didier, Tara D. Gibson, and Theresa L. Windus. New basis set exchange: An open, up-to-date resource for the molecular sciences community. *Journal of Chemical Information and Modeling*, 59(11):4814–4820, 2019. PMID: 31600445.

[6] M.J. van Setten, M. Giantomassi, E. Bousquet, M.J. Verstraete, D.R. Hamann, X. Gonze, and G.-M. Rignanese. The pseudodojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Computer Physics Communications*, 226:39–54, 2018.

[7] David E. Bernholdt, Benjamin A. Allan, Robert Armstrong, Felipe Bertrand, Kenneth Chiu, Tamara L. Dahlgren, Kostadin Damevski, Wael R. Elwasif, Thomas G. W. Epperly, Madhusudhan Govindaraju, Daniel S. Katz, James A. Kohl, Manoj Krishnan, Gary Kumfert, J. Walter Larson, Sophia Lefantzi, Michael J. Lewis, Allen D. Malony, Lois C. Mclnnes, Jarek Nieplocha, Boyana Norris, Steven G. Parker, Jaideep Ray, Sameer Shende, Theresa L. Windus, and Shujia Zhou. A component architecture for high-performance scientific computing. *The International Journal of High Performance Computing Applications*, 20(2):163–202, 2006.

[8] Joseph P. Kenny, Curtis L. Janssen, Edward F. Valeev, and Theresa L. Windus. Components for integral evaluation in quantum chemistry. *Journal of Computational Chemistry*, 29(4):562–577, 2008.

[9] Ryan M. Richard, Kristopher Keipert, Jonathan Waldrop, Murat Keçeli, David Williams-Young, Raymond Bair, Jeffery Boschen, Zachery Crandall, Kevin Gasperich, Quazi Ishtiaque Mahmud, Ajay Panyala, Edward Valeev, Hubertus van Dam, Wibe A. de Jong, and Theresa L. Windus. PluginPlay: Enabling exascale scientific software one module at a time. *The Journal of Chemical Physics*, 158(18):184801, 05 2023.

[10] Daniel G. A. Smith, Annabelle T. Lolinco, Zachary L. Glick, Jiyoung Lee, Asem Alenaizan, Taylor A. Barnes, Carlos H. Borca, Roberto Di Remigio, David L. Dotson, Sebastian Ehlert, Alexander G. Heide, Michael F. Herbst, Jan Hermann, Colton B. Hicks, Joshua T. Horton, Adrian G. Hurtado, Peter Kraus, Holger Kruse, Sebastian J. R. Lee, Jonathon P. Misiewicz, Levi N. Naden, Farhad Ramezanghorbani, Maximilian Scheurer, Jeffrey B. Schriber, Andrew C. Simmonett, Johannes Steinmetzer, Jeffrey R. Wagner, Logan Ward, Matthew Welborn, Doaa Altarawy, Jamshed Anwar, John D. Chodera, Andreas Dreuw, Heather J. Kulik, Fang Liu, Todd J. Martínez, Devin A. Matthews, Henry F. Schaefer, Jiří Šponer, Justin M. Turney, Lee-Ping Wang, Nuwan De Silva, Rollin A. King, John F. Stanton, Mark S. Gordon, Theresa L. Windus, C. David Sherrill, and Lori A. Burns. Quantum chemistry common driver and databases (QCDB) and quantum chemistry engine (QCEngine): Automation and interoperability among computational chemistry programs. *The Journal of Chemical Physics*, 155(20):204801, 2021.

IOP Roadmap: Education and Training in Electronic Structure Theory: Navigating an Evolving Landscape

Wibe Albert de Jong

Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Jessica A. Nash

The Molecular Sciences Software Institute, Virginia Tech University, Blacksburg, VA, USA

Nicola Seriani

Condensed Matter and Statistical Physics Section, The Abdus Salam ICTP, Trieste, Italy

Status

Experts building software for electronic structure based simulations in chemistry and materials are becoming a rare breed. The decline in the number of chemists and materials scientists with strong fundamentals in mathematics, electronic structure theory, and programming complex computer hardware is driven by the current undergraduate curriculum [1], and by students seeing limited career opportunities beyond their college education with the acquired skills.

Until the early 1990s programming was part of the undergraduate curriculum. To run an electronic structure simulation on a computer required one to have a good understanding of the programming environment and hardware. While basic programming skills were taught, many computational chemistry, materials and physics researchers were self-taught advanced programmers learning the latest programming languages and parallel computing.

Nowadays, most undergraduate degrees do not require a course in programming. Instead, the priority is given to using electronic structure software to teach chemistry and physics. Electronic structure-based simulations are used as virtual experiments.

This transition from teaching students to write programs to using programs started in the mid 1990s when many electronic structure software packages and high-level tools (think Python) started to become readily available.

This transition has led to the chemistry, materials, and physics community losing expertise in programming and the fundamentals of the methods underpinning electronic structure simulations. Push button electronic structure simulations allow practitioners to get away with not understanding the principles behind basis sets, pseudopotentials, exchange-correlation functionals in density functional theory, or electron correlation. All of this has led to issues with reproducibility and coding crisis [2].

These days, most groups developing electronic structure software teach their students limited programming skills to get the job done. The result is poor software coding practices. Many electronic structure software are developed over decades, with programming languages that were current at the time for hardware that existed at that time. With languages changing, few learn FORTRAN anymore and instead focus on C++ and Python, new hardware technologies that have been appearing, for example the GPU and increased parallel complexity, maintaining and expanding electronic structure software with new methods is becoming an increasingly daunting task for research groups.

More recently, efforts have been started to address the current state of affairs, with various efforts underway to address the challenges educating our next generation of electronic structure software developers and practitioners.

Current and Future Challenges

The main current challenge in the field of education regards the abundance of skills to learn. Traditionally, scientists in the field have had a solid background in mathematics, theory of electronic structure, and computer science. The set of skills is however expanding, as the theory is becoming more sophisticated, just as hardware and software. In the electronic structure theory, developments regard e.g. wavefunction topology and geometry, requesting access to mathematical concepts that are not traditionally in the toolbox of the practitioner in the field. On the side of computer skills, architectures used for computing are changing nature, shifting towards GPU and cloud-based computing. These issues are likely to become more pressing in the near future, also due to the rise of machine learning and quantum computing. On top of that, the success of electronic structure methods means that they are now employed in a variety of fields from molecular biology to functional materials, each requiring its own set of competencies.

As a result of the issues described above, nowadays the average scientist has become mostly a user of existing codes, leading to an employment of codes as black boxes in the worst cases. While the software has reached maturity allowing for semi-automatic usage, limitations of the methods make it however necessary for the user to have a deep understanding of theory and algorithms to interpret correctly the simulation results.

Finally, the impact of these problems is dampened in environments where students are in close contacts with experts in the field. On the contrary, problems are exacerbated when this environment is not present, as it is often the case in disadvantaged locations in developing countries [3]. It may then happen that people access to the codes and to on-line teaching material, but fail to reach the needed level of control on the simulations. Recent increase in on-line interactions due to the pandemic have been useful, but they cannot substitute sustained direct interaction. In this context, initiatives that foster direct interaction are crucial; in Africa, this role is taken by the African School on Electronic Structure Methods and Applications (ASESMA) [4].

Advances in Science and Technology to Meet Challenges

Though education in electronic structure software faces the many challenges outlined in the previous paragraphs, educators and researchers are innovating to meet these needs. The electronic structure community, and the wider computational molecular sciences community, are developing specialized training programs, utilizing open-source software, and increasing students' computational skills by integrating programming into the science curriculum. The field is also shifting to recognize research software development as its own scientific sub-discipline. Overall, many steps are being taken to improve computational competency.

A path the programming and computational proficiency can be established during the undergraduate career by integrating programming into existing science classes. Some examples of the incorporation of programming skills in chemistry classrooms were recently highlighted in the American Chemical Society Symposium Series book "Teaching Programming across the Chemistry Curriculum."[5]

Outside the formal curriculum, training efforts also occur through institutions, educational organizations, and specialized programs such as workshops and summer schools. One effort in this area is The Molecular Sciences Software Institute (MolSSI) [6]. To meet its goal of educating early-career researchers in programming and software development, MolSSI develops tutorials in programming, software design, and high-performance computing and holds workshops and summer schools. Other community efforts such as Psi4Education[7] provide educators with ready-made electronic structure and programming educational materials for classroom use. Resources like these are usually available online, making them accessible to a broad audience.

Learning materials and scientific communities are continually becoming more interconnected and accessible. Recent increases in the computing power of the average personal computer, coupled with the prevalence of free and open-source software in computational chemistry, make electronic structure education using personal computers possible [8]. Additionally, web-based computational platforms allow access to educational materials and computing resources and only require a computer with a web browser. Examples of cloud computing platforms for computational chemistry include chemcompute[9] and nanoHUB[10]. Both provide browser-based programming and computational environments and have several lessons introducing chemical concepts.

Within the scientific community, there is increasing recognition of software development as a scientific career. In Europe and the United States, there are organizations for Research Software Engineers, professionals who primarily write and maintain scientific software. A growing number of degrees also focus on high-performance computing or scientific software specifically.

More work may be needed, particularly at the undergraduate level, to establish core programming and computational competency. Programming and data analytics are increasingly a skill need by all scientists, not only computational scientists. Transformative change may only come through large scale efforts and continued shifts in educational culture and practices. However, the current developments and practices described in this section represent steps in the right direction toward increased programming and scientific competency.

Concluding Remarks

While the community has dealt for decades with a shortage of scientists with the broad, but much needed, expertise in mathematics, electronic structure theory, and the ability to program the latest computer highly parallel technologies able to build software for electronic structure-based simulations in chemistry and materials, the tides are changing. Academic institutions are recognizing the need to develop students that are ready for the multidisciplinary world that combines high-performance computing, software engineering, electronic structure, and data analytics, and are developing new ways to teach students the essential skills needed. In addition to formal educational degree programs, many organizations are taking full advantage of online training technologies to build broadly accessible curricula. The growing need for graduates with strong computational and analytical backgrounds in academia, national laboratories, and industry are making software development as a career an increasingly appealing choice for new students. In short, things are looking up.

Acknowledgements

WAdJ acknowledges support from the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration. J.A. Nash acknowledges support from The Molecular Sciences Software Institute, funded by the National Science Foundation under grant CHE-2136142.

References

[1] Ashley Ringer McDonald. *Teaching Programming across the Chemistry Curriculum: A Revolution or a Revival?*, chapter 1, pages 1–11.

[2] Jamie Durrani. Computational chemistry faces a coding crisis, July 2020. [Online; posted 1-July-2020].

[3] Nithaya Chetty and Richard Martin. Schooling Africa: Computational materials science education and research. *Journal of Physics: Conference Series*, 1512(1):012042, 2020.

[4] George Amolo, Nithaya Chetty, Ali Hassanali, Daniel Joubert, Richard Martin, and Sandro Scandolo. Growing materials science in Africa – the case of the African School for Electronic Structure Methods and Applications (ASESMA). *MRS Advances*, 3:2183, 2018.

[5] Ashley Ringer McDonald and Jessica A Nash, editors. *Teaching programming across the chemistry curriculum.* ACS SYMPOSIUM SERIES. Oxford University Press, New York, NY, August 2022.

[6] Jessica A. Nash, Mohammad Mostafanejad, T Daniel Crawford, and Ashley Ringer McDonald. Molssi education: Empowering the next generation of computational molecular scientists. *Computing in Science and Engineering*, 24(3), 2022.

[7] D Brandon Magers, Victor H Chavez, Benjamin G Peyton, Dominic A Sirianni, Ryan C Fortenberry, and Ashley Ringer McDonald. *PSI4EDUCATION: Free and open-source programing activities for chemical education with free and open-source software*, chapter 8, pages 107–122. ACS Publications, 2021.

[8] Susi Lehtola and Antti J. Karttunen. Free and open source software for computational chemistry education. *WIREs Computational Molecular Science*, page e1610.

[9] M. J. Perri and S. H. Weber. Web-based job submission interface for the gamess computational chemistry program. *Journal of Chemical Education*, 91(12):2206–2208, 2014.

[10] Krishna Madhavan, Lynn Zentner, Victoria Farnsworth, Swaroop Shivarajapura, Michael Zentner, Nathan Denny, and Gerhard Klimeck. nanohub.org: cloud-based services for nanoscale modeling, simulation, and education. *Nanotechnology Reviews*, 2(1):107–117, 2013.

Electronic structure theory facing industry and realistic modeling of experiments

Ryoji Asahi[1], Benoit Van Troeye[2], Pansy D. Patel[3], Geoffrey Pourtois[2]
[1]Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603, Japan
[2]imec, 75 Kapeldreef – 3001 Leuven
[3]PPG Industries, Monroeville Chemicals Center, 440 College Park Drive
Monroeville, PA, USA, 15146

Status

Modelling and simulation at the atomistic level are well-recognized tools in materials research to understand microscopic and macroscopic phenomena, as well as to guide the design of new molecules and materials. Among them, first-principles techniques based on Density Functional Theory (DFT), have become the norm thanks to their success to reproduce experimental observations in a broad range of systems and thanks to their transferability. The rapid developments of computational power and algorithms in the recent years have allowed a systematic investigation of different materials, input parameters, and environmental factors leading to efficient optimization and device refinement that cannot be obtained through experimentation alone. Furthermore, these developments have permitted the calculations of computationally intensive properties like electron-phonon-limited electrical conductivity and to push simulated system sizes towards the mesoscopic scale (see schematic representation of Fig. 1).

Current and Future Challenges

Accurate, versatile, and transferable atomistic simulation methods like DFT typically come at the expense of high computational cost as compared to simpler analytical descriptions. Computationally-accessible simulation size (~100-10000 atoms from regular DFT [1] to linear scaling methods[2]) and simulated time (<~100 ps) limit its applicability to the description of simple qualitative problems (well-defined surface reactions or the prediction of crystal phases). Therefore, predicting the evolution of extensive and complex physical/chemical/photoexcitation processes remains challenging with current state-of-the-art simulations.

As an additional layer of complexity, functional materials are multi-layered, macromolecular, multi-phasic, polycrystalline, alloyed, amorphous, or a combination thereof. Their descriptions require the development of accurate, but computationally low-cost methodologies to account for entropy and efficient statistical sampling. Furthermore, interfacial aspects in which strain, diffusion, and electrostatics need to be included to model the overall macroscopic properties. These aspects are conditioned by the chemical and thermal history and by the absence of explicit chemical reservoirs to allow degassing in materials, the introduction of stabilizing chemical reactants, or the activation of dislocation, etc. in simulations. Even when these problems are addressed, the translation of material properties into macroscopic observables remains challenging. Atomistic structures must be currently coupled to continuum formalisms, (semi-)classical condensed matter transport, or mesoscopic physics to enable multi-scale/physics modelling.

In recent years, materials informatics (MI) have been gaining momentum by applying machine learning (ML) to molecular/materials research. For such a data-driven exploration to be effective, systematic, and extensive collection of research data is necessary. Due to the limited available experimental data, the use of DFT calculations is extremely effective for data generation. In that regard, several initiatives have led to the development of global standard open repositories [3]-[4], that give one access to considerable amount of data to perform ML and screening. However, neither experimental nor computational data are sufficient for data-driven material exploration. Here, open/closed data strategies are being explored between academia and industry to alleviate intellectual property access limitations. The challenge is to enable a federated ML platform[5] with enough incentives and/or beneficial services for this to happen. The ML development would help in the development of intuitive workflow, a streamlined industry-leading predictive model with iterative data retrieval and transfer to achieve program objectives for early material candidates.

Advances in Science and Technology to Meet Challenges

Addressing the listed challenges requires the development of new algorithms and numerical approaches to reduce the computational cost. Algorithmics combined with alternative basis sets (local or mixed, like in the projector augmented-wave method), low cut-off pseudopotentials, software profiling and/or direct minimization scheme of the wavefunction have been paving the way to unlock these challenges, but should not come at the expense of transferability and accuracy. In that regard, recent initiatives have push for a more systematic assessment of the transferability of pseudopotentials with respect to all-electron computations [6,7].

Next, the DFT accurate description of physical properties relies heavily on the choice of exchange correlation (Exc) functional. Though being the object of intense investigations, Exc functionals that can accurately describe all properties of technological interest (electronic gap, mobilities, non-colinear spins coupled with magnetism...) have not been identified yet [8]. The same considerations hold for the time-dependent DFT kernel. In parallel, extending DFT to treat new perturbations, like external magnetic or chemical gradient, as well as considering the lattice temperature (electron-phonon and phonon-phonon coupling), would help driving numerous technological progresses.

In parallel, the introduction of ML is having a significant impact on enabling multiscale/large-scale simulations. The DFT data can be learned by ML to construct force-field (or ML potentials) with a minimum of DFT calculations, which flexibly adopt to the target model [9]. Once trained, a ML potential enables the evaluation of physical properties with a computational efficiency close to that of classical force fields, but with accuracy comparable to first-principles calculations. The ML algorithm could also be applied to generate efficient high-level quantum-mechanical calculations such as the coupled-clulster method and the random phase approximation by training on the difference from the DFT level. [10,11] As an example of application in industry, first-principles computations have recently been combined with ML to drive the discovery of new ovonic threshold switching materials for non-volatile resistive random-access memories [12,13]. The most-promising identified materials are then synthesized and tested internally to drive technological progresses.

Finally, the inherent errors related to low-cost computational techniques can be improved with the development of computational methods. A more recent approach is the use of data assimilation, which combines simulation and experimental data and used originally in meteorology to improve the accuracy of simulations by using measured data. This technique is now being applied to the field of materials science to find correlations between simulations and experimental data, as well as for prediction [14].

Concluding Remarks

Software has become increasingly complex with the evolution of materials science and the dramatic changes in societal needs. To continue the development, it is necessary to have a platform that allows subject matter experts to participate on a global scale rather than conventional development on an individual basis. Here it will also become increasingly important to develop human resources with knowledge of informatics in addition to fundamental physics and chemistry in materials science.

Acknowledgements

The authors would like to thank Michiel J. van Setten, Kiroubanand Sankaran, Sergiu Clima for their contributions and the Imec Industrial Affiliation Program (IIAP) for funding. The work was also partially supported by the JSPS Grant-in-Aid for Transformative Research Areas (A) (21H05560 and 23H04105).

References (separate from the two page limit)

[1] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P Seitsonen, A. Smogunov, P. Umari and R.M. Wentzcovitch, “QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials”, J. Phys.: Condens. Matter **21**, p. 395502 (2009).

[2] A. Nakata, J. S. Baker, S. Y. Mujahed, J. T. L. Poulton, S. Arapan, J. Lin, Z. Raza, S. Yadav, L. Truflandier, T. Miyazaki and D.R. Bowler, “Large scale and linear scaling DFT with the CONQUEST code”, J. Chem. Phys. **152**, p. 164112 (2020).

[3] https://nomad-lab.eu

[4] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, “The Materials Project: A materials genome approach to accelerating materials innovation”, APL Materials **1**, p. 011002 (2013).

[5] J. Konečný, B. McMahan, D. Ramage, “Federated Optimization: Distributed Optimization Beyond the Datacenter”. arXiv:1511.03575, (2015).

[6] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S.J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J.K. Dewhurst, I. Di Marco, C. Draxl, M. Dułak, O. Eriksson, J.A. Flores-Livas, K.F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grånäs, E. K. U. Gross, A. Gulans, F. Gygi, D.R. Hamann, P.J. Hasnip, N.A.W. Holzwarth, D. Iuşan, D.B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y.O. Kvashnin, I.L.M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C.J. Pickard, W. Poelmans, M.I.J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M.J. van Setten, V. Van Speybroeck, J.M. Wills, J.R. Yates, G.-X. Zhang, S. Cottenier, “Reproducibility in density functional theory calculations of solids”, Science 351, p. 6280 (2016)

[7] E. Maggio, P. Liu, M. J. van Setten, and G. Kresse “GW100: A Plane Wave Perspective for Small Molecules”, *Journal of Chemical Theory and Computation **13***, p. 635 (2017).

[8] N. Mardirossian and M. Head-Gordon, “Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals”, Molecular Physics **115**, p. 2315 (2017).

[9] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, “Machine Learning Force Fields”, Chem. Rev. **121**, p. 10142 (2021).

[10] M. Bogojeski, L. Vogt-Maranto, M.E. Tuckerman, K.-R. Müller, and K. Burke, “Quantum chemical accuracy from density functional approximations via machine learning”, Nat. Commun. **11**, p. 5223 (2020).

[11] P. Liu, C. Verdi, F. Karsai, and G. Kresse, “Phase transitions of zirconia: Machine-learned force fields beyond density functional theory”, Phys. Rev. B, **105**, L060102 (2022).

[12]S. Clima, T. Ravsher, D. Garbin, R. Degraeve, A. Fantini, R. Delhougne, G. Sankar Kar, and G. Pourtois, “Ovonic Threshold Switch Chalcogenides: Connecting the First-Principles Electronic Structure to Selector Device Parameters”, ACS Applied Electronic Materials **5**, p. 461 (2023).

[13]S. Clima, D. Matsubayashi, T. Ravsher, D. Garbin, R. Delhougne, G. Sankar Kar and G. Pourtois, “In silico screening for As/Se-free ovonic threshold switching materials”, npj Computational Materials **9**, p. 96 (2023)

[14]A. Karpatne, G. Atluri, J.H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar, “Theory-guided data science: A new paradigm for scientific discovery from data”, IEEE Transactions on Knowledge and Data Engineering **29**, p. 2318 (2017).

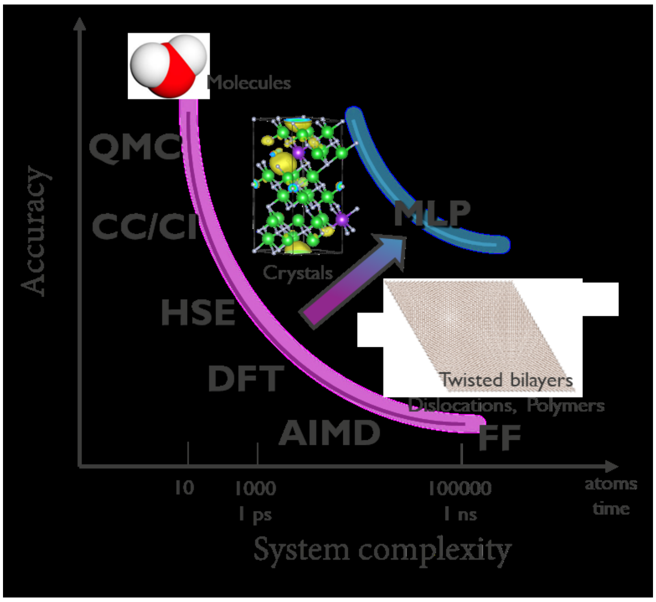


Figure I. State-of-the-art atomistic simulations with current limitations and future directions as already observed by employing the machine learning potential (MLP). Note that QMC, CC, CI, HSE, AIMD, and FF represent Quantum Monte Carlo, Coupled Cluster, Configuration Interaction, Hybrid Functional, Ab-Initio Molecular Dynamics, and Force Field, respectively.

List of Acronyms:

Acronym	Meaning
AFIR	artificial force induced reaction
AIMD	ab-Initio molecular dynamics
API	application programming interface
ARPES	angle-resolved photoemission spectroscopy
ASESMA	African School on Electronic Structure Methods and Applications
BBGKY	Bogoliubov-Born-Green-Kirkwood-Yvon
BO	Born- Oppenheimer
BSE	Bethe-Salpeter equation
CASSCF	complete active space self-consistent field
CC	coupled cluster
CCA	Common Component Architecture
CI	configuration interaction
CMS	content management system
CPU	central processing unit
DFT	density functional
DFT	density functional theory
DFTB	density functional tight binding
DMC	diffusion Monte Carlo
DMFT	dynamical mean-field theory
EELS	electron energy loss spectroscopy
EHT	extended Hückel theory
ES	electronic structure
Exc	exchange correlation
FAIR	Findable, Accessible, Interoperable and Reusable
FCI	full configuration interaction
FF	force field
FLOP	floating point operation
FPGA	field programmable gate array
GF	Green’s function
GGA	generalized gradient approximation
GPU	graphical processing unit
GUI	graphical user interface
HEG	homogeneous electron gas
HF	Hartree-Fock
HPC	high-performance computing
KS	Kohn-Sham

LDA	local-density approximation
LOSC	local-orbital scaling correction
LR	linear-response
MBPT	many-body perturbation theory
MCSCF	multi-configurational self-consistent-field
MCTDH	multiconfiguration time-dependent Hartree
MD	molecular dynamics
mGGA	meta-generalized gradient approximation
MI	materials informatics
ML	machine learning
MolSSI	Molecular Sciences Software Institute
MPI	modular path integral
NDDO	neglect of diatomic differential overlap
NEB	nudged elastic band
NEO	nuclear-electronic orbital
OEP	optimized effective potential
OLED	organic light emitting diodes
PBE	Perdew–Burke–Ernzerhof
PES	potential energy surface
PI	path integral
QC	quantum chemistry
QCPI	quantum-classical path integral
QED	quantum electrodynamics
QMC	Quantum Monte Carlo
QP	quasiparticle
QPU	quantum processing unit
QuAPI	quasi-adiabatic propagator path integral
RDF	reduced density matrix
RIXS	resonant inelastic x-rays scattering
RT-TDDFT	real-time time-dependent density functional theory
SaaS	Software-as-a-Service
SH	surface hopping
SIC	self-interaction correction
SMatPI	small matrix decomposition path integral
SOC	spin-orbit coupling
TD	time-dependent
TDHF	time-dependent Hartree-Fock
TS	transition state
VMC	variational Monte Carlo

WFT	wave-function-based electronic structure theory
X2C	exact two-component
XANES	x-ray absorption spectroscopy
XC	exchange-correlation
XDM	exchange-dipole moment
XX	exact exchange
ZDO	zero-differential overlap
ZPE	zero-point energy