ADVANCES IN DENSITY FUNCTIONAL THEORY AND BEYOND FOR COMPUTATIONAL CHEMISTRY

EDITED BY: Wei Hu and Mohan Chen
PUBLISHED IN: Frontiers in Chemistry

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-300-4
DOI 10.3389/978-2-88971-300-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ADVANCES IN DENSITY FUNCTIONAL THEORY AND BEYOND FOR COMPUTATIONAL CHEMISTRY

Topic Editors:
Wei Hu, Lawrence Berkeley National Laboratory, Italy
Mohan Chen, Peking University, China

Citation: Hu, W., Chen, M., eds. (2021). Advances in Density Functional Theory and Beyond for Computational Chemistry. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-300-4

Table of Contents

frontiers in Chemistry

EDITORIAL
published: 12 July 2021
doi: 10.3389/fchem.2021.705762

Editorial: Advances in Density Functional Theory and Beyond for Computational Chemistry

Wei Hu[1,2]* *and Mohan Chen*[3]*

[1]Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, China, [2]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [3]HEDPS, CAPT, College of Engineering, Peking University, Beijing, China

Keywords: density functional theory, low-scaling algorithms, machine learning, excited-state, high-performance computing

Editorial on the Research Topic

Advances in Density Functional Theory and Beyond for Computational Chemistry

The rapid development of modern computational chemistry has led to a growing need to understand the microscopic mechanisms determining the properties of molecular and solid materials at an atomic level. The interactions between atoms and electrons are governed by the laws of quantum mechanics; hence, accurate and efficient computational methods for solving the quantum-mechanical equations are needed. The Kohn-Sham density functional theory (DFT) Hohenberg and Kohn (1964), Kohn and Sham (1965) marks a decisive breakthrough in these efforts, and in the past few decades DFT has made an unparalleled impact on a variety of interesting and challenging problems in computational chemistry. The real forte of DFT is its favourable price and performance ratio as compared with electron-correlated wave-function-based methods, such as the Møller–Plesset perturbation theory Binkley and Pople (1975) or coupled cluster theory Čížek (1966). Thus, large-scale molecular and solid systems can be studied by DFT with sufficient accuracy, thereby expanding the predictive power inherent in electronic structure theory. As a result, DFT is now by far the most widely used electronic structure method. Although 50 years have passed since the formulation of the Kohn-Sham DFT, many open questions remain, including the mathematical issues in solving the Kohn-Sham equations, the developments of more accurate and efficient density functionals, and applying the DFT calculations to solve more scientific problems. This research topic focuses on covering recent advances within the framework of DFT.

Computational chemistry methods have become increasingly important in recent years, as manifested by their rapidly extending applications in a large number of diverse fields, such as computations of molecular structures and properties, the design of pharmaceutical drugs and novel materials, etc. In part as a result of this general trend, the size of the systems which can be computationally studied has also increased, generating even further needs for large-scale applications. This is because larger molecular systems show interesting phenomena and have important implications in modern biochemistry, biotechnology, and nanotechnology. Thus, it is of great importance to apply and further develop computational methods which provide physically sound models for large molecules at a reasonable computational cost. A representative approach is the linear scaling technique Goedecker (1999), which owns a computational cost that scales linearly $\mathcal{O}(N)$ with the size of the system. The linear-scaling DFT is an area of active research in computational chemistry, with the performances improve steadily over the years, especially on parallel high-performance machines. Historically, linear-scaling implementations were restricted to basic ground state energy and electron density calculations, but this has also improved in recent years

OPEN ACCESS

Edited by:
Thomas S. Hofer,
University of Innsbruck, Austria

Reviewed by:
Alexander Novikov,
Saint Petersburg State University, Russia

Correspondence:
Wei Hu
whuustc@ustc.edu.cn
Mohan Chen
mohanchen@pku.edu.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *06 May 2021*
Accepted: *25 June 2021*
Published: *12 July 2021*

Citation:
Hu W and Chen M (2021) Editorial: Advances in Density Functional Theory and Beyond for Computational Chemistry. Front. Chem. 9:705762. doi: 10.3389/fchem.2021.705762

with geometry optimizations and molecular dynamics (MD) becoming available. Moreover, recent developments of machine learning algorithms enable the large-scale MD simulations with *ab initio* accuracy, and has been applied to a variety of applications Jia et al. (2020). This research topic aims to report the state-of-the-art computational methods in several of the important questions related to the family of linear scaling methods.

A deep understanding of the excitations in molecules and solids are of fundamentally importance in many technological applications. There is already a rich set of theoretical and simulation methods for excited-state calculations, such as the GW plus Bethe-Salpeter equation Hedin (1965), time-dependent density functional theory (TDDFT) Runge and Gross (1984) and many-body coupled cluster (CC) theory Čížek (1966). Unfortunately, these post-Hartree-Fock and excited state methods for electronic excitations are all subject to computational bottlenecks, which are far more severe than those affecting the standard calculations of the ground-state energy, not only because of the system size, but also because the large number of excited states that need to be considered. A major difficulty for treating excited complex systems arises from the different nature of the various competing excited electronic states. For example, the localized neutral and delocalized charge transfer excitons, as a result of the relatively large length scale. Therefore, this research topic also aims to cover developments of novel electronic structure algorithms and scalable computational methods for excited states of complex systems.

The past several decades have witnessed tremendous strides in the capabilities of computational chemistry simulations, driven in large part by the extensive parallelism offered by powerful computer clusters and scalable programming methods on high performance computing (HPC) Hu et al. (2021), Kowalski et al. (2021). However, such massively parallel simulations increasingly require more advanced algorithms to achieve satisfactory performance across the vastly diverse ecosystem of modern heterogeneous computer systems. The design of efficient parallel codes proves to be difficult: the diversity of involved data structures and algorithms, as well as the frequently occurring inherent sequential control propose enormous challenges to efficiently use of a large number of processors. This research topic also focuses on the developments of more effective computational methods by use of high performance parallel computing.

This editorial sums up the contents of our Research Topic "Advances in Density Functional Theory and Beyond for Computational Chemistry" and a total of nine original research contributions have been included in this article collection, involving linear-scaling density functional theory, multiple scattering theory, ab initio molecular dynamics, deep potential model, hybrid and double-hybrid functional theory, second-order Møller–Plesset perturbation theory, coupled cluster theory and high performance computing.

Linear-scaling DFT Goedecker (1999) is an efficient method to yield the structural and electronic properties of molecules, semiconductors, and insulators to avoid the high cubic-scaling cost in conventional DFT calculations. Luo et al. described an efficient parallel implementation of linear-scaling density matrix trace correcting purification algorithm Niklasson (2002) to solve the Kohn–Sham equations with numerical atomic orbitals in the HONPAS Qin et al. (2015) package. The authors have performed large-scale DFT calculations on boron nitrogen nanotubes containing tens of thousands of atoms, which can scale up to hundreds of processing cores on modern heterogeneous supercomputers.

The Korringa–Kohn–Rostoker Green's function method Korringa (1947), Kohn and Rostoker (1954), also known as multiple scattering theory (MST) Lloyd and Smith (1972), provides equivalent information as solving the Kohn-Sham equation by employing the single-particle Green's function Economou (2006). Cao et al. investigated a reduced scaling full-potential DFT method based on the multiple scattering theory code MuST Rusanu et al. (2011). A significant advantage of the MST method is the reduced scaling in the calculations of metallic systems. The MST method shows the potential to simulate more complicated materials on massively parallel supercomputers and provides a reliable and accessible way to large-scale first-principle simulations of metals and alloys.

AIMD (*ab initio* molecular dynamics) has been extensively employed to explore the dynamical information of electronic systems. However, it remains extremely challenging to reliably predict electronic properties of systems with a radical nature using first-principles DFT calculations due to the presence of the static correlation. To address this challenge, Li and Chai proposed TAO-DFT (thermally-assisted-occupation density functional theory) with AIMD to explore the dynamical properties of nanosystems with a radical nature at finite temperatures. A variety properties of n-acenes (n = 2–8) at 300 K are presented including the instantaneous/average radical nature and infrared spectra of n-acenes containing n linearly fused benzene rings (n = 2–8).

Predicting crystal structure has been a challenging problem, which requires a reliable energy calculation engine and an efficient global search algorithm. Machine learning based inter-atomic potential energy surface models such as the deep potentials Jia et al. (2020) owns the DFT accuracy and the speed of empirical force fields and can be used as an energy calculator. Wang et al. employed the deep potential model to predict the intermetallic compound of the aluminum–magnesium system, and found six meta-stable phases with negative or nearly zero formation energy. The authors proposed a relatively robust structure screening criterion that selects potentially stable structures from the Deep Potential-based convex hull and performs DFT refinement. By using this criterion, the computational cost needed to construct the convex hull with ab initio accuracy can be dramatically reduced.

Accurate prediction of quasiparticle and excitation energies has been very challenging for ground-state density functional methods since the commonly adopted density functional approximations suffer from the delocalization error. Yang et al. propsed a new method which presumed a quantitative correspondence between the quasiparticle energies and the generalized Kohn–Sham orbital energies. Furthermore, the authors employed a previously developed global scaling

correction approach to achieve substantially improved prediction of molecular quasiparticle and excitation energies.

Interpretation of spectroscopic experiments is challenging because the results are affected by the interplay of stereo-electronic, dynamic and environmental effects. The work of Barone et al. showed that the last-generation hybrid and double-hybrid functionals Biczysko et al. (2010), which are described by partially augmented double- and triple-zeta basis sets, provided unprecedented accuracy for medium-size semi-rigid molecules under the framework of the second order vibrational perturbation theory.

The second-order Møller–Plesset perturbation theory (MP2) Binkley and Pople (1975) is a post-Hartree–Fock approach to taking the electron correlation effect into account. Despite its simple form, the MP2 method captures around 90% of the correlation energy Bartlett and Stanton (1994), but the $\mathcal{O}(N^5)$ computational scaling of the original (canonical) MP2 method has limited the application of the MP2 method in large systems. Shang and Yang implemented the canonical and Laplace-transformed algorithms to calculate the MP2 perturbation theory for the total energy and the band gap of periodic systems under periodic boundary conditions in HONPAS Qin et al. (2015) code with numerical atomic orbitals. The MP2 correction energy and band gaps presented in the work are in excellent agreement with the results of the canonical MP2 formulation. Moreover, the authors studied the binding-energy curves for the two stacked transpolyacetylene chains and demonstrated that the new method well describe the correlation energy and the long-range van der Waals interactions.

The coupled cluster (CC) theory Čížek (1966) has become one of the most accurate ab initio methods to yield the electronic structure information. Yang et al. presented a Newton Krylov method Knoll and Keyes (2004) for solving the coupled cluster equation. This new method used a Krylov subspace iterative method, such as the Generalized Minimum Residual (GMRES) method Saad and Schultz (1986), to compute the Newton correction to the approximate coupled cluster amplitude. Numerical results demonstrate the effectiveness and robustness of the Newton Krylov method not only for standard CCSD calculations but also for tailed CCSD calculations where the information for external correction is obtained from a density matrix renormalization group (DMRG) calculation Schollwöck (2005).

Williams-Young et al. proposed a three-level parallelism scheme for the distributed numerical integration of the exchange-correlation potential in the Gaussian basis set discretization of the Kohn–Sham equations on large computing clusters consisting of multiple graphics processing units (GPU) per compute node. They demonstrated that the performance and scalability of the implementation of the purposed method in the NWChemEx Kowalski et al. (2021) software package by comparing to the existing scalable CPU exchange-correlation integration in NWChem. The results show that the speedups are between 10× and 100× as compared to the analogous CPU implementation in NWChem.

The above article collection demonstrates that the DFT methods have broad impacts on a variety of subjects in computational chemistry and related disciplines. In conjunction with high-performance computation and machine-learning techniques, the DFT framework undergoes another round of fast developments. It can be expected that more accurate DFT approaches with more efficient algorithms will be available in the near future.

AUTHOR CONTRIBUTIONS

The editorial was edited by WH and MC.

ACKNOWLEDGMENTS

The authors thank Honghui Shang and Jie Liu for their help in the article collection.

REFERENCES

Bartlett, R. J., and Stanton, J. F. (1994). Applications of Post-Hartree-Fock Methods: A Tutorial. *Rev. Comput. Chem.* 5, 65–169. doi:10.1002/9780470125823.ch2

Biczysko, M., Panek, P., Scalmani, G., Bloino, J., and Barone, V. (2010). Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theor. Comput.* 6, 2115–2125. doi:10.1021/ct100212p

Binkley, J. S., and Pople, J. A. (1975). Møller-Plesset Theory for Atomic Ground State Energies. *Int. J. Quan. Chem.* 9, 229–236. doi:10.1002/qua.560090204

Čížek, J. (1966). On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* 45, 4256–4266. doi:10.1063/1.1727484

Economou, E. N. (2006). *Green's Functions in Quantum Physics*, Vol. 7. Springer, Berlin, Heidelberg: Springer Science & Business Media. doi:10.1007/3-540-28841-4

Goedecker, S. (1999). Linear Scaling Electronic Structure Methods. *Rev. Mod. Phys.* 71, 1085–1123. doi:10.1103/RevModPhys.71.1085

Hedin, L. (1965). New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.* 139, A796–A823. doi:10.1103/PhysRev.139.A796

Hohenberg, P., and Kohn, W. (1964). Inhomogeneous Electron Gas. *Phys. Rev.* 136, B864–B871. doi:10.1103/physrev.136.b864

Hu, W., Qin, X., Jiang, Q., Chen, J., An, H., Jia, W., et al. (2021). High Performance Computing of DGDFT for Tens of Thousands of Atoms Using Millions of Cores on Sunway TaihuLight. *Sci. Bull.* 66, 111–119. doi:10.1016/j.scib.2020.06.025

Jia, W., Wang, H., Chen, M., Lu, D., Lin, L., Car, R., et al. (2020). Pushing the Limit of Molecular Dynamics with Ab Initio Accuracy to 100 Million Atoms with Machine Learning. *Proc. Int. Conf. High Perform. Comput. Networking, Storage Anal.* 5, 14. doi:10.1021/ct100212p

Knoll, D. A., and Keyes, D. E. (2004). Jacobian-free Newton-Krylov Methods: a Survey of Approaches and Applications. *J. Comput. Phys.* 193, 357–397. doi:10.1016/j.jcp.2003.08.010

Kohn, W., and Rostoker, N. (1954). Solution of the Schrödinger Equation in Periodic Lattices with an Application to Metallic Lithium. *Phys. Rev.* 94, 1111–1120. doi:10.1103/PhysRev.94.1111

Kohn, W., and Sham, L. J. (1965). Self-consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* 140, A1133–A1138. doi:10.1103/physrev.140.a1133

Korringa, J. (1947). On the Calculation of the Energy of a Bloch Wave in a Metal. *Physica* 13, 392–400. doi:10.1016/0031-8914(47)90013-X

Kowalski, K., Bair, R., Bauman, N. P., Boschen, J. S., Bylaska, E. J., Daily, J., et al. (2021). From NWChem to NWChemEx: Evolving with the Computational

Chemistry Landscape. *Chem. Rev.* 121, 4962–4998. doi:10.1021/acs.chemrev.0c00998

Lloyd, P., and Smith, P. V. (1972). Multiple Scattering Theory in Condensed Materials. *Adv. Phys.* 21, 69–142. doi:10.1080/00018737200101268

Niklasson, A. M. N. (2002). Expansion Algorithm for the Density Matrix. *Phys. Rev. B* 66, 155115. doi:10.1103/PhysRevB.66.155115

Qin, X., Shang, H., Xiang, H., Li, Z., and Yang, J. (2015). HONPAS: A Linear Scaling Open-Source Solution for Large System Simulations. *Int. J. Quan. Chem.* 115, 647–655. doi:10.1002/qua.24837

Runge, E., and Gross, E. K. U. (1984). Density-functional Theory for Time-dependent Systems. *Phys. Rev. Lett.* 52, 997–1000. doi:10.1103/PhysRevLett.52.997

Rusanu, A., Stocks, G. M., Wang, Y., and Faulkner, J. S. (2011). Green's Functions in Full-Potential Multiple-Scattering Theory. *Phys. Rev.* B 84, 035102. doi:10.1103/PhysRevB.84.035102

Saad, Y., and Schultz, M. H. (1986). GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.* 7, 856–869. doi:10.1137/0907058

Schollwöck, U. (2005). The Density-Matrix Renormalization Group. *Rev. Mod. Phys.* 77, 259–315. doi:10.1103/RevModPhys.77.259

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 23 October 2020
doi: 10.3389/fchem.2020.584203

Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals

Vincenzo Barone, Giorgia Ceselin, Marco Fusè and Nicola Tasinato*

SMART Laboratory, Scuola Normale Superiore di Pisa, Pisa, Italy

Accuracy and interpretability are often seen as the devil and holy grail in computational spectroscopy and their reconciliation remains a primary research goal. In the last few decades, density functional theory has revolutionized the situation, paving the way to reliable yet effective models for medium size molecules, which could also be profitably used by non-specialists. In this contribution we will compare the results of some widely used hybrid and double hybrid functionals with the aim of defining the most suitable recipe for all the spectroscopic parameters of interest in rotational and vibrational spectroscopy, going beyond the rigid rotor/harmonic oscillator model. We will show that last-generation hybrid and double hybrid functionals in conjunction with partially augmented double- and triple-zeta basis sets can offer, in the framework of second order vibrational perturbation theory, a general, robust, and user-friendly tool with unprecedented accuracy for medium-size semi-rigid molecules.

Keywords: quantum chemistry, density functional theory, rotational spectroscopy, vibrational spectroscopy, benchmark, atmospheric molecules, astrochemical molecules

OPEN ACCESS

Edited by:
Wei Hu,
Lawrence Berkeley National Laboratory, United States

Reviewed by:
Igor Ying Zhang,
Fudan University, China
Honghui Shang,
Chinese Academy of Sciences, China

****Correspondence:***
Vincenzo Barone
vincenzo.barone@sns.it

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *16 July 2020*
Accepted: *17 August 2020*
Published: *23 October 2020*

Citation:
Barone V, Ceselin G, Fusè M and Tasinato N (2020) Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. Front. Chem. 8:584203. doi: 10.3389/fchem.2020.584203

1. INTRODUCTION

Spectroscopic techniques are unique tools to non-invasively probe the properties of complex molecular systems in a variety of environments and conditions. In fact, the increasing sophistication of well-established techniques like nuclear magnetic and electron paramagnetic resonance (NMR and EPR), microwave (MW), infrared (IR), Raman, visible (Vis), ultra-violet (UV), or fluorescence and the parallel blooming of new ones, e.g., vibrational, electronic, and magnetic circular dichroism (VCD, ECD, and MCD), Raman optical activity (ROA), circularly polarized luminescence (CPL), multi-photon and time-resolved methods have a huge impact in several fields of science and technology (He et al., 2007; Barone, 2012; Berova et al., 2012; Tasinato et al., 2015; Sugiki et al., 2017; Lane, 2018). In addition to being widely used to infer information about molecular structure and dynamics in both gas and condensed phases (Sugiki et al., 2017; Puzzarini and Barone, 2018), spectroscopy allows for the unequivocal identification of chemical species in hostile environments, e.g., the interstellar space (Baiano et al., 2020), or in samples of unknown composition (He et al., 2007; Lindon et al., 2017; Lane, 2018) and plays a pivotal role in the study of photochemical mechanisms in biological systems and in the development of new technological devices, including photovoltaic cells, optoelectronic devices, eco-sustainable solutions, UV-resistant materials, dyes, and fluorescent probes (Berova et al., 2012; Drummen, 2012; Lindon et al., 2017; Lane, 2018). Unfortunately, interpretation of experimental data is often

a difficult task: the observed spectroscopic behavior results from the subtle interplay of stereo-electronic, dynamic and environmental effects, whose specific roles are difficult to disentangle. Furthermore, although conveying additional information, spectral congestion makes quantitative interpretation of experimental data even more difficult. To disentangle these complex signatures and disclose the underlying molecular properties, detailed molecular simulations are crucial (Barone, 2012). The last decade has witnessed an increasing interaction between experiment, theory, and simulation in the field of molecular spectroscopy and in all related applications. These have revealed the need of computational tools and theoretical methods not only to interpret the spectra, but also to design new experiments that would be impossible or very expensive to perform in a blind way. The widespread use of computational techniques in many areas of science and by an ever increasing number of non-experienced users (e.g., experimental chemists) has prompted the development, in our group, of the Virtual Multifrequency Spectrometer (VMS, http://dreamslab.sns.it/vms/) (Barone, 2016). VMS features two interconnected tools: VMS-Draw (Licari et al., 2015) and VMS-Comp (Barone et al., 2012), the former providing a user-friendly, graphical user interface to pilot the latter, which is in charge of computationally intensive tasks. VMS-Comp includes a wide set of algorithms and calculation options and it allows the user to predict with remarkable accuracy many types of spectroscopic data for a vast range of molecular systems and environments (Zerbetto et al., 2013; Licari et al., 2017; Presti et al., 2017). Further developments are, however, needed to deal with new and more sophisticated experimental techniques (Quack and Merkt, 2011; Lane, 2018). For small semi-rigid molecules, the accuracy of state-of-the-art quantum mechanical (QM) methods often rivals that of experimental techniques, but extension to large flexible systems (not to mention condensed phases) faces a number of difficulties (even within the Born-Oppenheimer approximation) ranging from the very unfavorable scaling of those methods with the number of active electrons to the proper description of flat potential energy surfaces (PESs) with stationary points that are ill defined (Puzzarini et al., 2019a).

A possible route to obtain accurate results, even for relatively large molecular systems (a few dozens of atoms), is provided by hybrid QM/QM′ models, which combine accurate quantum-mechanical (QM) calculations of "primary" properties (e.g., molecular structures or harmonic force fields) with cheaper yet reliable electronic structure approaches (QM′) for "secondary" properties (e.g., vibrational corrections or anharmonic effects). At the same time, computation of spectroscopic parameters often requires purposely tailored basis sets, whose selection must be based on extensive benchmarks. Finally, the customary rigid rotor (RR) / harmonic oscillator (HO) approximation is not sufficient for quantitative work and more advanced models must be employed. In the last few years, the second-order vibrational perturbation theory (VPT2) (Mills, 1972; Papoušek and Aliev, 1982; Aliev and Watson, 1985) has been exploited with considerable success for semi-rigid molecules of increasing dimensions. However, the identification of resonances in VPT2 treatments remains a daunting task due to the arbitrariness of their definition and their indirect influence on the energy and intensities. This leads to two distinct issues: finding the true resonances and then correcting them appropriately. The resonance conditions are strongly related to the quality of the electronic structure calculation, but they also depend on the coupling between the potentially resonant states. The most obvious solution is to combine perturbative and variational models, with the generalized (G)VPT2 model offering a very good accuracy-to-cost ratio not only for energies, but also for transition moments (Bloino et al., 2012). In recent years, we have implemented analytical second derivatives for double-hybrid functionals (Biczysko et al., 2010), general equations for Abelian and non-Abelian symmetry groups (Piccardo et al., 2015a), and full treatment of intensities for all conventional (IR, Raman) and chiral (VCD, ROA) vibrational spectroscopies up to three-quanta excitations (Bloino et al., 2015). These features compose, together with the particular care devoted to robustness, ease of use, and computational efficiency, the mandatory background for systematic evaluations of all the spectroscopic parameters beyond the rigid-rotor/harmonic-oscillator level for medium- to large-size semi-rigid molecules. Next, together with further developments for the treatment of flexible molecules (Puzzarini et al., 2019b), the selection of the most effective electronic structure models remains a central issue of computational spectroscopy. Here, the ongoing developments of methods rooted in the density functional theory (DFT) come into play. For microwave and vibrational spectroscopic applications, global hybrid density functionals (DFs), such as B3LYP (Lee et al., 1988; Becke, 1993) coupled to a polarized double-zeta basis set supplemented by a set of *sp* diffuse functions (hereafter B3) can deliver the accuracy required for the interpretation of the vibrational characteristics of medium and large molecules beyond the harmonic approximation for both transition frequencies and intensities. An increased accuracy, at the price of a more demanding computational loading, is brought by the double-hybrid B2PLYP functional (Grimme, 2006) in conjunction with partially augmented triple-ζ basis sets (hereafter B2). In fact, it has been shown that B2 calculations of vibrational frequencies and intensities can reach an average accuracy often within 10 cm^{-1} and a few km mol^{-1}, respectively (Biczysko et al., 2010; Puzzarini et al., 2019a; Boussessi et al., 2020a), thus performing equally to, or even better than, the CCSD(T)/cc-pVTZ approach. Concerning the prediction of rotational spectroscopic parameters, the same B2/B3 approach has been validated in several studies concerning rotational constants (Spada et al., 2017a; Li et al., 2018) together with quartic centrifugal distortion parameters (Tasinato, 2014; Boussessi et al., 2020a) and, very recently, also for sextic centrifugal distortion parameters (Pietropolli Charmet et al., 2017; Boussessi et al., 2020b).

In the present work, hybrid and double-hybrid density functionals of the "last-generation" are analyzed to check if they provide improved performances with respect to the B3 and B2 paradigms. This is particularly significant since the improved performances for thermochemistry, kinetics and non-covalent interactions could have been obtained at the price of

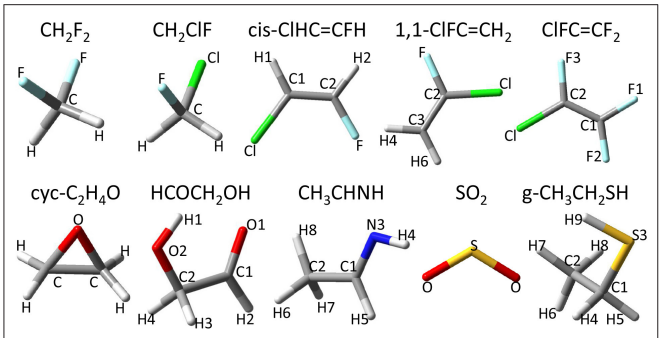

FIGURE 1 | Molecules in the benchmark set.

worsening other parameters of particular relevance for rotational and vibrational spectroscopy of closed- and open-shell species (Puzzarini et al., 2010). In parallel, the accuracy of different partially augmented correlation consistent basis sets is analyzed with the aim of defining the best choice in terms of the cost-to-accuracy ratio. We will consider, in particular, equilibrium geometries, ground state rotational constants and quartic centrifugal distortion constants, harmonic and anharmonic vibrational wavenumbers and IR intensities. The benchmark study is carried out on a set of ten molecules of atmospheric and astrochemical relevance, reported in **Figure 1**, which includes: difluoromethane (CH_2F_2) (Carlotti et al., 1988; Tasinato et al., 2012b; Piccardo et al., 2015b), chlorofluoromethane (CH_2ClF) (Blanco et al., 1995; Pietropolli Charmet et al., 2013), cis-1-chloro-2-fluoroethene (cis-ClHC=CHF) (Craig et al., 1970; Alonso et al., 1993; Gambi et al., 2002; Piccardo et al., 2015b), 1-chloro-1-fluoroethene ($ClFC=CH_2$) (Leung et al., 2009; Pietropolli Charmet et al., 2016; Gambi et al., 2019), chlorotrifluoroethene ($F_2C=CFCl$) (Hillig et al., 1988; Tasinato et al., 2012a), oxirane (cyc-C_2H_4O) (Russell and Wesendrup, 2003; Flaud et al., 2012; Medcraft et al., 2012; Lafferty et al., 2013; Puzzarini et al., 2014a; Piccardo et al., 2015b), glycolaldehyde ($HOCH_2CHO$) (Carroll et al., 2010; Johnson et al., 2013; Piccardo et al., 2015b; Boussessi et al., 2020a), E-ethanimine (CH_3CHNH) (Melli et al., 2018), sulfur dioxide (SO_2) (Flaud et al., 1993; Mller and Brnken, 2005; Tasinato et al., 2010; Boussessi et al., 2020a), and the gauche conformer of ethyl mercaptan (CH_3CH_2SH) (Smith et al., 1968; Wolff and Szydlowski, 1985; Miller et al., 2009; Kolesnikov et al., 2014; Puzzarini et al., 2014b; Hochlaf et al., 2015).

2. COMPUTATIONAL METHODOLOGY

Quantum chemical calculations at the DFT level were carried out using some hybrid and double-hybrid density functional approximations of the last generation, which are considered the best performing and transferable according to a very recent benchmark (Peveratti, 2020). Among hybrid functionals, the PW6B95 meta exchange-correlation functional proposed by Zhao and Truhlar (2005) and the ωB97 family of long-range corrected functionals introduced by Chai and Head-Gordon (Chai and Head-Gordon, 2008a,b), namely ωB97, ωB97X and ωB97X-D, were considered in conjunction with aug-cc-pVDZ (Dunning, 1989; Kendall et al., 1992; Woon and Dunning, 1993) and jul-cc-pVDZ (Papajak et al., 2011) double-ζ basis sets. For the PW6B95 functional, calculations were also carried out by using the jun-cc-pVDZ basis set (Papajak et al., 2011). The rev-DSD-PBEP86 double-hybrid density functional, recently proposed by Martin and coworkers (Santra et al., 2019), was employed together with the aug-cc-pVTZ and jun-cc-pVTZ basis sets. Indeed, triple-ζ basis sets in conjunction with the B2PLYP double-hybrid functional (Grimme, 2006) have been demonstrated to provide accurate predictions of geometries, rotational spectroscopic parameters and vibrational properties (Biczysko et al., 2010; Penocchio et al., 2015; Spada et al., 2017b; Tasinato et al., 2017; Boussessi et al., 2020a,b). Both PW6B95 and rev-DSD-PBEP86 were augmented for dispersion correlation through the Grimme's DFT-D3 scheme (Grimme et al., 2010) with Becke-Johnson damping (Grimme et al., 2011), even if the bare PW6B95 functional can already provide a satisfactory description of dispersion forces (Tasinato and Grimme, 2015). Since tight *d* functions are important for a quantitative representation of the electronic structure of second-row elements, partially augmented basis sets, namely aug-/jul-/jun-cc-pV(n+d)Z, including an additional set of *d* functions, were employed for sulfur and chlorine atoms. These basis sets were downloaded from the Basis Set Exchange library (Pritchard et al., 2019). At each level of theory, geometries were first optimized and then harmonic vibrational frequencies were computed by means of analytical Hessian matrices. While details for the calculation of analytical second-order derivatives of double-hybrid density functionals can be found in Biczysko et al. (2010), here it is only mentioned that their evaluation requires the full derivatives of the correlation contribution to the one-particle density matrix, γ^x. Its occupied-occupied and virtual-virtual blocks depend on the products of second-order perturbation amplitudes and amplitude derivatives, whereas the occupied-virtual block can be found from the solution of the so-called derivative *Z*-vector equations, that involve the derivatives of the MP2 Lagrangian. The cubic and semi-diagonal quartic force constants, and the second and third derivatives of the dipole moment surface were calculated by numerical differentiation of analytic quadratic force constants and dipole moment first derivatives, respectively. These quantities were then employed for the computation of rotational and vibrational spectroscopic parameters beyond the RR/HO approximation. In particular, quartic centrifugal distortion constants, vibrational contributions to rotational constants and vibrational frequencies were derived in the framework of second-order vibrational perturbation theory (Mills, 1972; Papoušek and Aliev, 1982; Barone, 2005; Bloino et al., 2012). In order to tackle the problem of resonances plaguing the anharmonic vibrational frequencies, the generalized second-order vibrational perturbation theory (GVPT2) was adopted, in which the (near-) singular terms are removed from the perturbative summations of anharmonicity constants and transition dipole moments (leading to the so called deperturbed approach, DVPT2) and the energy levels coupled by the resonances are treated in a second step by a proper

variational calculation of reduced dimensionality (Mills, 1972; Martin et al., 1995; Barone, 2005). All DFT calculations were carried out by using the Gaussian 16 suite of programs (Frisch et al., 2016), which was also employed for the perturbative treatment of the anharmonic force field in the framework of a general VPT2 engine. The latter, in addition to performing the calculation of anharmonic vibrational energies according to various flavors of VPT2 (namely, DVPT2, GVPT2 and the so-called hybrid-degeneracy corrected PT2, HDCPT2), allows the computation of transition integrals for a number of spectroscopic techniques (IR, Raman, VCD), from which the corresponding anharmonic transition intensities can be derived (Bloino et al., 2012, 2015). In addition, it should be noted that recently the VPT2 framework has been coupled to a 1-dimensional discrete-variable-representation (1D-DVR) approach for the treatment of molecular systems presenting one large-amplitude vibration (Baiardi et al., 2017), which has been applied to the simulation of the IR spectrum of the methyl-cyclopropenyl cation (Puzzarini et al., 2019b). The 1D-DVR method is currently implemented in the development version of the Gaussian code and it will be included in the next releases of the software, yet it is not required for simulating the spectroscopic properties of the molecules considered in this work, all being semi-rigid systems. Since rev-DSD-PBEP86 is not among the Gaussian built-in functionals, it has been defined by setting proper IOP flags on top of the DSD-PBEP86 functional.

3. RESULTS AND DISCUSSION

The performance of last-generation density functionals (DFs) for applications in the field of rotational and vibrational spectroscopy is investigated in relation to (a) equilibrium geometry, (b) rotational spectroscopic parameters, i.e., ground state rotational constants and quartic centrifugal distortion parameters, (c) harmonic vibrational frequencies and IR intensities, and (d) fundamental anharmonic wavenumbers and IR integrated absorption cross sections. The computed data are benchmarked against values determined experimentally and, in addition, they are also compared to high-level CCSD(T)-based results taken from the literature. Statistical indicators, such as mean deviation (MD), mean absolute deviation (MAD) and mean absolute percentage deviation (MAD%) are used to assess the accuracy of the different model chemistries. For comparison purposes, B3LYP/SNSD results as well as those obtained by the B2PLYP functional in conjunction with cc-pVTZ, aug-cc-pVTZ and may′-cc-pVTZ [i.e., may-cc-pVTZ (Papajak et al., 2009) with *d* functions removed from hydrogen atoms] basis sets, obtained in a previous work (Boussessi et al., 2020a), are also reported.

3.1. Equilibrium Geometry

Theoretical equilibrium geometries have been benchmarked against semi-experimental equilibrium structures that, in view of their high accuracy, represent optimal reference values to test the predictive power of new computational approaches. The MDs and MADSs obtained for bond lengths and bond angles over the whole set of molecules are reported in **Figures 2A,B**, respectively, while the results obtained for each molecule can be found in **Supplementary Tables 1–20**. All the hybrid DFs well describe the equilibrium geometries, with MADs within 0.01 Å and 0.35° for bond lengths and bond angles, respectively. Interestingly, both the ωB97 family and the PW6B95 functional yield more accurate equilibrium geometries than the B3LYP/SNSD model, that during the last years, has been proposed as a good tradeoff between accuracy and computational cost to predict the structure and spectroscopic properties of medium- to large-size molecules. However, it should be noted that for the functionals belonging to the ωB97 family this improvement can be mainly attributed to the use of additional polarization functions on second-row atoms. In fact, for difluoromethane, oxirane, glycolaldehyde and ethanimine bond lengths obtained from ωB97, ωB97X and ωB97X-D functionals are in line with B3LYP/SNSD ones. Significantly more accurate bond lengths are obtained for molecules containing sulfur and chlorine atoms: in particular, it should be noted that the description of the C-S bond length of ethyl-mercaptan improves by one order of magnitude, and for SO_2 the deviation from the semi-experimental equilibrium value for the S=O bond length decreases from 0.05 Å with the SNSD basis set to about 0.01 Å employing the ωB97(X-D) functional in conjunction with the aug-jul-cc-pV(D+d)Z basis sets. Similar conclusions can be drawn for the S-H and C-Cl bond lengths thus highlighting the importance of additional *d* functions for second-row elements. The improvement brought by PW6B95 over B3LYP is more systematic, indeed, in conjunction with aug- and jul-cc-pVDZ it attains lower deviations also for molecules containing only first-row elements. Moving to double-hybrid functionals, **Figures 2A,B**, demonstrate the good performance of B2PLYP and rev-DSD-PBEP86 approximations that, when coupled to triple-zeta basis sets, deliver an equivalent description of bond lengths (with MADs around 0.0035 Å), whereas for bond angles the rev-DSD-PBEP86 functional appears more accurate than B2PLYP, with excellent MDs and MADs of 0.03 and 0.15°, respectively.

3.2. Rotational Spectroscopic Parameters

Rotational constants are the leading terms for the prediction of rotational spectra as they rule the frequencies of the rotational transitions. While the rotational constants of the equilibrium configuration are straightforwardly derived from the equilibrium structure, for applications in the field of rotational spectroscopy, vibrational effects must be included in order to obtain the rotational constants of the molecule in a given (usually the ground) vibrational state. Even though vibrational corrections usually account only for ∼1–3% of the total rotational constant value, their inclusion is mandatory for quantitative predictions of rotational spectra and, furthermore, it is necessary for the comparison with experimentally determined rotational constants. As well-known, ground state rotational constants, (B_0), are obtained by adding vibrational corrections, ΔB_{vib} to equilibrium rotational constants (B_e):

$$B_0^i = B_e^i + \Delta B_{vib}^i = B_e^i - \frac{1}{2}\sum_r \alpha_r^i \quad (1)$$

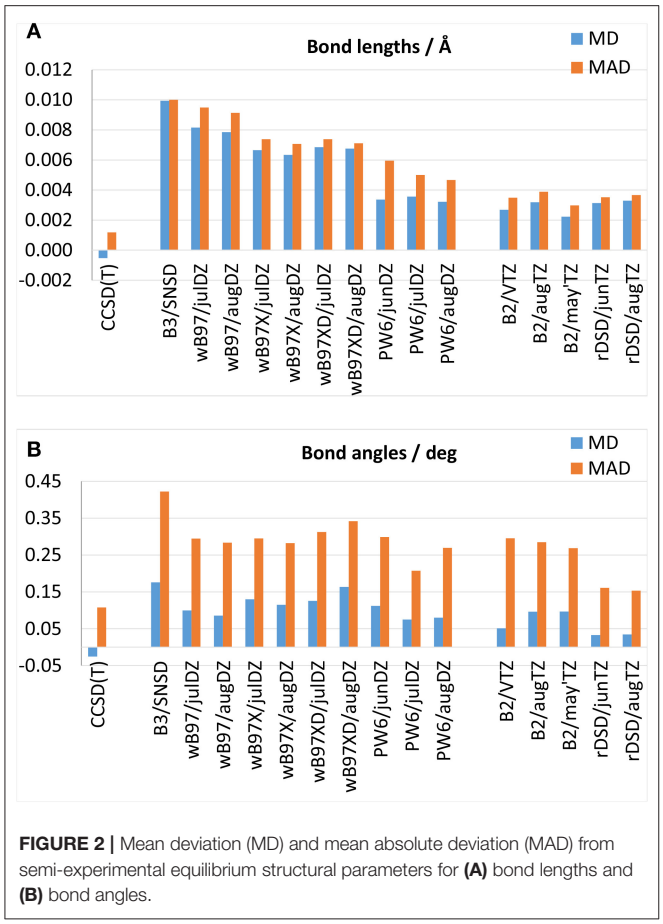

FIGURE 2 | Mean deviation (MD) and mean absolute deviation (MAD) from semi-experimental equilibrium structural parameters for **(A)** bond lengths and **(B)** bond angles.

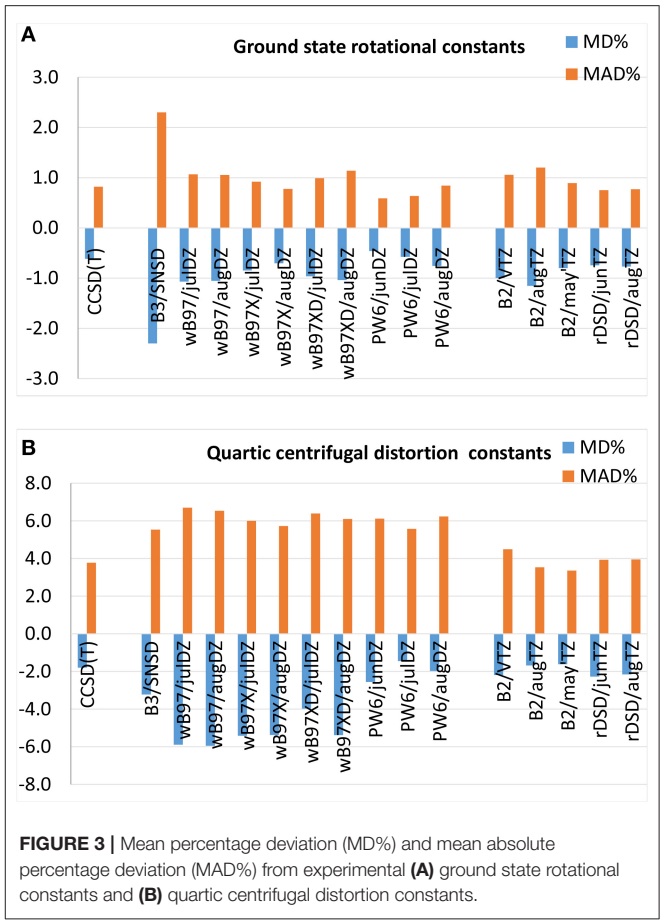

FIGURE 3 | Mean percentage deviation (MD%) and mean absolute percentage deviation (MAD%) from experimental **(A)** ground state rotational constants and **(B)** quartic centrifugal distortion constants.

where the summation runs over all the normal modes of vibration and the vibration-rotation interaction constants, α_r^i, are evaluated in the framework of VPT2 (Mills, 1972). Their computation requires to go beyond the RR/HO approximation as they depend on the semi-diagonal cubic force field, i.e., the third derivatives of the potential energy. Besides vibrational effects, also centrifugal distortions need to be accounted for, especially for predicting high-energy rotational transitions which are of interest in e.g., astrophysical and atmospheric applications of rotational and high-resolution IR spectroscopy. Given the different orders of magnitude that rotational constants and centrifugal distortion parameters can have, the performance of the different levels of theory considered in this work can be more conveniently evaluated by referring to percentage deviations.

3.2.1. Ground State Rotational Constants

Mean percentage deviations and mean absolute percentage deviations from experimental ground state rotational constants are reported in **Figure 3A** and the full list of results can be found in **Supplementary Tables 21–30**. At first, it should be noted that CCSD(T)-based computations, B3LYP/SNSD and B2PLYP/triple-ζ levels of theory reproduce experimental values with MAD%s around 0.8, 2.3, and 1%, respectively. However, the average errors for CCSD(T) computations may be slightly overestimated due to the inaccuracy of the rotational constants of cis-ClHC=CHF computed by a scaled-CCSD(T) force field, in particular A_0 which displays a deviation of −12% with respect to the experimental value. Indeed, by discarding this molecule, the global MD% and MAD% for CCSD(T)-based methods drop to −0.1 and 0.3%, respectively. The last-generation DFs examined in the present work provide excellent results in the prediction of ground state rotational constants: indeed, the ωB97 family shows MD%s and MAD%s around 1% and the PW6B95 functional performs even better, the MAD% being around 0.6% when coupled to the jun-cc-pVDZ and jul-cc-pVDZ basis sets and 0.8% in conjunction with the aug-cc-pVTZ basis set. A similar accuracy is delivered by the rev-DSD-PBEP86 double-hybrid functional, whose MAD% of 0.7% slightly improves the performance of the B2PLYP functional.

3.2.2. Centrifugal Distortion Constants

As shown in **Figure 3B**, quartic centrifugal distortion constants are reproduced by hybrid DFs with MAD%s around 6–7% independently of the functional or basis set employed (the full list of results is reported in **Supplementary Tables 31–40**). Despite the similar performance, the last generation DFs considered in the present work perform slightly worse than the B3LYP/SNSD model, with the exception of the PW6B95/jul-cc-VDZ level of

theory, which yields almost the same MAD% as B3LYP (5.5%) and a lower MD% (−2.0 vs. 3.2%), accompanied however by larger fluctuations (from −46 to 38% for PW6B95/jul-cc-pVDZ, in the −29 −19% range for B3LYP/SNSD). The rev-DSD-PBEP86 functional reproduces the quartic centrifugal distortion constants determined experimentally with a MD% and a MAD% around −2.2 and 3.9%, respectively, which are very similar to the scores of the B2PLYP functional in conjunction with augmented triple-ζ basis sets (MD% = −1.6% and MAD% = 3.6%). In passing, it should be noted that removal of diffuse functions worsens the accuracy of the results by about 1%.

At this point a few remarks concerning the comparison between experimental and theoretical centrifugal distortion constants are deserved. First, it has to be noted that computed constants refer to the equilibrium configuration of the molecule whereas rotational spectroscopy measurements provide those of the ground vibrational state. Even if the vibrational dependence of centrifugal distortion parameters is expected to amount to a few percent, it is an experimentally measurable quantity. However, only a few studies have been devoted to the theoretical treatment of vibrational effects on the centrifugal distortion (Watson, 2005). Second, attention has to be paid in comparing theory and experiment because measured centrifugal distortion constants can be affected by both limited accuracy and shortcomings in the fitting procedure used for their determination, as already pointed out by Boussessi et al. (2020a,b) Indeed, during the fitting procedure, centrifugal distortion constants may absorb the effects of resonances not fully treated (in particular high-order Coriolis or anharmonic interactions) that can be difficult to describe properly within the ro-vibrational Hamiltonian employed for inverting the measured transitions.

3.3. Harmonic Vibrational Properties

Since experimental harmonic vibrational frequencies are not available except for a few very simple molecules, DFT predictions are here benchmarked against CCSD(T) computations performed either in conjunction with large basis sets or within composite schemes. This comparison can be justified *a posteriori* on the basis of the good agreement between high-level CCSD(T) computations and experimental anharmonic wavenumbers and integrated absorption cross sections, which imply an accurate underlying harmonic force field. In the following discussion, the gauche conformer of ethyl mercaptan has been excluded from the data set in view of the huge discrepancies, up to 113 cm^{-1}, between CCSD(T)-F12 and experimental fundamental wavenumbers as discussed in (Boussessi et al., 2020a). MDs and MADs for harmonic vibrational wavenumbers and IR intensities are shown in **Figures 4A,B**, respectively, with the results for individual molecules being listed in the **Supplementary Tables 41–59**. As it can be seen from **Figure 4A**, concerning harmonic wavenumbers, among the functionals belonging to the ωB97 family there is a steady improvement of the performance on moving from ωB97, to ωB97X up to ωB97X-D which, with a MAD of around 14 cm^{-1} and a MD very close to zero, is the only one reaching an accuracy better than the B3 model (MAD = 16 cm^{-1}). The performance of the PW6B95 functional with different basis sets is similar to that of B3LYP in conjunction with the SNSD basis set, with MADs between 15 and 17 cm^{-1} and MDs around 0.7 cm^{-1}. On the other hand, the rev-DSD-PBEP86 functional represents a slight improvement over the already notable predictive power of the B2 model, being able to reproduce CCSD(T)-based reference data with a MAD of about 5 cm^{-1} to be compared with about 8 cm^{-1} at the B2PYLP level. A different picture is obtained for harmonic IR intensities: in fact, as shown in **Figure 4B**, all the hybrid functionals show comparable accuracy with MADs around 5 km mol^{-1}, the only exception being PW6B95 in conjunction with the jun-cc-pVDZ basis set, which provides poorer results. This is probably related to the lack of diffuse *d*-functions in the basis set, whose role in the computation of IR intensities is well-known: the MAD of 9 km mol^{-1} is indeed very similar to that obtained in Boussessi et al. (2020a) for the B3LYP functional in conjunction with the pcs-1 basis set, also lacking diffuse functions. Interestingly, also for intensities the rev-DSD-PBEP86 functional, with a MAD of 2.2 km mol^{-1} performs better than B2PLYP, whose MAD in conjunction with the aug-cc-pVTZ basis set is around 3 km mol^{-1} from reference CCSD(T)-based results. From the above discussion, it can be concluded that, concerning the calculation of vibrational frequencies and IR intensities within the double-harmonic approximation, the ωB97X-D and PW6B95 hybrid functionals can represent good alternatives to the B3LYP/SNSD level of theory, and rev-DSD-PBEP86 in conjunction with triple-ζ basis sets including diffuse functions appears even more accurate than the already well-performing B2 model.

3.4. Beyond the Double-Harmonic Approximation: Wave-Numbers and Absorption Cross Sections

The performance of the DFs of the last generation considered in this work for anharmonic fundamental frequencies are compared in **Figure 5** (**Supplementary Tables 60–69** report the results for each molecule). It is noted that, although CCSD(T)-based computations reach a MAD of 8 cm^{-1}, this result can be biased by the disagreement between experimental transition frequencies and CCSD(T)-F12/cc-pVTZ-F12 predictions previously reported for gauche-CH_3CH_2SH. As to the B3 and B2 models, they reproduce experimental observations with a MAD of 20 and 12 cm^{-1}, respectively. In comparison, all the model chemistries based on the ωB97 approximation, showing MADs between 26 and 36 cm^{-1} cannot compete with the B3LYP functional. Given the similar results obtained for harmonic properties, it is evident that the ωB97 DFs have problems with the computation of higher order derivatives of the potential energy (i.e., cubic and semi-diagonal quartic force constants). Some of the anharmonic contributions computed for glycolaldehyde at the ωB97X-D level are particularly disappointing: the harmonic frequencies of ν_6 and ν_8 normal modes are predicted at about 1,445 and 1,305 cm^{-1}, respectively, whereas at the ωB97X-D/jul-cc-pVDZ anharmonic level they are shifted at 904 and 719 cm^{-1} (1,012 and 832 cm^{-1} when using the aug-cc-pVTZ basis set). Even worse is the case of the ν_{12} vibration for which the anharmonic corrections evaluated by using jul-cc-pVDZ and aug-cc-pVDZ basis sets amount to −1,005 and

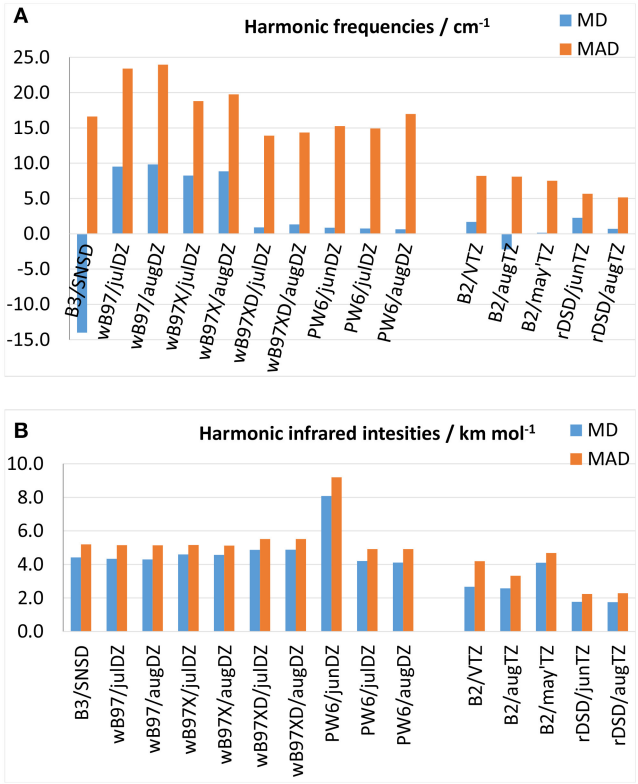

FIGURE 4 | Mean deviation (MD) and mean absolute deviation (MAD) from reference CCSD(T)-based theoretical values for **(A)** harmonic vibrational frequencies, **(B)** harmonic infrared intensities.

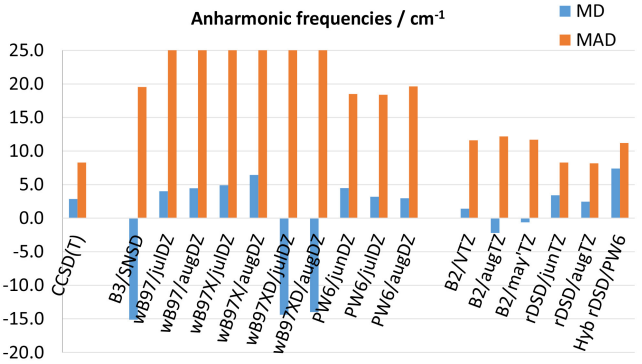

FIGURE 5 | Mean deviation (MD) and mean absolute deviation (MAD) from experimental fundamental frequencies. MADs for the ωB97 family are out of range.

-503 cm^{-1} thus resulting in a non-physical negative value of the ν_{12} fundamental transition frequency. Conversely, the PW6B95 DF turns out to be competitive with the B3 model, reaching a comparable MAD of 18 cm^{-1} when employed in conjunction with jul- and jun-cc-pVDZ basis sets and of 20 cm^{-1} together with the aug-cc-pVTZ basis set. Furthermore, recently Kreienborg and Merten (2019) pointed out that carbon-fluorine stretching vibrations are often strongly misplaced by common hybrid functionals with the possible exception of the M06-2X functional (restricting, however, the simulation to the double-harmonic approximation). Nevertheless, while this functional (and its predecessor M05-2X) predicts harmonic frequencies not far from experimental fundamentals, it becomes unreliable when anharmonic contributions are taken into the proper account (Puzzarini et al., 2010; Tasinato, 2014; Tasinato et al., 2018). Rather, it should be noted that the PW6B95 functional yields a consistent description of anharmonic C-F stretching frequencies, with deviations from experiment generally halved with respect to those of the B3 model. In fact, by focusing only on the C-F stretchings, the B3LYP/SNSD level of theory presents a MAD (computed over seven values) of 34 cm^{-1} and a maximum deviation of -50 cm^{-1}, while the PW6B95 DF (in conjunction with all the tested double-ζ basis sets) reproduces the experimental values with a MAD around 15 cm^{-1} and a maximum error of about -25 cm^{-1}. The rev-DSD-PBEP86 functional slightly improves over B2PLYP also for anharmonic fundamental wavenumbers: indeed it reproduces experimental data with a MD and a MAD around 2 and 8 cm^{-1}, respectively, in comparison to -1 and 12 cm^{-1} for the B2PLYP/aug-cc-pVTZ level of theory.

Some remarks are in order about the accuracy in the computation of IR integrated absorption cross sections (i.e., IR band intensities). It should be noted that their experimental determination is a daunting task prone to both systematic and random errors, hence a careful control of the experimental conditions and errors source needs to be performed during the measurements. Therefore, to assess the accuracy of DFT calculations, one must rely on precise experimental values, which are available only for a reduced number of molecules among those of the benchmark set, namely CH_2F_2, CH_2ClF, $ClFC{=}CH_2$, $ClFC{=}CF_2$, and SO_2. Moreover, in view of the poor results delivered by the ωB97-based functionals for anharmonic wavenumbers, the attention for IR anharmonic intensities is focused on PW6B95 and rev-DSD-PBEP86 DFs, whose performances are shown in **Figure 6** together with those of B3 and B2 models (the full list of results is given in **Supplementary Tables 70–74**). In passing it should be stressed that the accuracy reached for integrated absorption cross sections depends on the reliability of both the anharmonic potential energy and dipole moment surfaces. In fact, when overlapping IR bands cannot be resolved at the experimental level, the integration required for determining band intensities is performed by considering all the absorptions within a given spectral interval. Here, the same approach has been mimicked at the theoretical level, i.e., the intensities of all the bands predicted in a given (experimental) integration range have been summed. As it can be seen, as for anharmonic wavenumbers, PW6B95 in conjunction with the aug-cc-pVDZ or jul-cc-pVDZ basis sets performs on par with B3LYP/SNSD, the MDs and MADs being around 9 and 10 km mol^{-1}, respectively. Furthermore, as already pointed out for harmonic IR intensities, the lack of diffuse d functions in the jun-cc-pVDZ basis set results in a worsening of the predictions, thus highlighting

the necessity of diffuse polarization functions for a reliable description of the dipole moment surface. Moving to the double-hybrid DFs, **Figure 6** shows that, also for integrated absorption cross sections, the rev-DSD-PBEP86 functional, with both MD and MAD around 5 km mol^{-1}, improves over B2PLYP that, in turn, reproduces experimental measurements with a mean absolute deviation ranging between 8 and 10 km mol^{-1} for the different triple-ζ basis sets.

Finally, anharmonic vibrational properties have been computed by a hybrid QM/QM′ approach featuring harmonic frequencies and intensities calculated at the rev-DSD-PBEP86/jul-cc-pVTZ level and anharmonic effects obtained from PW6B95/jun-cc-pVDZ computations. The hybrid force field obtained in this way simulates fundamental wavenumbers with a MAD of 11 cm^{-1}, thus showing a worsening of only 3 cm^{-1} in comparison to full rev-DSD-PBEP86 computations, but still performing similarly to full B2 anharmonic predictions. This result shows that the PW6B95/jun-cc-pVDZ level of theory represents a reliable and cost-effective approach to compute anharmonic corrections when employed in conjunction with harmonic force fields obtained from the rev-DSD-PBEP86 functional. Conversely, PW6B95 should be used together with the larger jul-cc-pVDZ basis set when performing full anharmonic computations in order to obtain reliable predictions of band intensities.

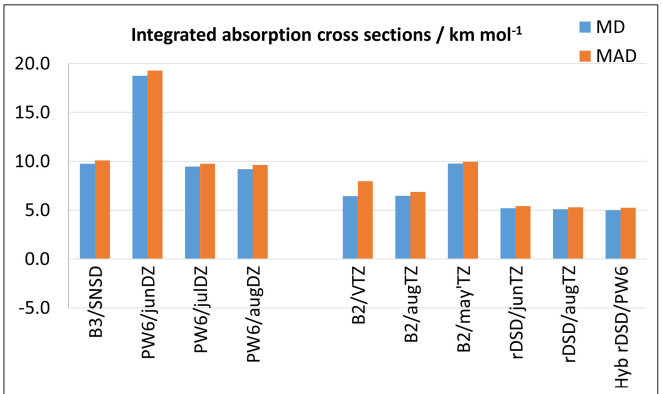

FIGURE 6 | Mean deviation (MD) and mean absolute deviation (MAD) from experimental integrated absorption cross sections for selected model chemistries.

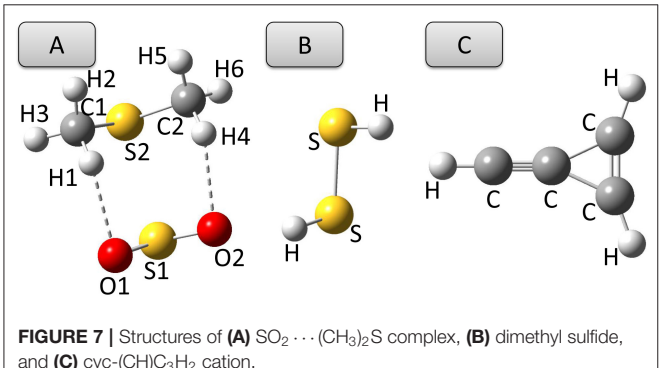

FIGURE 7 | Structures of **(A)** $SO_2 \cdots (CH_3)_2S$ complex, **(B)** dimethyl sulfide, and **(C)** cyc-$(CH)C_3H_2$ cation.

3.5. New Model Chemistries at Work

The previous sections have shown that, among the hybrid DFs of the last generation considered in this work, ωB97X-D and PW6B95 in conjunction with either aug- or jul-cc-pVDZ basis sets provide reliable predictions of equilibrium structures, rotational parameters and harmonic vibrational properties, sometimes even better than the well-tested B3 model. However, when anharmonic effects come into play, all the functionals of the ωB97 family yield unstable, sometimes disappointing, results. Conversely, the PW6B95 functional performs similarly to the B3 model for both fundamental transition frequencie and integrated absorption cross sections, provided that an additional set of *d*

TABLE 1 | Equilibrium structure of the $SO_2 \cdots S(CH_3)$ 1:1 complex[a].

	r_e^{SE} [b]	B3LYP-D3[c]	PW6B95-D3[d]	B2PLYP-D3[e]	rev-DSD-PBEP86-D3[f]
r(S1-S2)	2.947	2.8672	2.8676	2.9257	2.9288
r(O1-S1)	1.446	1.4966	1.4579	1.4532	1.4499
r(C1-S2)	1.790	1.8244	1.7954	1.8040	1.8029
r(H1-C1)	1.089	1.0941	1.0949	1.0883	1.0911
r(H2-C1)	1.086	1.0915	1.0936	1.0864	1.0895
r(H3-C1)	1.087	1.0922	1.0926	1.0870	1.0901
∠(O1S1S2)	95.0	96.10	95.62	95.20	94.92
∠(C1S2S1)	91.7	91.20	91.88	91.27	91.52
∠(H1C1S2)	110.6	110.08	110.72	110.67	110.60
∠(H2C1S2)	107.3	106.91	107.01	107.14	107.25
∠(H3C1S2)	109.8	109.48	109.98	109.97	109.93
δ(O2S1S2O1)	−118.2	−115.92	−117.51	−117.84	−117.86
δ(C1S2S1O1)	9.8	7.70	8.46	8.84	9.10
δ(H1C1C2S1)	27.2	21.71	26.60	26.06	27.91
δ(H2C1S2S1)	−91.6	−96.87	−91.94	−92.56	−90.77
δ(H3C1S2S1)	149.5	144.27	149.41	148.63	150.37

[a] *Bond lengths in Å, angles in deg. For atom labeling see* ***Figure 7****.*
[b] *Semi-experimental equilibrium structure from Obenchain et al. (2018).*
[c] *Employed in conjunction with the SNSD basis set.*
[d] *Employed in conjunction with the jul-cc-pV(D+d)Z basis set.*
[e] *Employed in conjunction with the may′-cc-pVTZ basis set. From Obenchain et al. (2018).*
[f] *Employed in conjunction with the jun-cc-pV(T+d)Z basis set.*

TABLE 2 | Equilibrium structure of hydrogen disulfide[a].

	PW6B95-D3/ jul-cc-pV(D+d)Z	rev-DSDPBEP86-D3/ jun-cc-pV(T+d)Z	SE[b]
r(S-S)	2.0596	2.0609	2.0513 (3;7)
r(S-H)	1.3493	1.3422	1.3401 (14;32)
∠(HSS)	98.35	98.24	98.07 (2;3)
δ(HSSH)	90.73	90.63	90.72 (2;5)

[a] *Bond lengths in Å, angles in deg.*
[b] *Semi-experimental structure from Ye et al. (2020). Values in parentheses are the standard deviation (first value) and the confidence interval at a 95% confidence level (second value).*

functions is employed on heavy elements, and in the case of intensities, at least a basis set of the jul-cc-pVDZ quality is used. Coming to double-hybrid functionals, the predictions of the rev-DSD-PBEP86 model are better than, or at least similar to, their B2 counterparts for both structural and rotational-vibrational spectroscopic properties.

Given these premises, in this section the model chemistries able of rivaling with the B2 and B3 paradigms are applied to selected case studies in the field of structural determination and IR spectroscopy. A convenient case-study for the first subject is the complex formed by SO_2 and dimethyl sulfide, $(CH_3)_2S$ (see **Figure 7A**), whose semi-experimental equilibrium structure has been recently determined (Obenchain et al., 2018). Then, the anharmonic IR spectrum of the cyclic-$(CH)C_3H_2^+$ cation (**Figure 7C**) has been simulated because of the astrophysical interest of this molecule pointed out very recently by Westbook et al. (2020).

The structural parameters of the equilibrium configuration of the $SO_2 \cdots (CH_3)_2S$ 1:1 complex evaluated by different DFs and basis sets are compared in **Table 1** with the semi-experimental structure obtained by Obenchain et al. (2018). Inspection of this table reveals that the B3 model reproduces bond lengths, valence and dihedral angles with average absolute errors of 0.03 Å, 0.6 and 4°, respectively. The PW6B95-D3/jul-cc-pV(D+*d*)Z model chemistry delivers significantly improved results, almost halving the deviations for bond lengths and valence angles and showing an absolute average error for dihedral angles around 0.6°. Concerning the B2PLYP-D3 (employed in conjunction with the may′-cc-pVTZ basis set) and rev-DSD-PBEP86-D3 (in conjunction with the jun-cc-pV(T+*d*) basis set) functionals, it can be seen that, on average, they perform equally well for both bond lengths (the mean absolute deviation over all the complex bond lengths are 0.014 and 0.013 Å, respectively) valence (0.2 vs. 0.1°) and dihedral (0.8 and 0.7°) angles. It is also noteworthy that the hybrid PW6B95-D3 functional reproduces the semi-experimental structure of the complex with an accuracy that, in spite of the considerably reduced computational cost, rivals that of the B2 and rev-DSD-PBEP86 double-hybrids. In order to understand whether the quite large deviation observed for the inter-molecular S-S distance is due to an intrinsic inaccuracy for S-S bonds or to an unbalanced treatment of inter-molecular interactions, the equilibrium geometry of hydrogen disulfide, HSSH (**Figure 7B**), has been computed at the PW6B95-D3/jul-cc-pV(D+*d*)Z level and compared with the semi-experimental structure recently reported by Ye et al. (2020). The obtained structural parameters, detailed in **Table 2**, show deviations of 8 and 9 mÅ for the S-S and S-H bond lengths, respectively, and of 0.3° for the HSS angle, while the HSSH dihedral is within the uncertainty of the semi-experimental value. Since the SS bond length of HSSH computed at PW6B95-D3 and rev-DSD-PBEP86-D3 levels is very close, the worse performance of the former functional for the SS distance in the complex is probably related to the description of non-covalent interactions.

Moving to the vibrational properties of the cyc-$(CH)C_3H_2^+$ cation, harmonic frequencies are listed in **Table 3** while anharmonic fundamental wavenumbers and IR intensities are reported in **Table 4**. In both Tables the results obtained by Westbrook et al. at the CCSD(T)-F12/aug-cc-pVTZ level (Westbook et al., 2020) are also reported for comparison purposes. In that work, the authors focused on the difficulties of post-Hartree-Fock methods in describing out-of-plane bending vibrations of molecules with C=C multiple bonds (especially, but not only, aromatic systems), an issue that should be of minor concern for DFT. For this reason, it is of some interest to compare their CCSD(T)-F12 predictions with the present DFT simulations. For hybrid DFs, the major differences with

TABLE 3 | Harmonic frequencies (in cm^{-1}) and intensities (in km mol^{-1}) for *cyc*-$CHC_3H_2^+$ computed at different levels of theory.

ω	*Sym.*	B3/SNSD		PW6/julDZ		B2/junTZ		rDSD/junTZ		TZ/augVTZ[a]	MP2[a]
		ω	*I*	*ω*	*I*	*ω*	*I*	*ω*	*I*	*ω*	*I*
ω_1	a_1	3,366	175.49	3,396	182.12	3,387	159.25	3,379	148.18	3370.2	121
ω_2	a_1	3,310	68.88	3,350	67.03	3,335	70.79	3,333	66.88	3325.9	76
ω_3	a_1	1,973	52.29	2,000	56.93	1,993	75.71	2,004	98.48	1977.3	186
ω_4	a_1	1,673	51.24	1,692	48.78	1,695	54.20	1,697	65.58	1681.3	67
ω_5	a_1	917	2.85	913	3.63	920	4.40	912	6.20	908.2	7
ω_6	a_1	798	22.40	825	19.42	772	33.55	766	42.04	768.6	63
ω_7	a_2	758	0.00	747	0.00	749	0.00	735	0.00	713.1	0
ω_8	b_1	731	65.29	723	63.95	735	65.00	731	64.92	711.5	77
ω_9	b_1	450	103.33	436	94.32	503	92.13	517	89.33	509.3	12
ω_{10}	b_1	361	3.08	323	13.42	382	3.55	376	4.96	355.9	0
ω_{11}	b_2	3,245	196.22	3,281	202.93	3,268	210.18	3,265	208.15	3260.5	234
ω_{12}	b_2	995	6.38	989	5.47	998	7.00	994	6.38	985.5	8
ω_{13}	b_2	944	0.69	933	0.45	937	0.04	935	0.11	920.6	3
ω_{14}	b_2	627	28.13	680	32.53	571	25.14	578	26.50	580.4	22
ω_{15}	b_2	200	24.47	190	23.01	170	26.96	160	27.56	155.7	46

[a] *Taken from Westbook et al. (2020).*

TABLE 4 | Anharmonic frequencies (in cm^{-1}) and intensities (in km mol^{-1}) for *cyc*-$CHC_3H_2^+$ obtained at different levels of theory.

ν	*Sym.*	B3/SNSD		PW6/julDZ		B2/junTZ		rDSD/junTZ		B2/B3		rDSD/PW6		rDSD/B3		TZ/aVTZ[a]
		ν	*I*	*ν*	*I*	*ν*	*I*	*ν*	*I*	*ν*	*I*	*ν*	*I*	*ν*	*I*	*ν*
ν_1	a_1	3,242	89.15	3,256	213.92	3,261	188.89	3,252	159.06	3,262	188.96	3,243	217.46	3,254	202.75	3237.1
ν_2	a_1	3,173	47.80	3,204	25.91	3,195	32.81	3,195	43.11	3,194	27.39	3,181	17.69	3,193	27.27	3182.2
ν_3	a_1	1,948	32.75	1,977	43.62	1,975	58.99	1,984	77.84	1,973	66.32	1,984	43.62	1,987	33.98	1947.8
ν_4	a_1	1,647	45.54	1,664	41.55	1,666	33.30	1,665	57.64	1,672	41.27	1,671	45.65	1,673	47.57	1641.9
ν_5	$_a1$	903	3.50	897	5.08	902	0.81	902	3.55	901	1.92	873	0.31	895	1.50	884.4
ν_6	a_1	761	22.08	804	25.51	744	6.46	766	14.54	751	22.70	719	5.33	746	14.77	760.2
ν_7	a_2	743	0.00	738	0.00	736	0.00	723	0.00	735	0.00	723	0.00	720	0	713
ν_8	b_1	717	67.04	713	66.81	722	68.65	719	68.28	723	67.48	720	67.62	719	67.62	729
ν_9	b_1	410	65.38	481	96.72	496	90.31	511	87.86	473	100.36	515	92.55	487	95.23	480.8
ν_{10}	b_1	315	35.31	303	8.67	380	4.26	372	5.64	343	0.08	388	12.01	336	13.79	154.3
ν_{11}	b_2	3,117	192.57	3,146	199.69	3,141	207.44	3,140	205.27	3,141	206.76	3,129	200.04	3,138	193.33	3130.1
ν_{12}	b_2	964	5.79	962	7.78	974	8.12	968	7.34	970	8.83	966	8.11	965	9.21	970.1
ν_{13}	b_2	909	0.02	903	0.03	901	0.91	921	0.48	896	1.23	907	0.03	908	0.02	894
ν_{14}	b_2	594	25.65	653	30.44	531	21.92	539	22.88	525	22.11	530	29.72	532	25.18	547.4
ν_{15}	b_2	199	23.79	189	22.13	172	26.35	163	26.95	172	26.33	160	22.46	162	23.91	175.9

[a] *Taken from Westbook et al. (2020).*

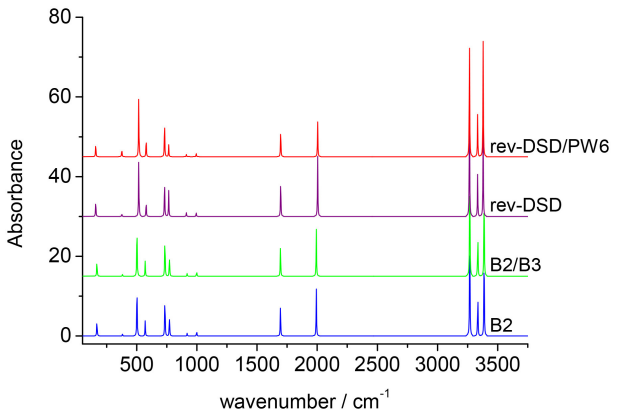

FIGURE 8 | Anharmonic infrared spectrum of cyc-$CHC_3H_2^+$ simulated by using different levels of theory (computed spectral transitions have been convoluted with a Lorentzian function with an half-width at half-maximum of 2 cm^{-1}). B2: full B2PLYP/jun-cc-pVTZ anharmonic force field; B2/B3: B2PLYP/jun-cc-pVTZ harmonic force field and B3LYP/SNSD anharmonic effects; rev-DSD: full rev-DSDPBEP86/jun-cc-pVTZ anharmonic force field; rev-DSD/PW6: rev-DSDPBEP86/jun-cc-pVTZ harmonic force field and PW6B95/jul-cc-pVDZ anharmonic effects. Some traces have been displaced fpr clarity.

respect to CCSD(T)-F12/aug-cc-pVTZ harmonic vibrational frequencies, can be observed, for both B3LYP and PW6B95 functionals, for the ω_7, ω_9, ω_{14} and ω_{15} normal modes which correspond to the H-C=C-H out-of-plane bending, the ≡C-H out-of-plane vibration, the in-plane ring deformation and the in-plane bending vibration of the C-H groups, respectively. Concerning double-hybrids, the largest difference (+36 cm^{-1}) is observed for ω_7 at B2PLYP/jun-cc-pVTZ level, while the worst rev-DSDPBEP86/jun-cc-pVTZ result concerns the ω_3 C≡C stretching vibration (+27 cm^{-1}). Moving to the anharmonic vibrational wavenumbers, a huge anharmonic correction (−202 cm^{-1}) for the ν_{10} fundamental and an unusual positive contribution (20 cm^{-1}) for the ν_{15} vibration were obtained by Westbook et al. (2020) Conversely, according to the present calculations, the anharmonic correction for the ν_{15} vibration amounts to a few wave-numbers and that of ν_{10} ranges from about −45 cm^{-1} at the B3LYP level to −2 cm^{-1} at the B2PLYP/jun-cc-pVTZ level. While a number of Fermi resonances both of type 1 and 2 have been found, none of them strongly alters the spectral structure, just shifting the transitions by about 5 cm^{-1} from their unperturbed values. The only exception is the ν_6 normal mode, whose excited $\nu_6 = 1$ level is involved in a resonant triad together with $\nu_{10} = 2$ and $\nu_{14} = \nu_{15} = 1$.

Three different hybrid force fields have been also computed: in the B2/B3 hybrid force field, B2PLYP/jun-cc-pVTZ harmonic frequencies have been corrected by B3LYP/SNSD anharmonic contributions, while in the revDSD/B3 and revDSD/PW6 hybrid QM/QM′ approaches, the harmonic force field at the rev-DSD-PBEP86/jun-cc-pVTZ level has been coupled to cubic- and semi-diagonal quartic force constants evaluated at B3LYP/SNSD and PW6B95/jul-cc-pVDZ levels, respectively. The results collected in **Table 4** show that the hybrid force fields deliver, with a few exceptions, the same results as the corresponding force field obtained by full anharmonic calculations at B2PLYP or rev-DSD-PBEP86 levels. The spectra of cyc-$(CH)C_3H_2^+$ cation, simulated beyond the double-harmonic approximation at different levels of theory, which can be useful to guide experimental measurements on this species or even astronomical campaigns, are reported in **Figure 8**. It is quite apparent that the only major differences among the different theoretical models are found for the integrated absorption cross sections, whereas transition frequencies are only marginally affected.

4. CONCLUSIONS

In the last decade we have witnessed the increasing accuracy of methods rooted in the density functional theory due to ongoing improvements in both methodological and numerical aspects. However, attention has mainly been focused on thermochemistry and kinetics, whereas theoretical support to rotational and vibrational spectroscopy requires accurate predictions of molecular geometries, harmonic force fields and leading anharmonic contributions, not to mention dipole moments and their derivatives. On these grounds, the present work has analyzed the performance of last-generation hybrid and double-hybrid functionals in conjunction with partially augmented correlation consistent basis sets for a benchmark set of 10 molecules of both atmospheric and astrochemical relevance. Equilibrium molecular geometries issued from DFT computations have been benchmarked against accurate semi-experimental equilibrium structures, while rotational constants, centrifugal distortion parameters, and vibrational frequencies have been compared to the experimental data available in the literature and with high level CCSD(T)-based *ab initio* calculations. The following conclusions can be drawn:

(1) The jun-cc-pVDZ basis set performs remarkably well for hybrid functionals with the possible exception of IR intensities, which require diffuse *d*-functions, namely the jul-cc-pVDZ (or SNSD) basis set. In the case of double-hybrid functionals, the jun-cc-pVTZ basis set represents a nearly optimum cost/performance balance, but also the may—-cc-pVTZ basis set can be safely employed for larger systems. An additional set of *d*-functions is always mandatory for second-row atoms.

(2) Among hybrid functionals, B3LYP-D3 (B3) is still very competitive, although the PW6B95 (PW6) model significantly improves equilibrium geometries.

(3) Concerning double-hybrid functionals, the rev-DSD-PBEP86-D3 functional (rDSD) systematically improves the already reliable results delivered by the B2PLYP (B2) model, the enhancement being especially significant for non-covalent complexes.

(4) Composite methods employing geometries and harmonic contributions evaluated by double-hybrid functionals coupled to anharmonic corrections, obtained with hybrid functionals, always lead to accurate results. In this connection the previously employed B2/B3 model remains very useful, but the new rDSD/PW6 variant seems capable of delivering even better results with the same cost.

In summary, even with further pending developments and validation, thanks to effective implementations in general electronic structure codes, last-generation hybrid and double-hybrid functionals provide unprecedented accuracy for all the parameters ruling rotational and vibrational spectroscopy with computer requirements well within current standards and, coupled to generalized second-order vibrational perturbation theory (GVPT2), can also be used by non-specialists to complement experimental studies of medium- and, even, large-size molecules of current fundamental and technological interest.

DATA AVAILABILITY STATEMENT

The original contributions generated for the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

GC and MF performed the computations and analyzed the results. NT supervised the work, analyzed the results, and wrote part of the paper. VB defined the general strategy and wrote part of the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been supported by MIUR PRIN 2015 (Grant Number 2015F59J3R), PRIN 2017 (Grant Number 2017A4XRCA), and by Scuola Nornale Superiore (Grant Number SNS18B-TASINATO).

ACKNOWLEDGMENTS

The technical staff of the SMART laboratory at Scuola Normale Superiore di Pisa is thanked for help with use of HPC resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.584203/full#supplementary-material

REFERENCES

Aliev, M. R., and Watson, J. K. G. (1985). "Higher-order effects in the vibration-rotation spectra of semirigid molecules," in *Molecular Spectroscopy: Modern Research*, Vol. 3, ed K. N. Rao (New York, NY: Academic Press), 2–67. doi: 10.1016/B978-0-12-580643-5.50006-3

Alonso, J. L., Lesarri, A., Leal, L. A., and López, J. C. (1993). The millimeter-wave spectra of 1-chloro-1-fluoroethylene and cis-1-chloro-2-fluoroethylene. *J. Mol. Spectrosc.* 162, 4–9. doi: 10.1006/jmsp.1993.1265

Baiano, C., Lupi, J., Tasinto, N., Puzzarini, C., and Barone, V. (2020). The role of state-of-the-art quantum-chemical calculations in astrochemistry: formation route and spectroscopy of ethanimine as a paradigmatic case. *Molecules* 25:2873. doi: 10.3390/molecules25122873

Baiardi, A., Bloino, J., and Barone, V. (2017). Simulation of vibronic spectra of flexible systems: hybrid DVR-harmonic approaches. *J. Chem. Theory Comput.* 13, 2804–2822. doi: 10.1021/acs.jctc.7b00236

Barone, V. (2005). Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.* 122:014108. doi: 10.1063/1.1824881

Barone, V. (2012). *Computational Strategies for Spectroscopy: From Small Molecules to Nano Systems*. Hoboken, NJ: John Wiley & Sons.

Barone, V. (2016). The virtual multifrequency spectrometer: a new paradigm for spectroscopy. *WIREs Comput. Mol. Sci.* 6, 86–110. doi: 10.1002/wcms.1238

Barone, V., Baiardi, A., Biczysko, M., Bloino, J., Cappelli, C., and Lipparini, F. (2012). Implementation and validation of a multi-purpose virtual spectrometer for large systems in complex environments. *Phys. Chem. Chem. Phys.* 14, 12404–12422. doi: 10.1039/c2cp41006k

Becke, A. D. (1993). Density functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 98, 5648–5652. doi: 10.1063/1.464913

Berova, N., Polavarapu, P. L., Nakanishi, K., and Woody, R. W. (2012). *Comprehensive Chiroptical Spectroscopy: Instrumentation, Methodologies and Theoretical Simulations*. Hoboken, NJ: John Wiley & Sons.

Biczysko, M., Scalmani, G., Bloino, J., and Barone, V. (2010). Harmonic and anharmonic vibrational frequency calculations with the double-hybrid B2PLYP method. Analytic second derivatives and benchmark studies. *J. Chem. Theory Comput.* 6, 2115–2125. doi: 10.1021/ct100212p

Blanco, S., Lesarri, A., Alonso, J. L., and Guarnieri, A. (1995). The rotational spectrum of chlorofluoromethane. *J. Mol. Spectrosc.* 174, 397–416. doi: 10.1006/jmsp.1995.0011

Bloino, J., Biczysko, M., and Barone, V. (2012). General perturbative approach for spectroscopy, thermodynamics, and kinetics: methodological background and benchmark studies. *J. Chem. Theory Comput.* 8, 1015–1036. doi: 10.1021/ct200814m

Bloino, J., Biczysko, M., and Barone, V. (2015). Anharmonic effects on vibrational spectra intensities: infrared, Raman, vibrational circular dichroism and raman optical activity. *J. Phys. Chem. A* 119, 11862–11874. doi: 10.1021/acs.jpca.5b10067

Boussessi, R., Ceselin, G., Tasinato, N., and Barone, V. (2020a). DFT meets the segmented polarization consistent basis sets: Performances in the computation of molecular structures, rotational and vibrational spectroscopic properties. *J. Mol. Struct.* 1208:127886. doi: 10.1016/j.molstruc.2020.127886

Boussessi, R., Tasinato, N., Pietropolli Charmet, A., Stoppa, P., and Barone, V. (2020b). Sextic centrifugal distortion constants: interplay of density functional and basis set for accurate yet feasible computations. *Mol. Phys.* 118:e1734678. doi: 10.1080/00268976.2020.1734678

Carlotti, M., Nivellini, G., Tullini, F., and Carli, B. (1988). The far-infrared spectrum of methylene fluoride. *J. Mol. Spectrosc.* 132, 158–165. doi: 10.1016/0022-2852(88)90065-3

Carroll, P. B., Drouin, B. J., and Widicus-Weaver, S. L. (2010). The submillimeter spectrum of glycolaldehyde. *Astrophys. J.* 723, 845–849. doi: 10.1088/0004-637X/723/1/845

Chai, J.-D., and Head-Gordon, M. (2008a). Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* 10, 6615–6620. doi: 10.1039/b810189b

Chai, J.-D., and Head-Gordon, M. (2008b). Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* 128:084106. doi: 10.1063/1.2834918

Craig, N. C., Lo, Y. S., Piper, L. G., and Wheeler, J. C. (1970). Vibrational assignments and potential constants for cis- and trans-1-chloro-2-fluoroethylenes and their deuterated modifications. *J. Phys. Chem.* 74, 1712–1727. doi: 10.1021/j100720a010

Drummen, G. P. C. (2012). Fluorescent probes and fluorescence (microscopy) techniques-illuminating biological and biomedical research. *Molecules* 17, 14067–14090. doi: 10.3390/molecules171214067

Dunning, T. H. J. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* 90, 1007–1023. doi: 10.1063/1.456153

Flaud, J.-M., Lafferty, W. J., Kwabia-Tchana, F., Perrin, A. M., and Landsheere, X. (2012). First high-resolution analysis of the $\nu_1 5$, $\nu_1 2$, ν_5, $\nu_1 0$ and ν_2 bands of oxirane. *J. Mol. Spectrosc.* 271, 38–43. doi: 10.1016/j.jms.2011.11.005

Flaud, J.-M., Perrin, A. M., Salah, L. M., Lafferty, W. J., and Guelachvili, G. (1993). A reanalysis of the (010), (020), (100), and (001) rotational levels of $^3 2 S^1 6 O_2$. *J. Mol. Spectrosc.* 160, 272–278. doi: 10.1006/jmsp.1993.1174

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian16 Revision C.01*. Wallingford, CT: Gaussian Inc.

Gambi, A., Pietropolli Charmet, A., Stoppa, P., Tasinato, N., Ceselin, G., and Barone, V. (2019). Molecular synthons for accurate structural determinations: the equilibrium geometry of 1-chloro-1-fluoroethene. *Phys. Chem. Chem. Phys.* 21, 3615–3625. doi: 10.1039/C8CP04888F

Gambi, A., Puzzarini, C., Cazzoli, G., Dore, L., and Palmieri, P. (2002). The anharmonic force field of *cis*-1-chloro-2-fluoroethylene. *Mol. Phys.* 100, 3535–3543. doi: 10.1080/00268970210130155

Grimme, S. (2006). Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* 124:034108. doi: 10.1063/1.2148954

Grimme, S., Anthony, J., Ehrlich, S., and Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 132:154104. doi: 10.1063/1.3382344

Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32, 1456–1465. doi: 10.1002/jcc.21759

He, Y., Tang, L., Wu, X., Hou, X., and Lee, Y. (2007). Spectroscopy: the best way toward green analytical chemistry? *Appl. Spectrosc. Rev.* 42, 119–138. doi: 10.1080/05704920601184259

Hillig, K. W., Bittner, E. R., Kuczkowski, R. L., Lewis-Bevan, W., and Gerry, M. C. L. (1988). The chlorine nuclear quadrupole coupling tensor in chlorotrifluoroethylene. *J. Mol. Spectrosc.* 132, 369–379. doi: 10.1016/0022-2852(88)90332-3

Hochlaf, M., Puzzarini, C., and Senent, M. L. (2015). Towards the computations of accurate spectroscopic parameters and vibrational spectra for organic compounds. *Mol. Phys.* 113, 1661–1673. doi: 10.1080/00268976.2014.1003986

Johnson, T. J., Sams, R. L., Profeta, L. T. M., Akagi, S. K., Burling, I. R., Yokelson, R. J., et al. (2013). Quantitative IR spectrum and vibrational assignments for glycolaldehyde vapor: glycolaldehyde measurements in biomass burning plumes. *J. Phys. Chem. A* 117, 4096–4107. doi: 10.1021/jp311945p

Kendall, R. A., Dunning, T. H. J., and Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* 96, 6796–6806. doi: 10.1063/1.462569

Kolesnikov, L., Tercero, B., Cernicharo, J., Alonso, J. L., Daly, A. M., Gordon, B. P., et al. (2014). Spectroscopic characterization and detection of ethyl mercaptan in orion. *Astrophys. J. Lett.* 784:L7. doi: 10.1088/2041-8205/784/1/L7

Kreienborg, N. M., and Merten, C. (2019). How to treat C-F stretching vibrations? A vibrational CD study on chiral fluorinated molecules. *Phys. Chem. Chem. Phys.* 21, 3506–3511. doi: 10.1039/C8CP02395F

Lafferty, W. J., Flaud, J. M., Tchana, F. K., and Fernandez, J. M. (2013). Raman and infrared spectra of the ν_1 band of oxirane. *Mol. Phys.* 111, 1983–1986. doi: 10.1080/00268976.2013.775516

Lane, J. L. (2018). *Frontiers and Advances in Molecular Spectroscopy*. Amsterdam: Elsevier.

Lee, C., Yang, W., and Parr, R. G. (1988). Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37, 785–789. doi: 10.1103/PhysRevB.37.785

Leung, H. O., Marshall, M. D., Vasta, A. L., and Craig, N. C. (2009). Microwave spectra of eight isotopic modifications of 1-chloro-1-fluoroethylene. *J. Mol. Spectrosc.* 253, 116–121. doi: 10.1016/j.jms.2008.11.002

Li, W., Spada, L., Tasinato, N., Rampino, S., Evangelisti, L., Gualandi, A., et al. (2018). Theory meets experiment for noncovalent complexes: the puzzling case of pnicogen interactions. *Angew. Chem. Int. Ed.* 57, 13853–13857. doi: 10.1002/anie.201807751

Licari, D., Baiardi, A., Egidi, F., Latouche, C., and Barone, V. (2015). Implementation of a graphical user interface for the virtual multifrequency spectrometer: the VMS draw tool. *J. Comput. Chem.* 36, 321–334. doi: 10.1002/jcc.23785

Licari, D., Tasinato, N., Spada, L., Puzzarini, C., and Barone, V. (2017). VMS-ROT: a new module of the virtual multifrequency spectrometer for simulation, interpretation, and fitting of rotational spectra. *J. Chem. Theory Comput.* 13, 4382–4396. doi: 10.1021/acs.jctc.7b00533

Lindon, J., Tranter, G. E., and Koppenaal, D. (2017). *Encyclopedia of Spectroscopy and Spectrometry*. Oxford: Academic Press.

Martin, J. M. L., Lee, T. J., Taylor, P. R., and Françoise, J.-P. (1995). The anharmonic force field of ethylene, C_2H_4, by means of accurate *ab initio* calculations. *J. Chem. Phys.* 103:2589. doi: 10.1063/1.469681

Medcraft, C., Thompson, C. D., Robertson, E. G., Appadoo, D. R. T., and McNaughton, D. (2012). The far-infrared rotational spectrum of ethylene oxide. *Astrophys. J.* 753:18. doi: 10.1088/0004-637X/753/1/18

Melli, A., Melosso, M., Tasinato, N., Bosi, G., Spada, L., Bloino, J., et al. (2018). Rotational and infrared spectroscopy of ethanimine: a route toward its astrophysical and planetary detection. *Astrophys. J.* 855:123. doi: 10.3847/1538-4357/aaa899

Miller, B. J., Howard, D. L., Lane, J. R., Kjaergaard, H. G., Dunn, M. E., and Vaida, V. (2009). SH-stretching vibrational spectra of ethanethiol and *tert*-butylthiol. *J. Phys. Chem. A* 113, 7576–7583. doi: 10.1021/jp9017162

Mills, I. M. (1972). "Vibration-rotation structure in asymmetric- and symmetric-top molecules," in *Molecular Spectroscopy: Modern Research*, eds K. N. Rao and C. W. Mathews (New York, NY: Academic Press), 115–140. doi: 10.1016/B978-0-12-580640-4.50013-3

Mller, H. S., and Brnken, S. (2005). Accurate rotational spectroscopy of sulfur dioxide, SO_2, in its ground vibrational and first excited bending states, v_2 = 0, 1, up to 2 thz. *J. Mol. Spectrosc.* 232, 213–222. doi: 10.1016/j.jms.2005.04.010

Obenchain, D. A., Spada, L., Alessandrini, S., Rampino, S., Herbers, S., Tasinato, D. N., et al. (2018). Unveiling the sulfur-sulfur bridge: Accurate structural and energetic characterization of a homochalcogen intermolecular bond. *Angew. Chem. Int. Ed.* 57, 15822–15826. doi: 10.1002/anie.201810637

Papajak, E., Leverentz, H. R., Zheng, J., and Truhlar, D. G. (2009). Efficient diffuse basis sets: cc-pVxZ+ and maug-cc-pVxZ. *J. Chem. Theory Comput.* 5, 1197–1202. doi: 10.1021/ct800575z

Papajak, E., Zheng, J., Xu, X. R., Leverentz, H., and Truhlar, D. G. (2011). Perspectives on basis sets beautiful: seasonal plantings of diffuse basis functions. *J. Chem. Theory Comput.* 7, 3027–3034. doi: 10.1021/ct200106a

Papoušek, D., and Aliev, M. R. (1982). *Molecular Vibrational/Rotational Spectra*. Amsterdam: Elsevier.

Penocchio, E., Piccardo, M., and Barone, V. (2015). Semiexperimental equilibrium structures for building blocks of organic and biological molecules: the B2PLYP route. *J. Chem. Theory Comput.* 11, 4342–4363. doi: 10.1021/acs.jctc.5b00622

Peveratti, R. (2020). Fitting elephants in the density functional zoo: statistical criteria for the evaluation of density functional theory methods as a suitable replacement for counting parameters. *Int. J. Quant. Chem.* 120, 1–15. doi: 10.1002/qua.26379

Piccardo, M., Bloino, J., and Barone, V. (2015a). Generalized vibrational perturbation theory for rotovibrational energies of linear, symmetric, and asymmetric tops: theory, approximations, and authomated approaches to deal with medium-to-large molecular systems. *Int. J. Quant. Chem.* 115, 948–982. doi: 10.1002/qua.24931

Piccardo, M., Penocchio, E., Puzzarini, C., Biczysko, M., and Barone, V. (2015b). Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: validation and application to semirigid organic molecules. *J. Phys. Chem. A* 119, 2058–2082. doi: 10.1021/jp511432m

Pietropolli Charmet, A., Stoppa, P., Tasinato, N., and Giorgianni, S. (2017). Computing sextic centrifugal distortion constants by DFT: a benchmark analysis on halogenated compounds. *J. Mol. Spectrosc.* 335, 117–125. doi: 10.1016/j.jms.2017.02.006

Pietropolli Charmet, A., Stoppa, P., Tasinato, N., Giorgianni, S., Barone, V., Biczysko, M., et al. (2013). An integrated experimental and quantum-chemical investigation on the vibrational spectra of chlorofluoromethane. *J. Chem. Phys.* 139:164302. doi: 10.1063/1.4825380

Pietropolli Charmet, A., Stoppa, P., Tasinato, N., Giorgianni, S., and Gambi, A. (2016). Study of the vibrational spectra and absorption cross sections of 1-chloro-1-fluoroethene by a joint experimental and *ab initio* approach. *J. Phys. Chem. A* 120, 8369–8386. doi: 10.1021/acs.jpca.6b07426

Presti, D., Pedone, A., Licari, D., and Barone, V. (2017). A modular implementation for the simulation of 1D and 2D solid-state NMR spectra of quadrupolar nuclei in the virtual multifrequency spectrometer-draw graphical interface. *J. Chem. Theory Comput.* 13, 2215–2229. doi: 10.1021/acs.jctc.7b00154

Pritchard, B. P., Altarawy, D., Didier, B., Gibson, T. D., and Windus, T. L. (2019). New basis set exchange: an open, up-to-date resource for the molecular sciences community. *J. Chem. Inform. Model.* 59, 4814–4820. doi: 10.1021/acs.jcim.9b00725

Puzzarini, C., and Barone, V. (2018). Diving for accurate structures in the ocean of molecular systems with the help of spectroscopy and quantum chemistry. *Acc. Chem. Res.* 51, 548–556. doi: 10.1021/acs.accounts.7b00603

Puzzarini, C., Biczysko, M., and Barone, V. (2010). Accurate harmonic/anharmonic vibrational frequencies for open-shell systems: performances of the B3LYP/N07D model for semirigid free radicals benchmarked by CCSD(T) computations. *J. Chem. Theory Comput.* 6, 828–838. doi: 10.1021/ct900594h

Puzzarini, C., Biczysko, M., Bloino, J., and Barone, V. (2014a). Accurate spectroscopic characterization of oxirane: a valuable route to its identification in titan's atmosphere and the assignment of unidentified infrared bands. *Astrophys. J.* 785:107. doi: 10.1088/0004-637X/785/2/107

Puzzarini, C., Bloino, J., Tasinato, N., and Barone, V. (2019a). Accuracy and interpretability: the devil and the holy grail. New routes across old boundaries in computational spectroscopy. *Chem. Rev.* 119, 8131–8191. doi: 10.1021/acs.chemrev.9b00007

Puzzarini, C., Senent, M. L., Domnguez-Gme, R., Carvajal, M., Hochlaf, M., and Al-Mogren, M. M. (2014b). Accurate spectroscopic characterization of ethyl mercaptan and dimethyl sulfide isotopologues: a route toward their astrophysical detection. *Astrophys. J.* 796:50. doi: 10.1088/0004-637X/796/1/50

Puzzarini, C., Tasinato, N., Bloino, J., Spada, L., and Barone, V. (2019b). State-of-the-art computation of the rotational and IR spectra of the methyl-cyclopropyl cation: hints on its detection in space. *Phys. Chem. Chem. Phys.* 21, 3615–3625. doi: 10.1039/C8CP04629H

Quack, M., and Merkt, F. (2011). *Handbook of High-Resolution Molecular Spectroscopy*. Chichester: John Wiley & Sons.

Russell, D. K., and Wesendrup, R. (2003). Tunable diode laser study of the ν_3 band of oxirane. *J. Mol. Spectrosc.* 217, 59–71. doi: 10.1016/S0022-2852(02)00010-3

Santra, G., Sylvetsky, N., and Martin, J. M. L. (2019). Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* 123, 5129–5143. doi: 10.1021/acs.jpca.9b03157

Smith, D., Devlin, J. P., and Scott, D. W. (1968). Conformational analysis of ethanethiol and 2-propanethiol. *J. Mol. Spectrosc.* 25, 174–184. doi: 10.1016/0022-2852(68)80004-9

Spada, L., Tasinato, N., Bosi, G., Vazart, F., Barone, V., and Puzzarini, C. (2017a). On the competition between weak OH$\cdots$F and CH$\cdots$F hydrogen bonds, in cooperation with CH$\cdots$O contacts, in the difluoromethane–tert-butyl alcohol cluster. *J. Mol. Spectrosc.* 337, 90–95. doi: 10.1016/j.jms.2017.04.001

Spada, L., Tasinato, N., Vazart, F., Barone, V., Caminati, W., and Puzzarini, C. (2017b). Noncovalent interactions and internal dynamics in pyridine-ammonia: a combined quantum-chemical and microwave spectroscopy study. *Chem. Eur. J.* 23, 4876–4883. doi: 10.1002/chem.201606014

Sugiki, T., Kobayashi, N., and Fujiwara, T. (2017). Modern technologies of solution nuclear magnetic resonance spectroscopy for three-dimensional structure determination of proteins open avenues for life scientists. *Comput. Struct. Biotechnol. J.* 15, 328–329. doi: 10.1016/j.csbj.2017.04.001

Tasinato, N. (2014). What are the spectroscopic properties of HFC 32? Answers from DFT. *Int. J. Quant. Chem.* 114, 1472–1485. doi: 10.1002/qua.24716

Tasinato, N., Ceselin, G., Stoppa, P., Pietropolli Charmet, A., and Giorgianni, S. (2018). A bit of sugar on TiO_2: quantum chemical insights on the interfacial interaction of glycolaldehyde over titanium dioxide. *J. Phys. Chem. C* 122, 6041–6051. doi: 10.1021/acs.jpcc.7b11911

Tasinato, N., and Grimme, S. (2015). Unveiling the non-covalent interactions of molecular homodimers by dispersion-corrected DFT calculations and collision-induced broadening of ro-vibrational transitions: application to $(CH_2F_2)_2$ and $(SO_2)_2$. *Phys. Chem. Chem. Phys.* 17, 5659–5669. doi: 10.1039/C4CP05680A

Tasinato, N., Moro, D., Stoppa, P., Pietropolli Charmet, A., Toninello, P., and Giorgianni, S. (2015). Adsorption of F_2CCFCl on TiO_2 nano-powder: structures, energetics and vibrational properties from DRIFT spectroscopy and periodic quantum chemical calculations. *Appl. Surf. Sci.* 353, 986–994. doi: 10.1016/j.apsusc.2015.07.006

Tasinato, N., Pietropolli Charmet, A., Stoppa, P., Giorgianni, S., and Buffa, G. (2010). Spectroscopic measurements of SO_2 line parameters in the 9.2 μm atmospheric region and theoretical determination of self-broadening coefficients. *J. Chem. Phys.* 132:044315. doi: 10.1063/1.3299274

Tasinato, N., Pietropolli Charmet, A., Stoppa, P., Giorgianni, S., and Gambi, A. (2012a). Quantum-chemical ab initio investigation of the vibrational spectrum of halon 1113 and its anharmonic force field: a joint experimental and computational approach. *Chem. Phys.* 397, 55–64. doi: 10.1016/j.chemphys.2011.12.015

Tasinato, N., Puzzarini, C., and Barone, V. (2017). Correct modeling of cisplatin: a paradigmatic case. *Angew. Chem. Int. Ed.* 56, 13838–13841. doi: 10.1002/anie.201707683

Tasinato, N., Regini, G., Stoppa, P., Pietropolli Charmet, A., and Gambi, A. (2012b). Anharmonic force field and vibrational dynamics of CH_2F_2 up to 5000 $cm^{-}1$ studied by Fourier transform infrared spectroscopy and state-of-the-art ab initio calculations. *J. Chem. Phys.* 136:214302. doi: 10.1063/1.4720502

Watson, J. K. (2005). The vibrational dependence of quartic centrifugal distortion. *J. Mol. Struct.* 742, 91–98. doi: 10.1016/j.molstruc.2005.01.009

Westbook, B. R., Rio, W. A. D., Lee, T. J., and Fortenberry, R. C. (2020). Overcoming the out-of-plane bending issue in an aromatic hydrocarbon: the anharmonic vibrational frequencies of c-$(CH)C_3H_2^+$. *Phys. Chem. Chem. Phys.* 22, 12951–12958. doi: 10.1039/D0CP01889A

Wolff, H., and Szydlowski, J. (1985). Vibrational spectra and rotational isomerism of ethanethiol and ethanethiol-d_1. *Can. J. Chem.* 63, 1708–1712. doi: 10.1139/v85-287

Woon, D. E., and Dunning, T. H. J. (1993). Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* 98, 1358–1371. doi: 10.1063/1.464303

Ye, H., Mendolicchio, M., Kruse, H., Puzzarini, C., Biczysko, M., and Barone, V. (2020). The challenging equilibrium structure of HSSH: another success of the rotational spectroscopy/quantum chemistry synergism. *J. Mol. Struct.* 1211:127933. doi: 10.1016/j.molstruc.2020.127933

Zerbetto, M., Licari, D., Barone, V., and Polimeno, A. (2013). Computational tools for the interpretation of electron spin resonance spectra in solution. *Mol. Phys.* 111, 2746–2756. doi: 10.1080/00268976.2013.800602

Zhao, Y., and Truhlar, D. G. (2005). Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A* 109, 5656–5667. doi: 10.1021/jp050536c

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 05 November 2020
doi: 10.3389/fchem.2020.589432

TAO-DFT-Based *Ab Initio* Molecular Dynamics

Shaozhi Li[1] *and Jeng-Da Chai*[1,2]*

[1] *Department of Physics, National Taiwan University, Taipei, Taiwan,* [2] *Center for Theoretical Physics and Center for Quantum Science and Engineering, National Taiwan University, Taipei, Taiwan*

Recently, AIMD (*ab initio* molecular dynamics) has been extensively employed to explore the dynamical information of electronic systems. However, it remains extremely challenging to reliably predict the properties of nanosystems with a radical nature using conventional electronic structure methods (e.g., Kohn-Sham density functional theory) due to the presence of static correlation. To address this challenge, we combine the recently formulated TAO-DFT (thermally-assisted-occupation density functional theory) with AIMD. The resulting TAO-AIMD method is employed to investigate the instantaneous/average radical nature and infrared spectra of *n*-acenes containing *n* linearly fused benzene rings ($n = 2$–8) at 300 K. According to the TAO-AIMD simulations, on average, the smaller *n*-acenes (up to $n = 5$) possess a nonradical nature, and the larger *n*-acenes ($n = 6$–8) possess an increasing radical nature, showing remarkable similarities to the ground-state counterparts at 0 K. Besides, the infrared spectra of *n*-acenes obtained with the TAO-AIMD simulations are in qualitative agreement with the existing experimental data.

Keywords: TAO-DFT, AIMD, static correlation, radical nature, infrared spectra

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Chongjie Mo,
Beijing Computational Science Research Center, China
Fujie Tang,
Temple University, United States

***Correspondence:**
Jeng-Da Chai
jdchai@phys.ntu.edu.tw

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *30 July 2020*
Accepted: *02 September 2020*
Published: *05 November 2020*

Citation:
Li S and Chai J-D (2020) TAO-DFT-Based Ab Initio Molecular Dynamics. Front. Chem. 8:589432. doi: 10.3389/fchem.2020.589432

1. INTRODUCTION

Molecular dynamics (MD) is a computational method for simulating dynamical processes that occur in a system consisting of atoms (for example, atoms, molecules, solids, and liquids) (Lifson and Warshel, 1968; Levitt and Lifson, 1969; Karplus and Petsko, 1990; Kresse and Hafner, 1993; Sprik et al., 1996; Silvestrelli et al., 1997; Putrino and Parrinello, 2002; Tuckerman, 2002; Chai et al., 2003; Kuo and Mundy, 2004; Jensen, 2007; Marx and Hutter, 2009; Gaigeot, 2010; Ramírez-Solís et al., 2011; Vitale et al., 2015; Hollingsworth et al., 2018). The motion of the atomic nuclei in the system is described by the classical Newtonian equations of motion, starting from pre-specified initial conditions (e.g., initial nuclear positions and velocities) and subject to boundary conditions suitable for the system studied. By performing MD simulations, one can calculate both the dynamical and equilibrium thermodynamic properties associated with a system at non-zero temperatures and can simultaneously monitor the microscopic movements of the atomic nuclei in the system. However, the specification of the forces acting on the atomic nuclei in MD simulations remains challenging, yielding two popular types of MD methods: classical MD and *ab initio* MD (AIMD).

In classical MD, the forces that act on the atomic nuclei are calculated using an empirical potential energy function (i.e., the potential energy of a system is expressed as a function of the nuclear coordinates) that is defined by molecular mechanics (Tuckerman, 2002; Jensen, 2007; Marx and Hutter, 2009). Because of a simple analytical function for the potential energy, classical MD is computationally very efficient, and it has thus been widely employed for many applications

(Lifson and Warshel, 1968; Levitt and Lifson, 1969; Karplus and Petsko, 1990; Hollingsworth et al., 2018). For example, classical MD can reveal the details of the motion of macromolecules (e.g., proteins) (Lifson and Warshel, 1968; Levitt and Lifson, 1969), which can be very challenging using experimental approaches. However, in classical MD, empirical potential energy functions are typically parameterized for some peculiar systems and may not be transferable to others (Tuckerman, 2002). Furthermore, in classical MD, electrons are not present explicitly (i.e., their effects are approximated by empirical potential energy functions), and the electronic properties (e.g., electron density and radical nature) of systems therefore cannot be explored.

To overcome the shortcomings of classical MD, one can resort to AIMD, wherein the forces that act on the atomic nuclei are calculated on-the-fly according to the potential energy obtained with an electronic structure method along an AIMD trajectory (Tuckerman, 2002; Jensen, 2007; Marx and Hutter, 2009). Clearly, the accuracy of AIMD simulations can be limited due to the accuracy of the underlying electronic structure method. On the other hand, for AIMD simulations, it is also essential to adopt a computationally efficient electronic structure method that can provide reasonably accurate potential energy for the systems of interest, since there are typically 10^4–10^6 time steps per AIMD trajectory. Owing to the decent balance between performance and cost, KS-DFT (i.e., Kohn-Sham density functional theory) (Kohn and Sham, 1965) is currently the most popular electronic structure method adopted in AIMD simulations. Over the past three decades, KS-DFT-based AIMD (KS-AIMD) has been commonly used to investigate the equilibrium thermodynamic and dynamical properties associated with electronic systems at finite temperatures, and has been a powerful AIMD method for modeling various phenomena (Kresse and Hafner, 1993; Sprik et al., 1996; Silvestrelli et al., 1997; Putrino and Parrinello, 2002; Chai et al., 2003; Kuo and Mundy, 2004; Gaigeot, 2010; Ramírez-Solís et al., 2011; Vitale et al., 2015).

Nevertheless, KS-DFT using traditional exchange-correlation (XC) energy functionals could yield qualitative failures for the various properties associated with electronic systems possessing radical nature due to the lack of static correlation effects (Cohen et al., 2008, 2012). For example, consider linear n-acene (with the chemical formula $C_{4n+2}H_{2n+4}$), which contains n linearly fused benzene rings (for example, see **Figure 1**), $N_e = 26n + 16$ electrons, and $N = 6n+6$ nuclei. According to the recent findings (Hachmann et al., 2007; Chai, 2012, 2014, 2017; Rivero et al., 2013; Wu and Chai, 2015; Fosso-Tande et al., 2016), the larger n-acenes ($n \geq 6$) possess increasing radical nature in their ground states. However, KS-DFT adopting commonly used XC energy functionals could perform poorly for the larger n-acenes ($n \geq 6$) due to the lack of static correlation effects (Hachmann et al., 2007; Chai, 2012, 2017). It can therefore be anticipated that KS-AIMD simulations at finite temperatures may yield unreliable dynamical information for electronic systems with a radical nature [e.g., the larger n-acenes ($n \geq 6$)].

To study the properties of electronic systems with a radical nature, one generally resorts to multi-reference (MR) electronic structure methods (Andersson et al., 1992; Hachmann et al., 2007; Gidofalvi and Mazziotti, 2008; Pelzer et al., 2011; Gryn'ova et al., 2015; Fosso-Tande et al., 2016; Battaglia et al., 2017; Mullinax et al., 2019). Although MR electronic structure methods can reliably predict the various properties associated with electronic systems with radical nature, they could be prohibitively expensive for a single-point energy + nuclear gradient calculation on large electronic systems, not to mention the respective AIMD simulations where such calculations should be performed about 10^4–10^6 times per AIMD trajectory. To explore the dynamical properties of nanosystems with radical nature using AIMD simulations at finite temperatures, therefore, it is essential to employ an efficient electronic structure method that can properly describe the static correlation effects during the AIMD simulations.

Recently, TAO-DFT (thermally-assisted-occupation density functional theory) (Chai, 2012) has been formulated for investigating the ground-state (GS) properties associated with nanosystems possessing a radical nature at 0 K. Note that TAO-DFT is a density functional theory adopting fractional orbital occupation numbers (i.e., generated by the Fermi-Dirac distribution using a fictitious temperature θ), which is rather different from KS-DFT. In TAO-DFT, the fictitious temperature θ is intimately correlated with the effect of the configuration mixing on the electron density of a GS system (e.g., see sections 2, 3 of Chai, 2012), and it is thus completely unrelated to the physical temperature (i.e., 0 K) of the GS system. Note that KS-DFT corresponds to TAO-DFT with $\theta = 0$. However, for a general GS system at 0 K, the θ value in TAO-DFT can be nonzero (e.g., see sections 3.5, 4 of Chai, 2012). In TAO-DFT, an entropy contribution term, which is dependent on the fictitious temperature θ and orbital occupation numbers, can offer a proper description for static correlation even when the simplest LDA (i.e., local density approximation) XC energy functional is used (Chai, 2012). At the LDA level, even with a properly defined system-independent θ value (e.g., see section 5 of Chai, 2012), TAO-DFT, which has similar cost as KS-DFT in computation, has

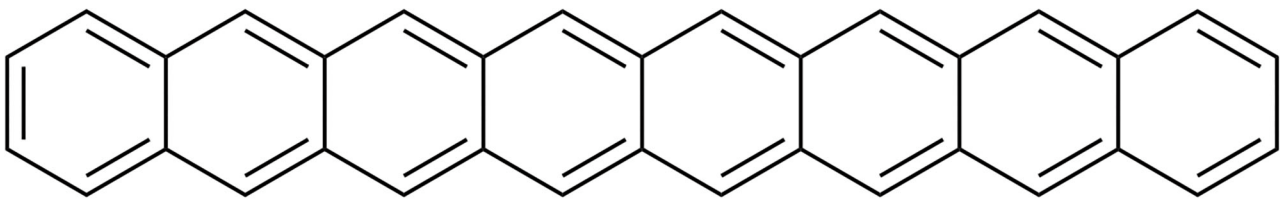

FIGURE 1 | Structure of 8-acene, containing eight linearly fused benzene rings.

been shown to consistently outperform KS-DFT for electronic systems with a radical nature, while performing similarly to KS-DFT for electronic systems with a nonradical nature (Chai, 2012).

More complicated semilocal (Chai, 2014), global hybrid (Chai, 2017), and range-separated hybrid (Chai, 2017; Xuan et al., 2019) XC energy functionals could also be used in TAO-DFT. Moreover, in order to enhance its accuracy for a wide range of applications, a method that determines the θ value in TAO-DFT in a self-consistent manner has recently been developed (Lin et al., 2017). Very recently, a frequency-domain formulation of linear-response time-dependent TAO-DFT (Yeh et al., 2020) has been formulated to explore the properties of electronic excited states within TAO-DFT.

Owing to its computational efficiency and decent accuracy for exploring the properties of electronic systems at the nanoscale, TAO-DFT has recently been adopted to investigate the electronic properties (Wu and Chai, 2015; Seenithurai and Chai, 2016, 2017, 2018, 2019, 2020; Wu et al., 2016; Yeh and Chai, 2016; Yeh et al., 2018; Chung and Chai, 2019; Deng and Chai, 2019; Hanson-Heine and Hirst, 2020; Hanson-Heine et al., 2020; Huang et al., 2020; Manassir and Pakiari, 2020), hydrogen storage properties (Seenithurai and Chai, 2016, 2017, 2018), and vibrational frequencies (Hanson-Heine, 2020) of several electronic systems at the nanoscale, especially for those possessing a radical nature.

Therefore, in the present work, we propose to combine TAO-DFT (Chai, 2012) with AIMD, yielding TAO-DFT-based AIMD (TAO-AIMD). Since analytical nuclear gradients for TAO-DFT are available (Chai, 2012), TAO-AIMD is as computationally efficient as KS-AIMD. Accordingly, it is feasible to study the equilibrium thermodynamic and dynamical properties of nanosystems with a radical nature using TAO-AIMD simulations at finite temperatures. To highlight some of the present capabilities of TAO-AIMD, we perform TAO-AIMD simulations to explore the instantaneous/average radical nature and infrared (IR) spectra of n-acenes with $n = 2$–8 fused benzene rings at 300 K. The rest of this paper is organized as follows. The TAO-AIMD method is defined in section 2. We then describe the computational details in section 3, discuss the results in section 4, and give our conclusions in section 5.

2. TAO-AIMD

Consider a system containing N_e electrons (described by coordinates $\mathbf{r}_1, ..., \mathbf{r}_{N_e}$) and N nuclei (described by coordinates $\mathbf{R}_1, ..., \mathbf{R}_N$). Here, we resort to the adiabatic or Born-Oppenheimer (BO) approximation (Born and Oppenheimer, 1927). Because the electrons are much lighter than the nuclei, it is assumed that the nuclei move relatively slowly, and hence the electrons are able to respond to the nuclear motion almost instantaneously. In other words, the electronic motion and nuclear motion can be treated separately.

Accordingly, in the first step of the BO approximation, the kinetic energy of the nuclei is ignored. For fixed nuclear positions $\mathbf{R}_1, ..., \mathbf{R}_N$, the electronic Hamiltonian is expressed as (Tuckerman, 2002; Jensen, 2007; Marx and Hutter, 2009)

$$\hat{H}_{\text{elec}} = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{N_e}\nabla_i^2 - \frac{e^2}{4\pi\epsilon_0}\sum_{i=1}^{N_e}\sum_{A=1}^{N}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \frac{e^2}{4\pi\epsilon_0}\sum_{i=1}^{N_e}\sum_{j>i}^{N_e}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1)$$

where m_e and $-e$ are the mass and charge, respectively, of an electron, and $Z_A e$ is the charge of nucleus A. On the right-hand side of Equation (1), the first term is the kinetic energy of electrons, the second term is the nuclear-electron attraction energy, and the third term is the electron-electron repulsion energy. The time-independent electronic Schrödinger equation

$$\hat{H}_{\text{elec}}\Psi_k = E_k\Psi_k \quad (2)$$

is subsequently solved for the electronic energy E_k and electronic wavefunction Ψ_k. In particular, the lowest eigenvalue E_0 is the GS electronic energy, and the corresponding eigenfunction Ψ_0 is the GS electronic wavefunction. By adding the nuclear-nuclear repulsion energy to the electronic energy, one obtains the potential energy for the k-th electronic eigenstate of the system

$$U_k = E_k + \frac{e^2}{4\pi\epsilon_0}\sum_{A=1}^{N}\sum_{B>A}^{N}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}. \quad (3)$$

By varying the nuclear positions and solving the corresponding time-independent electronic Schrödinger equation, U_k can be expressed as a function of the nuclear positions, also known as the potential energy surface of the k-th electronic eigenstate. In the second step of the BO approximation, the kinetic energy of the nuclei is reintroduced, and in principle, the nuclear dynamics is described by the time-dependent nuclear Schrödinger equation, evolving on a potential energy surface, e.g., $U_k(\mathbf{R}_1, ..., \mathbf{R}_N)$.

Similar to most AIMD simulations (Tuckerman, 2002; Jensen, 2007; Marx and Hutter, 2009; Gaigeot, 2010; Ramírez-Solís et al., 2011; Vitale et al., 2015), in this work, we assume that nuclear quantum effects could be ignored, and nuclear motion occurs only on the GS potential energy surface (i.e., the potential energy surface of the electronic GS), $U_0(\mathbf{R}_1, ..., \mathbf{R}_N)$. Since the time-independent electronic Schrödinger equation is only applicable to very small electronic systems, a sufficiently efficient electronic structure method for the determination of electronic ground state is typically required for AIMD simulations. As mentioned previously, KS-AIMD simulations can be unreliable for electronic systems with a radical nature. On the other hand, AIMD simulations employing MR electronic structure methods are very likely to be computationally intractable for most electronic systems.

To resolve this issue with minimum computational expense, in the present work, we combine TAO-DFT (Chai, 2012) with AIMD, yielding the TAO-AIMD method. Specifically, in TAO-AIMD, the nuclei are treated as classical particles, obeying the

classical nuclear Hamiltonian

$$H_{\text{nucl}}(\mathbf{P}_1(t),...,\mathbf{P}_N(t),\mathbf{R}_1(t),...,\mathbf{R}_N(t)) = \sum_{A=1}^{N}\frac{|\mathbf{P}_A(t)|^2}{2M_A} + U_0^{\text{TAO-DFT}}(\mathbf{R}_1(t),...,\mathbf{R}_N(t)), \quad (4)$$

where M_A is the mass of nucleus A, and $\mathbf{P}_A(t)$ is the momentum of nucleus A at time t. The left-hand side of Equation (4) gives the total energy at time t [i.e., $E(t)$]. On the right-hand side of Equation (4), the first term is the nuclear kinetic energy at time t, and the second term is the potential energy of the electronic GS obtained with TAO-DFT for the nuclear positions at time t, i.e., $\mathbf{R}_1(t),...,\mathbf{R}_N(t)$:

$$U_0^{\text{TAO-DFT}}(\mathbf{R}_1(t),...,\mathbf{R}_N(t)) = E_0^{\text{TAO-DFT}}(\mathbf{R}_1(t),...,\mathbf{R}_N(t)) + \frac{e^2}{4\pi\epsilon_0}\sum_{A=1}^{N}\sum_{B>A}^{N}\frac{Z_A Z_B}{|\mathbf{R}_A(t)-\mathbf{R}_B(t)|}, \quad (5)$$

where $E_0^{\text{TAO-DFT}}(\mathbf{R}_1(t),...,\mathbf{R}_N(t))$ (e.g., see Equation 27 of Chai, 2012) is the corresponding GS electronic energy obtained with TAO-DFT. On the basis of Equation (4), the nuclei move based on Newton's equations of motion on the GS potential energy surface generated by TAO-DFT:

$$\dot{\mathbf{R}}_A(t) = \frac{\mathbf{P}_A(t)}{M_A} \quad (6)$$

$$\dot{\mathbf{P}}_A(t) = -\nabla_A U_0^{\text{TAO-DFT}}(\mathbf{R}_1(t),...,\mathbf{R}_N(t)), \quad (7)$$

where $\dot{\mathbf{R}}_A(t)$ is the velocity of nucleus A at time t, $\dot{\mathbf{P}}_A(t)$ is the time derivative of the momentum of nucleus A at time t, and the right-hand side of Equation (7) gives the force acting on nucleus A at time t [i.e., $\mathbf{F}_A(t)$].

Equations (4)–(7) form the theoretical basis of the TAO-AIMD method. Given the initial nuclear positions $\mathbf{R}_1(0),...,\mathbf{R}_N(0)$ and velocities $\dot{\mathbf{R}}_1(0),...,\dot{\mathbf{R}}_N(0)$, all the future nuclear positions $\mathbf{R}_1(t),...,\mathbf{R}_N(t)$ and velocities $\dot{\mathbf{R}}_1(t),...,\dot{\mathbf{R}}_N(t)$ are determined by Equations (5)–(7), generating a TAO-AIMD trajectory (i.e., TAO-AIMD is deterministic). Note that the GS potential energy surface and the forces that act on the nuclei can be computed on-the-fly using TAO-DFT, as needed along the TAO-AIMD trajectory. According to the definitions presented here (Marx and Hutter, 2009), TAO-AIMD can also be regarded as TAO-DFT-based Born-Oppenheimer MD (BOMD), i.e., TAO-BOMD.

Note that the GS potential energy surface and the forces that act on the nuclei are computed using TAO-DFT in TAO-AIMD, while they are computed using KS-DFT in KS-AIMD. Since TAO-DFT is as computationally efficient as KS-DFT (e.g., see Chai, 2012 for details), TAO-AIMD is similar to KS-AIMD in computational expense. On the other hand, AIMD simulations employing MR electronic structure methods are very likely to be computationally infeasible for most electronic systems. Accordingly, TAO-AIMD can be a very promising method for exploring the equilibrium thermodynamic and dynamical properties of nanosystems with a radical nature at finite temperatures. In addition, existing XC energy functionals defined in KS-DFT could also be employed in TAO-DFT (Chai, 2012, 2014, 2017) and TAO-AIMD. For electronic systems with a nonradical nature, TAO-DFT performs similarly to KS-DFT, and TAO-AIMD should therefore perform similarly to KS-AIMD.

3. COMPUTATIONAL DETAILS

All calculations are performed using TAO-LDA (Chai, 2012), i.e., TAO-DFT employing the LDA XC and θ-dependent energy functionals, where the suggested fictitious temperature θ = 7 mhartree is employed (Chai, 2012). Almost all numerical data are obtained with Q-Chem 4.4 (Shao, 2015), using the 6-31G(d) basis set and a numerical grid consisting of 75 Euler-Maclaurin radial grid points and 302 Lebedev angular grid points, wherein isolated boundary conditions (i.e., well-suited for the study of atoms and molecules) are employed. The TAO-AIMD-based IR spectra are computed using the program TRAVIS (Brehm and Kirchner, 2011; Thomas et al., 2013, 2015; Brehm et al., 2020).

In several recent investigations (Chai, 2012; Wu and Chai, 2015; Fosso-Tande et al., 2016; Yeh and Chai, 2016; Mullinax et al., 2019), the orbital occupation numbers obtained from TAO-LDA (with θ = 7 mhartree) have been found to be similar to the natural orbital occupation numbers obtained from a very accurate MR electronic structure method that can be applied to treat relatively large active spaces, leading to a qualitatively similar tendency for the radical nature associated with various PAHs (i.e., polycyclic aromatic hydrocarbons).

On the other hand, because of the constraint of symmetry, the spin-unrestricted and spin-restricted calculations based on an exact electronic structure method should lead to the same energy values for the lowest singlet state (i.e., GS) of n-acene (Chai, 2012; Rivero et al., 2013; Gryn'ova et al., 2015). Nevertheless, KS-DFT employing conventional XC energy functionals fails to obey this constraint for the larger n-acenes (which are electronic systems possessing radical nature), leading to the unphysical symmetry-breaking effects in the corresponding spin-unrestricted calculations (Cohen et al., 2008, 2012). In our previous studies (Chai, 2012, 2014; Wu and Chai, 2015), the spin-unrestricted and spin-restricted GS (i.e., lowest singlet state) energy values of n-acene (up to n = 100), calculated by TAO-LDA (with θ = 7 mhartree), have been found to be essentially the same, leading to essentially no unphysical symmetry-breaking effects in the corresponding spin-unrestricted calculations. For computational efficiency, therefore, all TAO-LDA calculations in this work are spin-restricted calculations unless noted otherwise.

For all TAO-AIMD simulations, a time step of 20 a.u. ($\approx$ 0.484 fs) is adopted for the integration of the equations of motion. For each TAO-AIMD simulation, the initial geometry of n-acene is chosen as the GS geometry of n-acene obtained with TAO-LDA, and the initial nuclear velocities of n-acene are randomly selected from the Maxwell-Boltzmann (MB) distribution at T = 300 K. To equilibrate n-acene at T = 300 K, the TAO-AIMD simulation

is first performed in the canonical (*NVT*) ensemble with the aid of the Nosé-Hoover (NH) chain thermostat (Martyna et al., 1992) (with the NH chain length of three auxiliary variables and NH timescale of 200 fs in the implementation of Q-Chem) for about 10.2 ps. To avoid any interference with the dynamics (Gaigeot and Sprik, 2003), we subsequently remove the NH chain thermostat and continue the TAO-AIMD simulation in the microcanonical (*NVE*) ensemble for a total of 10,500 time steps (≈ 5.1 ps), at which time the total energy of *n*-acene is well converged. After equilibration, we continue the TAO-AIMD simulation in the *NVE* ensemble and collect relevant data along the equilibrated TAO-AIMD trajectory for a total of 42,000 time steps (≈ 20.3 ps), where the average temperature is 300 ± 1 K. As the initial nuclear velocities of *n*-acene for each TAO-AIMD simulation are *randomly selected* from the MB distribution at 300 K, the aforementioned processes are repeated to generate a total of four different equilibrated TAO-AIMD trajectories (≈ 20.3 ps per trajectory).

Furthermore, to demonstrate the significance of TAO-AIMD simulations for exploring the dynamical information of large molecules with radical nature, we also perform preliminary calculations on 8-acene to examine the possible symmetry-breaking effects in the spin-unrestricted TAO-AIMD/KS-AIMD simulations (see **Supplementary Section 1** and **Supplementary Figures 1–3**).

4. RESULTS AND DISCUSSION

4.1. Symmetrized von Neumann Entropy

In TAO-DFT, the radical nature of a GS molecule can be examined by the symmetrized von Neumann entropy (for example, see Equation 5 of Chung and Chai, 2019 for the spin-unrestricted case). For the spin-restricted case, the symmetrized von Neumann entropy of a GS molecule can be expressed as

$$S_{\mathrm{vN}} = -\sum_{i=1}^{\infty}\left\{\frac{f_i}{2}\ln\left(\frac{f_i}{2}\right)+\left(1-\frac{f_i}{2}\right)\ln\left(1-\frac{f_i}{2}\right)\right\}, \quad (8)$$

where f_i (i.e., a number between 0 and 2) is the occupation number of the i-th orbital of the GS molecule, obtained with spin-restricted TAO-DFT. Note that f_i is closely related to the corresponding natural orbital occupation number (Löwdin and Shull, 1956; Chai, 2012, 2014, 2017). For a GS molecule possessing a nonradical nature, the occupation numbers associated with all orbitals are very close to either 0 or 2, yielding a vanishingly small S_{vN} value. Nonetheless, for a GS molecule with a significant radical nature, the active orbital occupation numbers can deviate significantly from 0 and 2 (for example, 0.2–1.8); hence, the corresponding S_{vN} value can greatly increase as the number of active orbitals increases and/or the active orbital occupation numbers are closer to 1 (Rivero et al., 2013; Chai, 2014, 2017; Wu and Chai, 2015; Seenithurai and Chai, 2016, 2017, 2018, 2019; Wu et al., 2016; Yeh et al., 2018; Chung and Chai, 2019; Deng and Chai, 2019; Huang et al., 2020).

On the basis of Equation (8), in a spin-restricted TAO-AIMD simulation, the symmetrized von Neumann entropy of a molecule at time t along a TAO-AIMD trajectory can be defined as

$$S_{\mathrm{vN}}(t) = -\sum_{i=1}^{\infty}\left\{\frac{f_i(t)}{2}\ln\left(\frac{f_i(t)}{2}\right)+\left(1-\frac{f_i(t)}{2}\right)\ln\left(1-\frac{f_i(t)}{2}\right)\right\}, \quad (9)$$

where $f_i(t)$ (i.e., a value between 0 and 2) is the occupation number of the i^{th} orbital of the molecule, obtained with spin-restricted TAO-DFT for the nuclear positions at time t, i.e., $\mathbf{R}_1(t), ..., \mathbf{R}_N(t)$, along the TAO-AIMD trajectory. Accordingly, the time average of $S_{\mathrm{vN}}(t)$ along the TAO-AIMD trajectory is calculated by

$$\overline{S_{\mathrm{vN}}} = \frac{1}{\tau}\int_0^{\tau} S_{\mathrm{vN}}(t)dt, \quad (10)$$

where τ is the total time duration of the TAO-AIMD trajectory.

To investigate the radical nature of *n*-acene in the TAO-AIMD simulations, the $\overline{S_{\mathrm{vN}}}$ value of *n*-acene along each equilibrated TAO-AIMD trajectory is computed using Equation (10), and the reported $\overline{S_{\mathrm{vN}}}$ value of *n*-acene is an average over four different equilibrated TAO-AIMD trajectories (≈ 20.3 ps per trajectory). For comparison, we also report the S_{vN} value of GS *n*-acene (given by Equation 8), corresponding to the $\overline{S_{\mathrm{vN}}}$ value of *n*-acene at 0 K in this work.

As presented in **Figure 2**, the $\overline{S_{\mathrm{vN}}}$ value of *n*-acene obtained with the TAO-AIMD simulations at 300 K is very close to the S_{vN} value of GS *n*-acene Chai (2014), increasing monotonically with increasing *n* (also see **Supplementary Table 1**). This suggests that similar to the GS counterparts at 0 K (Hachmann et al., 2007; Chai, 2012, 2014, 2017), on average, the smaller *n*-acenes (up to $n = 5$) possess nonradical nature, and the larger *n*-acenes ($n = 6$–8) possess increasing radical nature in AIMD simulations at 300 K.

Along each equilibrated TAO-AIMD trajectory, the instantaneous $S_{\mathrm{vN}}(t)$ value of *n*-acene fluctuates over time and can thus be even larger than the $\overline{S_{\mathrm{vN}}}$ value of *n*-acene (see **Figure 3** for $n = 8$, and **Supplementary Figures 4–9** for others). Therefore, for the larger *n*-acenes ($n = 6$–8), it is essential to perform AIMD simulations at 300 K with an efficient electronic structure method that can reliably describe strong static correlation effects, well-justifying the use of TAO-AIMD in this work.

4.2. Active Orbital Occupation Numbers

As mentioned previously, in a spin-restricted TAO-AIMD simulation, $f_i(t)$ (i.e., a value between 0 and 2) is the occupation number of the i-th orbital of a molecule, obtained with spin-restricted TAO-DFT for the nuclear positions at time t, i.e., $\mathbf{R}_1(t), ..., \mathbf{R}_N(t)$, along the TAO-AIMD trajectory. Consequently, the time average of $f_i(t)$ along the TAO-AIMD trajectory is calculated by

$$\overline{f_i} = \frac{1}{\tau}\int_0^{\tau} f_i(t)dt, \quad (11)$$

where τ is the total time duration of the TAO-AIMD trajectory.

To further assess the radical nature of *n*-acene in the TAO-AIMD simulations, the $\overline{f_i}$ value of *n*-acene along each equilibrated

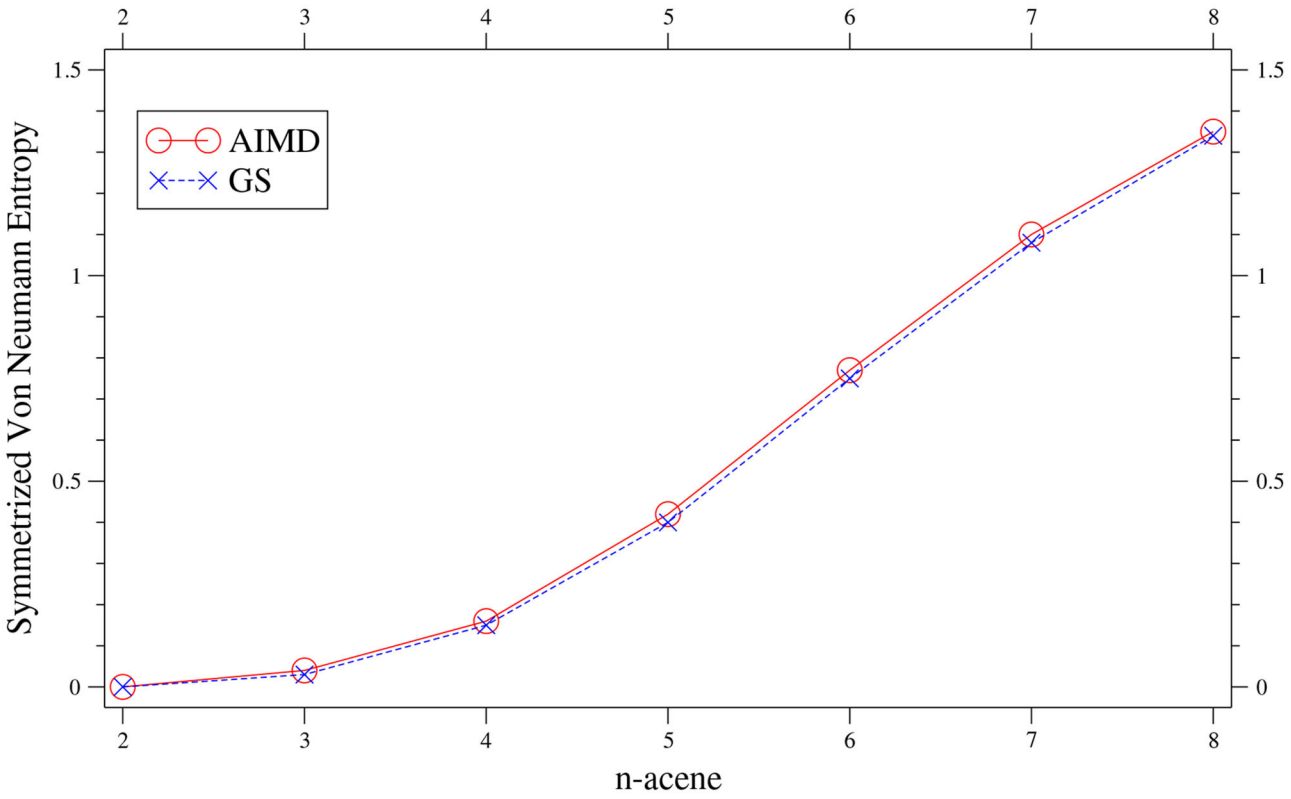

FIGURE 2 | Symmetrized von Neumann entropy of *n*-acene, obtained with the TAO-AIMD simulations at 300 K and GS calculation, calculated by TAO-LDA.

TAO-AIMD trajectory is computed using Equation (11), and the reported $\overline{f_i}$ value of *n*-acene is an average over four different equilibrated TAO-AIMD trajectories ($\approx$ 20.3 ps per trajectory). For comparison, we also present the f_i value of GS *n*-acene, which corresponds to the $\overline{f_i}$ value of *n*-acene at 0 K in this work.

For *n*-acene (containing N_e electrons), we define the HOMO (i.e., highest occupied molecular orbital) as the $(N_e/2)$th orbital, the LUMO (i.e., lowest unoccupied molecular orbital) as the $(N_e/2 + 1)$th orbital, and so forth (Chai, 2012, 2017; Wu and Chai, 2015; Wu et al., 2016; Yeh et al., 2018; Chung and Chai, 2019; Deng and Chai, 2019; Seenithurai and Chai, 2019; Huang et al., 2020). For brevity, HOMO and LUMO are denoted as H and L, respectively. In addition, the orbitals with an occupation number (0.2–1.8) are regarded as the active orbitals.

As presented in **Figure 4**, the $\overline{f_i}$ value of *n*-acene obtained with the TAO-AIMD simulations at 300 K is very close to the f_i value of GS *n*-acene (Chai, 2012) (also see **Supplementary Table 2**). This implies that on average, the radical nature of *n*-acene obtained with the TAO-AIMD simulations at 300 K is very similar to that obtained with the GS calculation at 0 K, showing consistency with the analysis of symmetrized von Neumann entropy. For smaller *n* values (e.g., up to $n = 5$), the occupation numbers associated with all orbitals are very close to either 0 or 2. Therefore, the smaller *n*-acenes should exhibit nonradical nature. However, as *n* increases, the number of active orbitals increases and/or the active orbital occupation numbers are closer to 1, apparently showing that the larger *n*-acenes ($n = 6$–8) should exhibit increasing radical character. This clearly indicates that similar to the GS counterparts at 0 K (Hachmann et al., 2007; Chai, 2012, 2014, 2017), on average, the larger *n*-acenes ($n = 6$–8) should possess increasing radical nature in AIMD simulations at 300 K.

In addition, the instantaneous $f_i(t)$ value of *n*-acene along each equilibrated TAO-AIMD trajectory fluctuates around an average, implying that the instantaneous radical nature of *n*-acene can be more pronounced than the average radical nature of *n*-acene in the TAO-AIMD simulations (see **Figure 5** for $n = 8$, and **Supplementary Figures 10–15** for others). According to our findings, the radical nature of the larger *n*-acenes ($n = 6$–8) can persist in AIMD simulations at 300 K. For such molecules, KS-AIMD simulations can therefore be unreliable, and AIMD simulations employing MR electronic structure methods are very likely to be computationally intractable. Accordingly, this highlights the significance of TAO-AIMD simulations for exploring the dynamical information of large molecules with radical nature.

4.3. IR Spectra

IR spectroscopy, which involves the interaction of IR radiation with matter, has been extensively used to explore the structure of materials nondestructively in recent years (Roggo et al., 2007; Joblin et al., 2011; Beć and Huck, 2019). Recently, the IR spectra of PAHs (e.g., *n*-acenes) have attracted much interest since PAHs

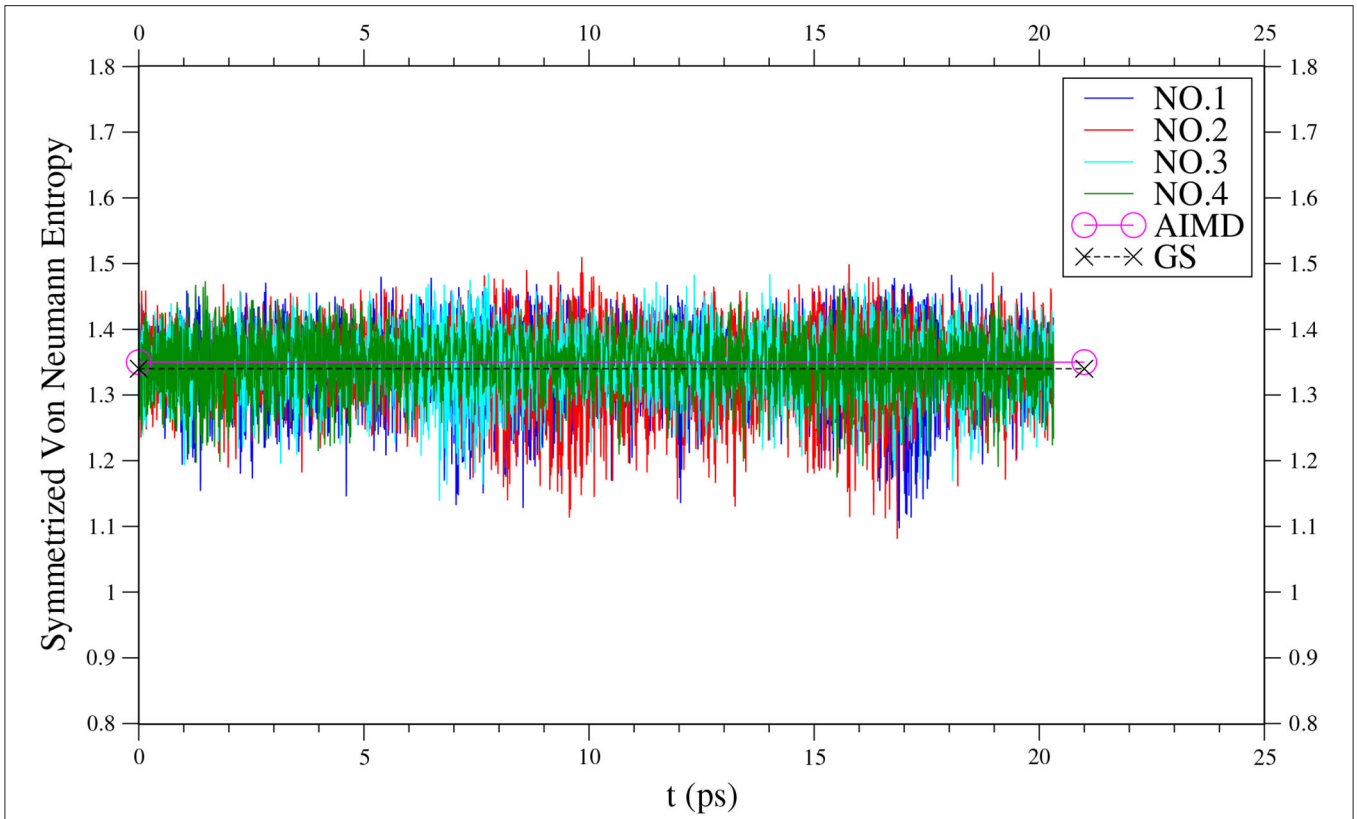

FIGURE 3 | Time evolution of the symmetrized von Neumann entropy of 8-acene, obtained from four different equilibrated TAO-AIMD trajectories (No.1 to No.4) at 300 K, calculated by TAO-LDA. The TAO-AIMD average and GS values are also shown for comparison.

can be responsible for the unidentified infrared emission (UIR) bands from interstellar media (Allamandola et al., 1985, 1999; Hudgins and Sandford, 1998a,b; Kim et al., 2001; Joblin et al., 2011).

On the basis of the Fermi golden rule of time-dependent perturbation theory (McQuarrie, 1976), the IR spectrum of a molecule is related to the Fourier transform of the autocorrelation function (ACF) (e.g., see Equation 15 of Brehm et al., 2020) of the dipole moment (i.e., the sum of the electronic and nuclear contributions) $\vec{\mu}$ of the molecule (Thomas et al., 2013, 2015; Dutta and Chowdhury, 2019):

$$I(\omega) \propto \omega^2 \int_{-\infty}^{\infty} \left\langle \vec{\mu}(0) \cdot \vec{\mu}(t) \right\rangle e^{-i\omega t} dt, \tag{12}$$

where $I(\omega)$ is the IR intensity, and ω is the vibrational frequency. Note that a quantum correction factor $\beta\hbar\omega/(1-e^{-\beta\hbar\omega})$ (Ramírez et al., 2004; Joalland et al., 2010; Thomas et al., 2013) has been taken into account in Equation (12), where $\beta = 1/(k_B T)$ and k_B is the Boltzmann constant. According to the properties of the Fourier transform (Thomas et al., 2013; Lawson Daku, 2018), $I(\omega)$ can also be expressed as

$$I(\omega) \propto \int_{-\infty}^{\infty} \left\langle \dot{\vec{\mu}}(0) \cdot \dot{\vec{\mu}}(t) \right\rangle e^{-i\omega t} dt, \tag{13}$$

which is directly proportional to the Fourier transform of the ACF of the time derivative of the dipole moment $\dot{\vec{\mu}}$ of the molecule.

Given the $\vec{\mu}(t)$ of n-acene along an equilibrated TAO-AIMD trajectory (obtained with Q-Chem 4.4, Shao, 2015), we compute the IR spectrum of n-acene using the TRAVIS program package (Brehm and Kirchner, 2011; Thomas et al., 2013, 2015; Brehm et al., 2020). To reduce the numerical noise, the IR spectrum has been smoothed using a window function applied in the time domain [specifically, each term of the ACF is multiplied by a Gaussian function $\exp(-\frac{\sigma t^2}{2\tau^2})$, where $\sigma = 10$ (as suggested in Gaigeot et al., 2005; Gaigeot, 2010; Vitale et al., 2015 for gas phase simulations) and τ is the total time duration of the equilibrated TAO-AIMD trajectory ($\approx$ 20.3 ps)]. The reported IR spectrum of n-acene is an average over four different equilibrated TAO-AIMD trajectories (with the average temperature being 300 $\pm$ 1 K).

On the other hand, NMA (i.e., normal mode analysis) is a commonly adopted approach to compute the vibrational frequencies and intensities of GS (Harris and Bertolucci, 1978; Wilson et al., 1980; Gaigeot et al., 2007) and excited-state (ES) (Liu and Liang, 2011a,b) molecules in the harmonic approximation at 0 K. For comparison purposes, we also compute the IR spectra of GS n-acenes using NMA. To perform a GS-NMA, the computation of nuclear second derivatives of energy (i.e., the nuclear Hessian) at the GS molecular geometry

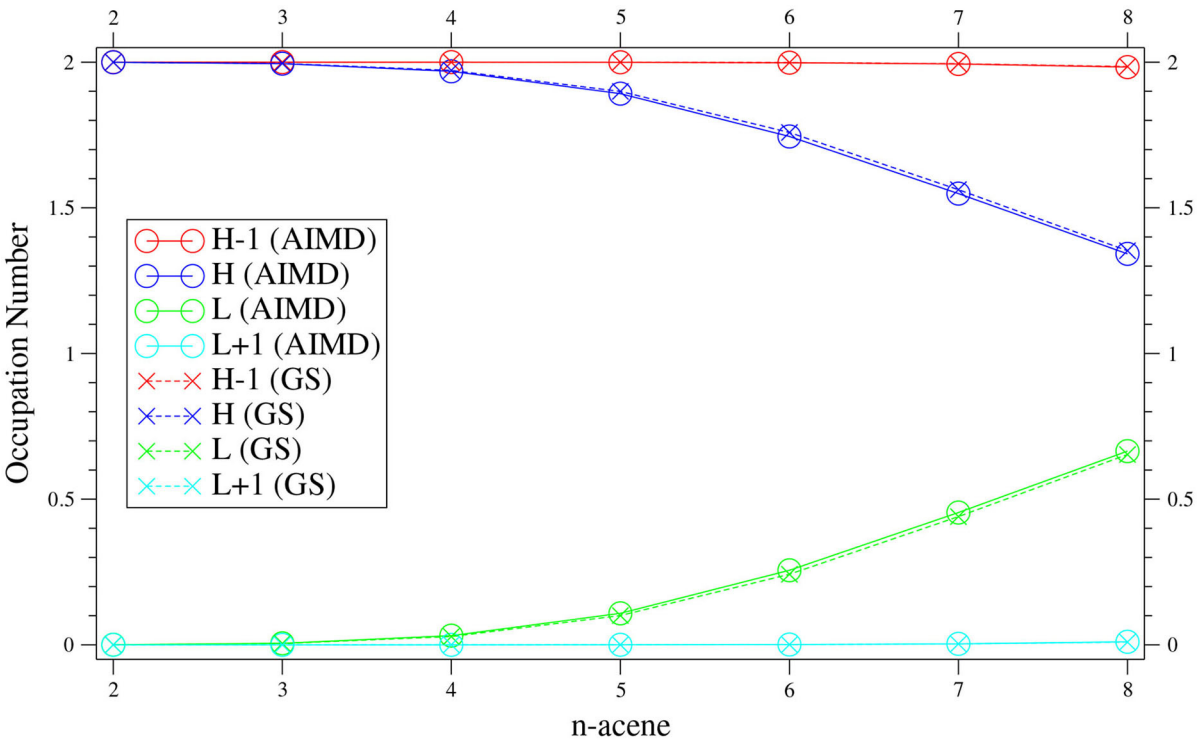

FIGURE 4 | Active orbital occupation numbers (HOMO−1, HOMO, LUMO, and LUMO+1) of *n*-acene, obtained with the TAO-AIMD simulations at 300 K and GS calculation, calculated by TAO-LDA.

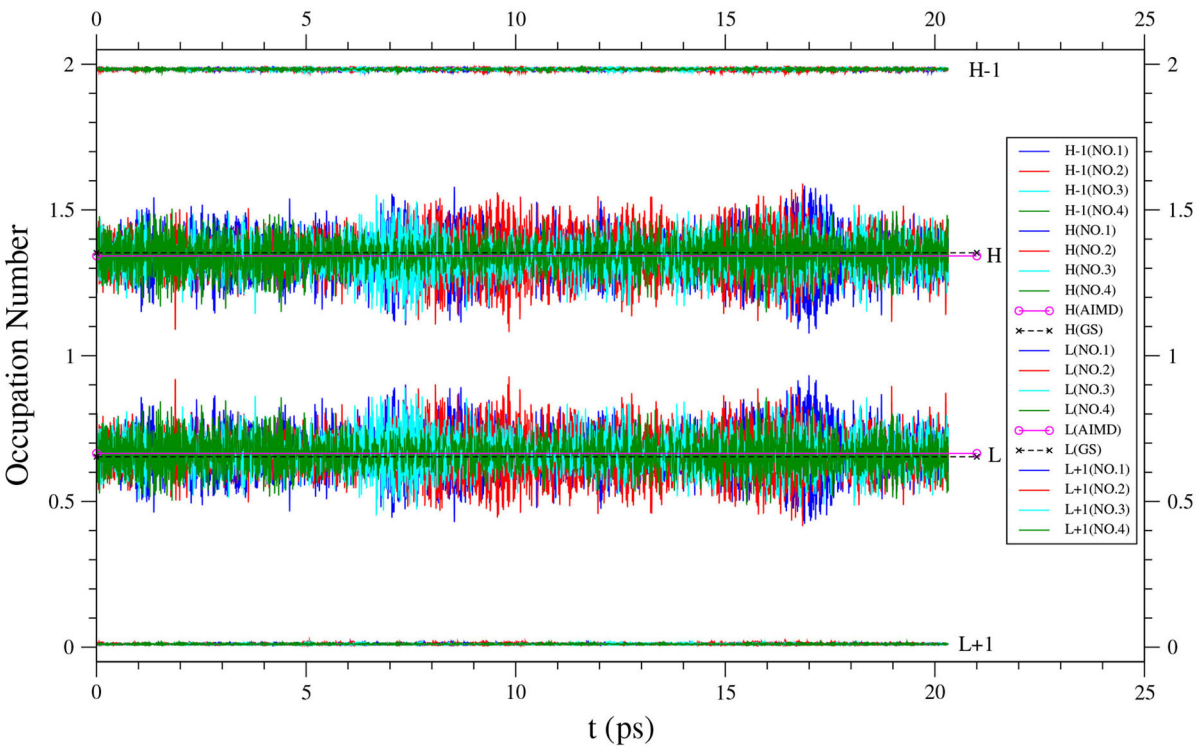

FIGURE 5 | Time evolution of the active orbital occupation numbers (HOMO−1, HOMO, LUMO, and LUMO+1) of 8-acene, obtained from four different equilibrated TAO-AIMD trajectories (No.1 to No.4) at 300 K, calculated by TAO-LDA. For HOMO/LUMO, the TAO-AIMD average and GS values are also shown for comparison.

is required. Since analytical nuclear Hessians for TAO-DFT are not yet available in Q-Chem, numerical nuclear Hessians are computed using finite differences of analytical nuclear gradients (with a step size of 0.001 Å, i.e., the default setting of Q-Chem) for all the GS-NMA performed in this work.

As shown in **Figures 6–9** (for comparison purposes, also see **Supplementary Figures 16–19** for the TAO-AIMD results obtained with different values of σ), the IR spectra of the smaller n-acenes (n = 2–5), obtained with the TAO-AIMD simulations at 300 K and GS-NMA, are in qualitative agreement with the available experimental IR spectra (Hudgins and Sandford, 1998a,b; Boersma et al., 2014; Bauschlicher et al., 2018; NIST mass spectrometry data center, 2020; Mattioda et al., unpublished). Note that a number of experimental IR bands (e.g., those in the 1,700–2,600 cm^{-1} range for 2-acene and 3-acene and those in the 1,700–2,000 cm^{-1} range for 4-acene and 5-acene) are completely missing from the corresponding IR spectra obtained with the commonly used GS-NMA approach, indicating that anharmonic effects on the IR spectra cannot be completely ignored. By contrast, these IR bands can be obtained with the TAO-AIMD simulations at 300 K, going beyond the harmonic approximation of GS-NMA. The overall discrepancies between theoretical and experimental results for the IR spectra can be attributed to several factors, such as the purity of samples and environmental factors (temperature, background noise level, etc.) in the experiments and the approximate nature of the theoretical methods adopted.

To our knowledge, there are no experimental IR spectra for 6-acene, 7-acene, and 8-acene. It remains highly challenging to synthesize and isolate the larger n-acenes (n = 6–8), possibly due to their radical character (Hachmann et al., 2007; Chai, 2012, 2014, 2017; Rivero et al., 2013; Wu and Chai, 2015; Fosso-Tande et al., 2016). In **Figures 10–12** (for comparison purposes, also see **Supplementary Figures 20–22** for the TAO-AIMD results obtained with different values of σ), we thus only compare the IR spectra of the larger n-acenes (n = 6–8), obtained with the TAO-AIMD simulations at 300 K and GS-NMA. As shown, both theoretical results are in reasonably good agreement, except for the 1,700–2,000 cm^{-1} range, which may be attributed to anharmonic effects on the IR spectra [cf., the IR spectra of the smaller n-acenes (n = 2–5)].

5. CONCLUSIONS

In conclusion, we have proposed TAO-AIMD (i.e., TAO-DFT-based AIMD) for the study of the equilibrium thermodynamic and dynamical properties of nanosystems with a radical nature at finite temperatures. To highlight some of the present capabilities of TAO-AIMD, we have performed TAO-AIMD simulations to investigate the instantaneous/average radical nature and IR spectra of n-acenes (n = 2–8) at 300 K. According to the TAO-AIMD simulations, on average, the smaller n-acenes (up to n = 5) possess a nonradical nature, and the larger n-acenes (n = 6–8) possess increasing radical nature, showing remarkable similarities to the GS counterparts at 0 K. Besides, the IR spectra of n-acenes obtained with the TAO-AIMD simulations are in qualitative agreement with the existing experimental data.

For GS molecules with radical nature [e.g., the larger n-acenes (n = 6–8)], on average, the radical nature of the molecules can persist in AIMD simulations at finite temperatures. For these molecules, conventional KS-AIMD simulations can therefore be unreliable, and AIMD simulations employing MR electronic structure methods are very likely to be computationally intractable for most molecules. It is thus certainly justified to perform TAO-AIMD simulations for exploring the dynamical information of these molecules.

While only TAO-AIMD (or more specifically, TAO-BOMD) is presented and discussed in this work, it is also possible to combine TAO-DFT with the popular Car-Parrinello MD (CPMD) (Car and Parrinello, 1985) method (i.e., an approximation of the BOMD method) for improved computational efficiency. In addition, it is worth mentioning that a brilliant simulation method combining the advantages of AIMD (for accuracy) and classical MD (for efficiency) has been developed, i.e., the hybrid QM/MM (quantum mechanics/molecular mechanics) method (Warshel and Levitt, 1976), where a small portion (e.g., the reactive portion) of a system is treated with QM and the remaining portion is treated with MM. The QM/MM method has been widely used for the study of very large systems where AIMD simulations are prohibitively expensive (van der Kamp and Mulholland, 2013). To further improve the efficiency of TAO-AIMD, the TAO-DFT-based QM/MM method is thus expected to be useful for the simulations of very large systems (e.g., biomolecules).

While TAO-AIMD is computationally efficient, it could be a very promising approach for the study of the equilibrium thermodynamic and dynamical properties of nanosystems with a radical or non-radical nature at finite temperatures. Nevertheless, a few limitations remain due to a number of assumptions made in TAO-AIMD. For example, the BO approximation is assumed to be valid, and the motion of the nuclei in a system is assumed to evolve only on the GS potential energy surface, wherein non-adiabatic effects are completely ignored. Besides, it is assumed that nuclear quantum effects can be neglected: the nuclear motion is assumed to be governed by the classical equations of motion, rather than the time-dependent nuclear Schrödinger equation. Moreover, it is assumed that TAO-DFT employing the approximate XC and θ-dependent energy functionals (with a system-independent fictitious temperature θ) can provide the exact GS potential energy surface. In addition, similar to KS-AIMD, the real electronic temperature is zero in TAO-AIMD, and it is thus assumed that TAO-AIMD remains applicable for systems at nonzero electronic temperatures. For a system at room temperature, TAO-AIMD should remain applicable (Gaigeot, 2010; Ramírez-Solís et al., 2011; Vitale et al., 2015). However, for a system at a considerably high electronic temperature (e.g., warm dense matter), TAO-AIMD may no longer be applicable (Rüter and Redmer, 2014; Karasiev et al., 2018; Bonitz et al., 2020), wherein AIMD

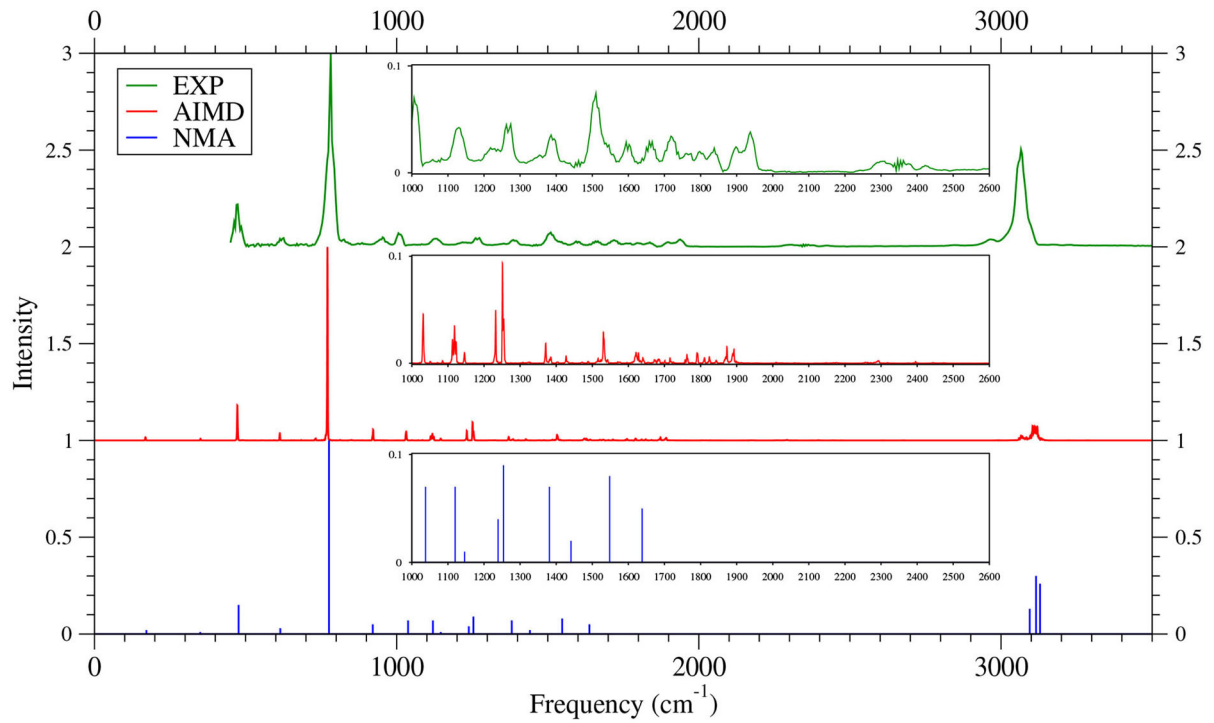

FIGURE 6 | IR spectra of 2-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. Experimental (EXP) data (NIST mass spectrometry data center, 2020) are included for comparison. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 1,000–2,600 cm^{-1} range.

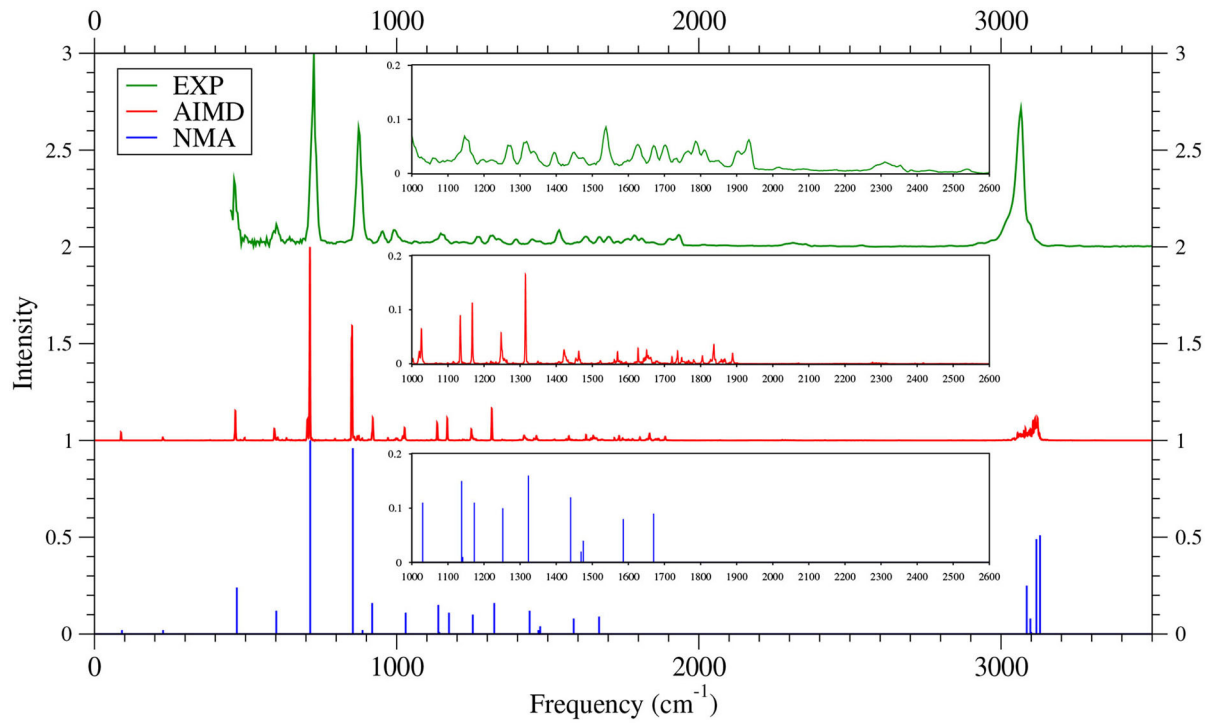

FIGURE 7 | IR spectra of 3-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. Experimental (EXP) data NIST mass spectrometry data center (2020) are included for comparison. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 1,000–2,600 cm^{-1} range.

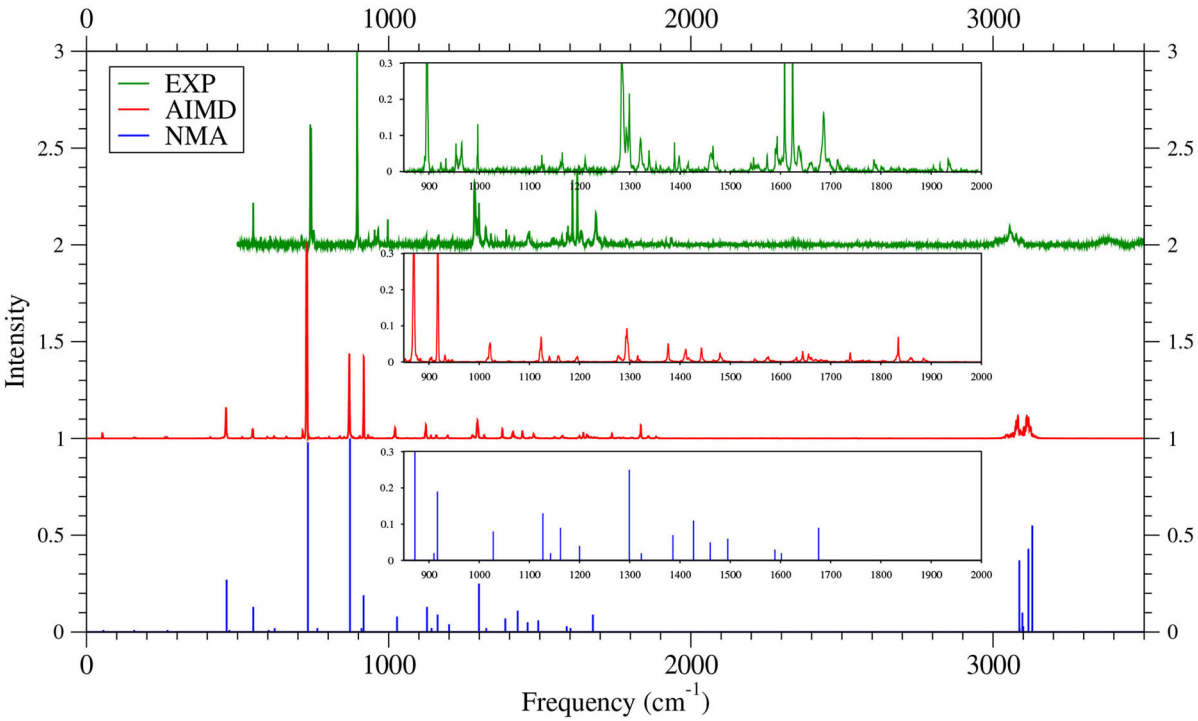

FIGURE 8 | IR spectra of 4-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. Experimental (EXP) data (Hudgins and Sandford, 1998a; Boersma et al., 2014; Bauschlicher et al., 2018; Mattioda et al., unpublished) are included for comparison. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 850–2,000 cm^{-1} range.

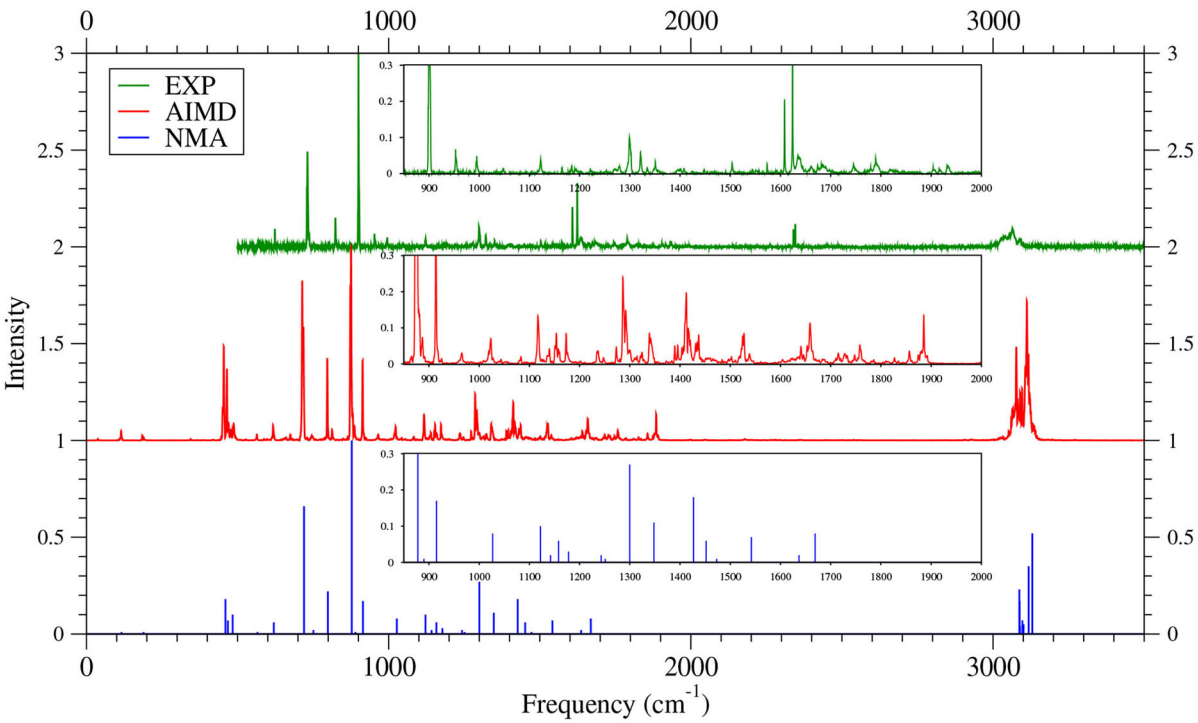

FIGURE 9 | IR spectra of 5-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. Experimental (EXP) data (Hudgins and Sandford, 1998b; Boersma et al., 2014; Bauschlicher et al., 2018; Mattioda et al., unpublished) are included for comparison. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 850–2,000 cm^{-1} range.

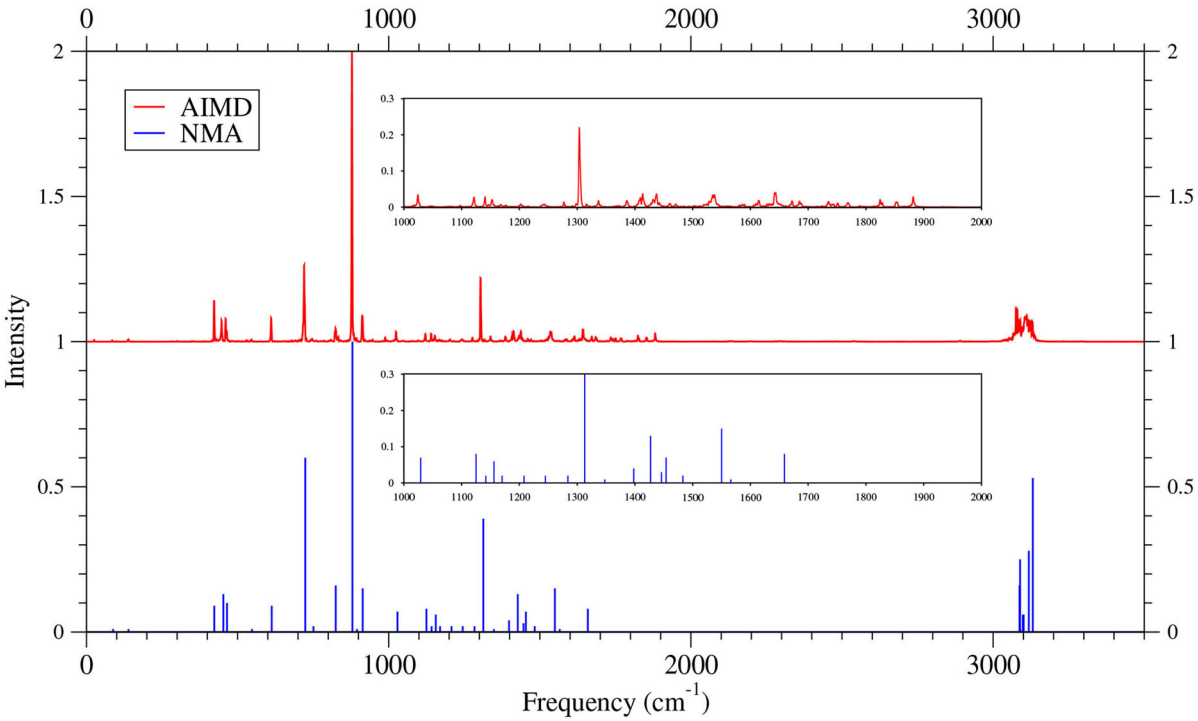

FIGURE 10 | IR spectra of 6-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 1,000–2,000 cm^{-1} range.

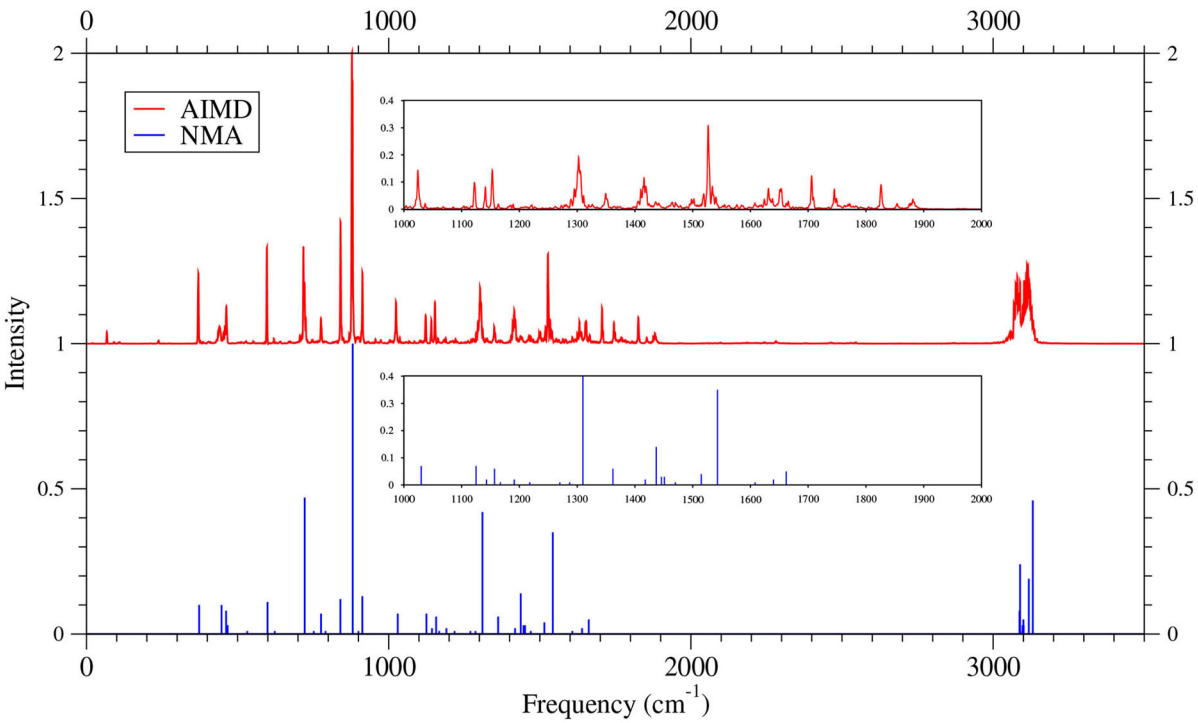

FIGURE 11 | IR spectra of 7-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 1,000–2,000 cm^{-1} range.

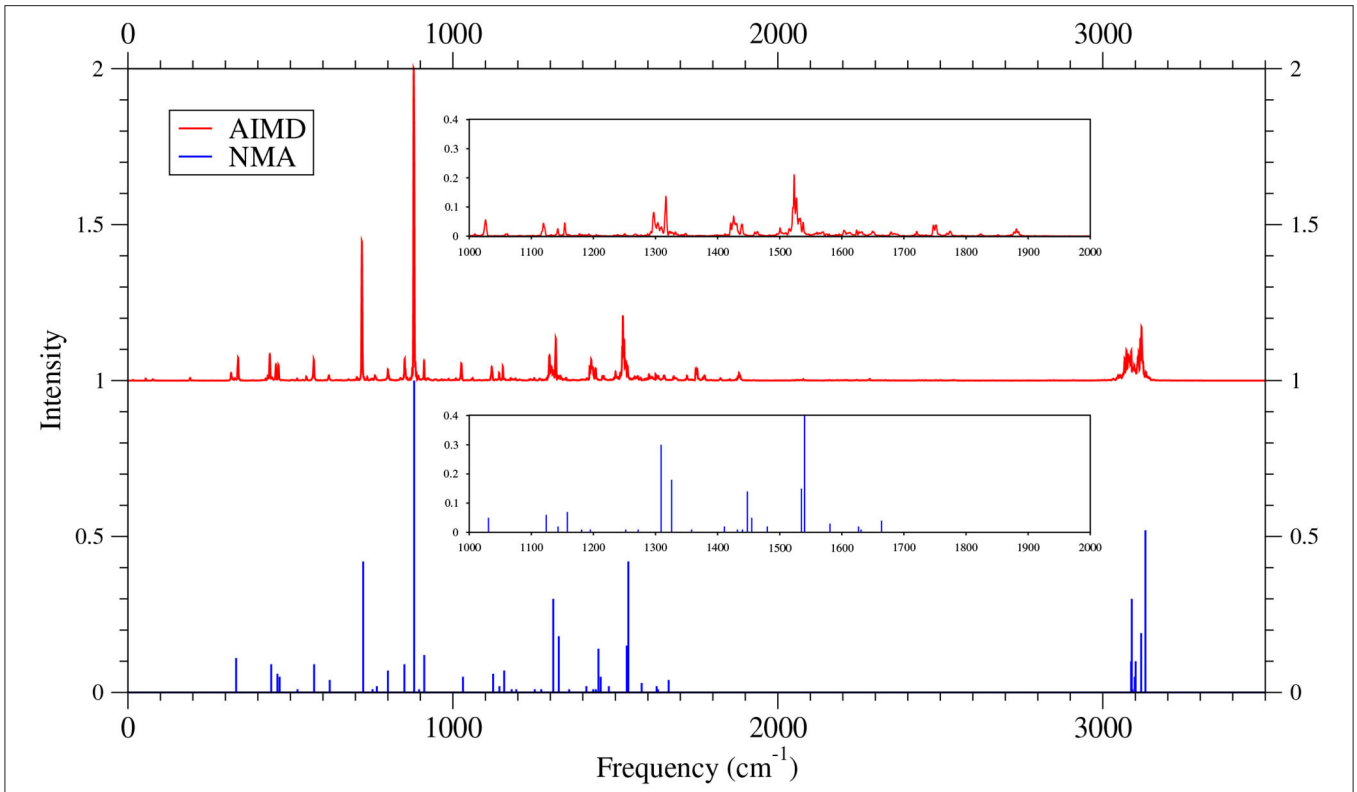

FIGURE 12 | IR spectra of 8-acene, obtained with the TAO-AIMD simulations at 300 K and GS-NMA, calculated by TAO-LDA. The IR spectra are normalized to have a maximum intensity of one and, for clarity, are vertically offset from each other by the same value. Subfigures show the IR spectra in the 1,000–2,000 cm^{-1} range.

based on the finite-electronic-temperature extension of TAO-DFT will be needed. To lift these limitations, we plan to investigate along some of these lines, and results may be reported elsewhere.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

J-DC conceived the project. SL performed the calculations. SL and J-DC designed the project, performed the data analysis, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Ministry of Science and Technology of Taiwan (Grant No. MOST107-2628-M-002-005-MY3), National Taiwan University (Grant No. NTU-CDP-105R7818), and the NCTS of Taiwan.

ACKNOWLEDGMENTS

We thank Dr. Martin Brehm for useful discussions on the evaluation of the IR spectra.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.589432/full#supplementary-material

REFERENCES

Allamandola, L. J., Hudgins, D. M., and Sandford, S. A. (1999). Modeling the unidentified infrared emission with combinations of polycyclic aromatic hydrocarbons. *Astrophys. J* 511, L115–L119. doi: 10.1086/311843

Allamandola, L. J., Tielens, A. G. G. M., and Barker, J. R. (1985). Polycyclic aromatic hydrocarbons and the unidentified infrared emission bands: auto exhaust along the milky way. *Astrophys. J.* 290, L25–L28. doi: 10.1086/184435

Andersson, K., Malmqvist, P.-Å., and Roos, B. O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* 96, 1218–1226. doi: 10.1063/1.462209

Battaglia, S., Faginas-Lago, N., Andrae, D., Evangelisti, S., and Leininger, T. (2017). Increasing radical character of large [*n*]cyclacenes unveiled by wave function theory. *J. Phys. Chem. A* 121, 3746–3756. doi: 10.1021/acs.jpca.7b00123

Bauschlicher, C. W. Jr., Ricca, A., Boersma, C., and Allamandola, L. J. (2018). The NASA Ames PAH IR spectroscopic database: computational version 3.00 with

updated content and the introduction of multiple scaling factors. *ApJS* 234:32. doi: 10.3847/1538-4365/aaa019

Beć, K. B., and Huck, C. W. (2019). Breakthrough potential in near-infrared spectroscopy: spectra simulation. A review of recent developments. *Front. Chem.* 7:48. doi: 10.3389/fchem.2019.00048

Boersma, C., Bauschlicher, C. W. Jr., Ricca, A., Mattioda, A. L., Cami, J., Peeters, E., et al. (2014). The NASA Ames PAH IR spectroscopic database version 2.00: updated content, web site, and on(off)line tools. *ApJS* 211:8. doi: 10.1088/0067-0049/211/1/8

Bonitz, M., Dornheim, T., Moldabekov, Zh. A., Zhang, S., Hamann, P., et al. (2020). *Ab initio* simulation of warm dense matter. *Phys. Plasmas* 27:042710. doi: 10.1063/1.5143225

Born, M., and Oppenheimer, J. R. (1927). Zur Quantentheorie der Molekeln. *Ann. Phys.* 84, 457–484. doi: 10.1002/andp.19273892002

Brehm, M., and Kirchner, B. (2011). TRAVIS – a free analyzer and visualizer for monte carlo and molecular dynamics trajectories. *J. Chem. Inf. Model.* 51, 2007–2023. doi: 10.1021/ci200217w

Brehm, M., Thomas, M., Gehrke, S., and Kirchner, B. (2020). TRAVIS – a free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* 152:164105. doi: 10.1063/5.0005078

Car, R., and Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55, 2471–2474. doi: 10.1103/PhysRevLett.55.2471

Chai, J.-D. (2012). Density functional theory with fractional orbital occupations. *J. Chem. Phys.* 136:154104. doi: 10.1063/1.3703894

Chai, J.-D. (2014). Thermally-assisted-occupation density functional theory with generalized-gradient approximations. *J. Chem. Phys.* 140:18A521. doi: 10.1063/1.4867532

Chai, J.-D. (2017). Role of exact exchange in thermally-assisted-occupation density functional theory: a proposal of new hybrid schemes. *J. Chem. Phys.* 146:044102. doi: 10.1063/1.4974163

Chai, J.-D., Stroud, D., Hafner, J., and Kresse, G. (2003). Dynamic structure factor of liquid and amorphous Ge from *ab initio* simulations. *Phys. Rev. B* 67:104205. doi: 10.1103/PhysRevB.67.104205

Chung, J.-H., and Chai, J.-D. (2019). Electronic properties of Möbius cyclacenes studied by thermally-assisted-occupation density functional theory. *Sci. Rep.* 9:2907. doi: 10.1038/s41598-019-39524-4

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2008). Insights into current limitations of density functional theory. *Science* 321, 792–794. doi: 10.1126/science.1158722

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2012). Challenges for density functional theory. *Chem. Rev.* 112, 289–320. doi: 10.1021/cr200107z

Deng, Q., and Chai, J.-D. (2019). Electronic properties of triangle-shaped graphene nanoflakes from TAO-DFT. *ACS Omega* 4, 14202–14210. doi: 10.1021/acsomega.9b01259

Dutta, B., and Chowdhury, J. (2019). Existence of dimeric hydroxylamine-O-sulfonic acid: experimental observations aided by ab initio, DFT, Car-Parrinello and Born-Oppenheimer on the fly dynamics. *Chem. Phys. Lett.* 732:136645. doi: 10.1016/j.cplett.2019.136645

Fosso-Tande, J., Nguyen, T.-S., Gidofalvi, G., and DePrince III, A. E. (2016). Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods. *J. Chem. Theory Comput.* 12, 2260–2271. doi: 10.1021/acs.jctc.6b00190

Gaigeot, M.-P. (2010). Theoretical spectroscopy of floppy peptides at room temperature. A DFTMD perspective: gas and aqueous phase. *Phys. Chem. Chem. Phys.* 12, 3336–3359. doi: 10.1039/B924048A

Gaigeot, M.-P., Martinez, M., and Vuilleumier, R. (2007). Infrared spectroscopy in the gas and liquid phase from first principle molecular dynamics simulations: application to small peptides. *Mol. Phys.* 105, 2857–2878. doi: 10.1080/00268970701724974

Gaigeot, M.-P., and Sprik, M. (2003). Ab initio molecular dynamics computation of the infrared spectrum of aqueous uracil. *J. Phys. Chem. B* 107, 10344–10358. doi: 10.1021/jp034788u

Gaigeot, M. P., Vuilleumier, R., Sprik, M., and Borgis, D. (2005). Infrared spectroscopy of *N*-Methylacetamide revisited by ab Initio molecular dynamics simulations. *J. Chem. Theory Comput.* 1, 772–789. doi: 10.1021/ct050029z

Gidofalvi, G., and Mazziotti, D. A. (2008). Active-space two-electron reduced-density-matrix method: complete active-space calculations without diagonalization of the *N*-electron hamiltonian. *J. Chem. Phys.* 129:134108. doi: 10.1063/1.2983652

Gryn'ova, G., Coote, M. L., and Corminboeuf, C. (2015). Theory and practice of uncommon molecular electronic configurations. *WIREs Comput. Mol. Sci.* 5, 440–459. doi: 10.1002/wcms.1233

Hachmann, J., Dorando, J. J., Aviles, M., and Chan, G. K. L. (2007). The radical character of the acenes: a density matrix renormalization group study. *J. Chem. Phys.* 127:134309. doi: 10.1063/1.2768362

Hanson-Heine, M. W. D. (2020). Static correlation in vibrational frequencies studied using thermally-assisted-occupation density functional theory. *Chem. Phys. Lett.* 739:137012. doi: 10.1016/j.cplett.2019.137012

Hanson-Heine, M. W. D., and Hirst, J. D. (2020). Möbius and Hückel cyclacenes with Dewar and Ladenburg defects. *J. Phys. Chem. A* 124, 5408–5414. doi: 10.1021/acs.jpca.0c04137

Hanson-Heine, M. W. D., Rogers, D. M., Woodward, S., and Hirst, J. D. (2020). Dewar benzenoids discovered in carbon nanobelts. *J. Phys. Chem. Lett.* 11, 3769–3772. doi: 10.1021/acs.jpclett.0c01027

Harris, D. C., Bertolucci, M. D. (1978). *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy*. New York: Oxford University Press.

Hollingsworth, S. A., and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron* 99, 1129–1143. doi: 10.1016/j.neuron.2018.08.011

Huang, H.-J., Seenithurai, S., and Chai, J.-D. (2020). TAO-DFT study on the electronic properties of diamond-shaped graphene nanoflakes. *Nanomaterials* 10:1236. doi: 10.3390/nano10061236

Hudgins, D. M., and Sandford, S. A. (1998a). Infrared spectroscopy of matrix isolated polycyclic aromatic hydrocarbons. 1. PAHs containing two to four rings. *J. Phys. Chem. A* 102, 329–343. doi: 10.1021/jp9834816

Hudgins, D. M., and Sandford, S. A. (1998b). Infrared spectroscopy of matrix isolated polycyclic aromatic hydrocarbons. 2. PAHs containing five or more rings. *J. Phys. Chem. A* 102, 344–352. doi: 10.1021/jp983482y

Jensen, F. (2007). *Introduction to Computational Chemistry*. New York, NY: Wiley.

Joalland, B., Rapacioli, M., Simon, A., Joblin, C., Marsden, C. J., and Spiegelman, F. (2010). Molecular dynamics simulations of anharmonic infrared spectra of $[SiPAH]^+$ π-complexes. *J. Phys. Chem. A* 114, 5846–5854. doi: 10.1021/jp911526n

Joblin, C., Tielens, A., and Oomens, J. (2011). Laboratory infrared spectroscopy of PAHs. *EAS Publ. Ser.* 46, 61–73. doi: 10.1051/eas/1146007

Karasiev, V. V., Dufty, J. W., and Trickey, S. B. (2018). Nonempirical semilocal free-energy density functional for matter under extreme conditions. *Phys. Rev. Lett.* 120:076401. doi: 10.1103/PhysRevLett.120.076401

Karplus, M., and Petsko, G. A. (1990). Molecular dynamics simulations in biology. *Nature* 347, 631–639. doi: 10.1038/347631a0

Kim, H. S., Wagner, D. R., and Saykally, R. J. (2001). Single photon infrared emission spectroscopy of the gas phase pyrene cation: support for a polycyclic aromatic hydrocarbon origin of the unidentified infrared emission bands. *Phys. Rev. Lett.* 86, 5691–5694. doi: 10.1103/PhysRevLett.86.5691

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140, A1133–A1138. doi: 10.1103/PhysRev.140.A1133

Kresse, G., and Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* 47, 558–561. doi: 10.1103/PhysRevB.47.558

Kuo, I. F. W., and Mundy, C. J. (2004). An ab Initio molecular dynamics study of the aqueous liquid-vapor interface. *Science* 303, 658–660. doi: 10.1126/science.1092787

Lawson Daku, L. M. (2018). Spin-state dependence of the structural and vibrational properties of solvated iron(ii) polypyridyl complexes from AIMD simulations: aqueous [Fe(bpy)3]Cl2, a case study. *Phys. Chem. Chem. Phys.* 20, 6236–6253. doi: 10.1039/C7CP07862E

Levitt, M., and Lifson, S. (1969). Refinement of protein conformations using a macromolecular energy minimization procedure. *J. Mol. Biol.* 46, 269–279. doi: 10.1016/0022-2836(69)90421-5

Lifson, S., and Warshel, A. (1968). Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. *J. Chem. Phys.* 49, 5116–5129. doi: 10.1063/1.1670007

Lin, C.-Y., Hui, K., Chung, J.-H., and Chai, J.-D. (2017). Self-consistent determination of the fictitious temperature in thermally-assisted-occupation density functional theory. *RSC Adv.* 7, 50496–50507. doi: 10.1039/C7RA10241K

Liu, J., and Liang, W. Z. (2011a). Analytical Hessian of electronic excited states in time-dependent density functional theory with Tamm-Dancoff approximation. *J. Chem. Phys.* 135:014113. doi: 10.1063/1.3605504

Liu, J., and Liang, W. Z. (2011b). Analytical approach for the excited-state Hessian in time-dependent density functional theory: formalism, implementation, and performance. *J. Chem. Phys.* 135:184111. doi: 10.1063/1.3659312

Löwdin, P.-O., and Shull, H. (1956). Natural orbitals in the quantum theory of two-electron systems. *Phys. Rev.* 101, 1730–1739. doi: 10.1103/PhysRev.101.1730

Manassir, M., and Pakiari, A. H. (2020). Total non-Lewis structures: an application to predict the stability and reactivity of linear and angular polyacenes. *J. Mol. Graph. Model.* 99:107643. doi: 10.1016/j.jmgm.2020.107643

Martyna, G. J., Klein, M. L., and Tuckerman, M. (1992). Nosé-Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97, 2635–2643. doi: 10.1063/1.463940

Marx, D., and Hutter, J. (2009). *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge: Cambridge University Press.

McQuarrie, D. A. (1976). *Statistical Mechanics*. New York, NY: Harper and Row Publishers.

Mullinax, J. W., Maradzike, E., Koulias, L. N., Mostafanejad, M., Epifanovsky, E., Gidofalvi, G., and DePrince III, A. E. (2019). Heterogeneous CPU + GPU algorithm for variational two-electron reduced-density matrix-driven complete active-space self-consistent field theory. *J. Chem. Theory Comput.* 15, 6164–6178. doi: 10.1021/acs.jctc.9b00768

NIST mass spectrometry data center, Wallace, W. E., director, "Infrared Spectra" in NIST chemistry webbook, NIST standard reference database number 69, eds. Linstrom, P. J., and Mallard, W. G., National Institute of Standards and Technology, Gaithersburg MD, 20899, https://doi.org/10.18434/T4D303, (retrieved June 17, 2020).

Pelzer, K., Greenman, L., Gidofalvi, G., and Mazziotti, D. A. (2011). Strong correlation in acene sheets from the active-space variational two-electron reduced density matrix method: effects of symmetry and size. *J. Phys. Chem. A* 115, 5632–5640. doi: 10.1021/jp2017192

Putrino, A., and Parrinello, M. (2002). Anharmonic Raman spectra in high-pressure ice from ab initio simulations. *Phys. Rev. Lett.* 88:176401. doi: 10.1103/PhysRevLett.88.176401

Ramírez, R., López-Ciudad, T., Kumar P, P., and Marx, D. (2004). Quantum corrections to classical time-correlation functions: hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.* 121, 3973–3983. doi: 10.1063/1.1774986

Ramírez-Solís, A., Jolibois, F., and Maron, L. (2011). Born-Oppenheimer DFT molecular dynamics studies of S_2O_2: non-harmonic effects on the lowest energy isomers. *Chem. Phys. Lett.* 510, 21–26. doi: 10.1016/j.cplett.2011.04.091

Rivero, P., Jiménez-Hoyos, C. A., and Scuseria, G. E. (2013). Entanglement and polyradical character of polycyclic aromatic hydrocarbons predicted by projected Hartree-Fock theory. *J. Phys. Chem. B* 117, 12750–12758. doi: 10.1021/jp401478v

Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., and Jent, N. (2007). A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J. Pharm. Biomed. Anal.* 44, 683–700. doi: 10.1016/j.jpba.2007.03.023

Rüter, H. R., and Redmer, R. (2014). *Ab initio* simulations for the ion-ion structure factor of warm dense aluminum. *Phys. Rev. Lett.* 112:145007. doi: 10.1103/PhysRevLett.112.145007

Seenithurai, S., and Chai, J.-D. (2016). Effect of Li adsorption on the electronic and hydrogen storage properties of acenes: a dispersion-corrected TAO-DFT study. *Sci. Rep.* 6:33081. doi: 10.1038/srep33081

Seenithurai, S., and Chai, J.-D. (2017). Effect of Li termination on the electronic and hydrogen storage properties of linear carbon chains: a TAO-DFT study. *Sci. Rep.* 7:4966. doi: 10.1038/s41598-017-05202-6

Seenithurai, S., and Chai, J.-D. (2018). Electronic and hydrogen storage properties of Li-terminated linear boron chains studied by TAO-DFT. *Sci. Rep.* 8:13538. doi: 10.1038/s41598-018-31947-9

Seenithurai, S., and Chai, J.-D. (2019). Electronic properties of linear and cyclic boron nanoribbons from thermally-assisted-occupation density functional theory. *Sci. Rep.* 9:12139. doi: 10.1038/s41598-019-48560-z

Seenithurai, S., and Chai, J.-D. (2020). TAO-DFT investigation of electronic properties of linear and cyclic carbon chains. *Sci. Rep.* 10:13133. doi: 10.1038/s41598-020-70023-z

Shao, Y., Gan, Z., Epifanovsky, E., Gilbert, T. B., Wormit, M., Kussmann, J., et al. (2015). Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* 113, 184–215. doi: 10.1080/00268976.2014.952696

Silvestrelli, P. L., Bernasconi, M., and Parrinello, M. (1997). *Ab initio* infrared spectrum of liquid water. *Chem. Phys. Lett.* 277, 478–482. doi: 10.1016/S0009-2614(97)00930-5

Sprik, M., Hutter, J., and Parrinello, M. (1996). Ab initio molecular dynamics simulation of liquid water: comparison of three gradient–corrected density functionals. *J. Chem. Phys.* 105, 1142–1152. doi: 10.1063/1.471957

Thomas, M., Brehm, M., Fligg, R., Vöhringer, P., and Kirchner, B. (2013). Computing vibrational spectra from *ab initio* molecular dynamics. *Phys. Chem. Chem. Phys.* 15, 6608–6622. doi: 10.1039/C3CP44302G

Thomas, M., Brehm, M., and Kirchner, B. (2015). Voronoi dipole moments for the simulation of bulk phase vibrational spectra. *Phys. Chem. Chem. Phys.* 17, 3207–3213. doi: 10.1039/C4CP05272B

Tuckerman, M. E. (2002). Ab initio molecular dynamics: basic concepts, current trends and novel applications. *J. Phys.* 14, R1297–R1355. doi: 10.1088/0953-8984/14/50/202

van der Kamp, M. W., and Mulholland, A. J. (2013). Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 52, 2708–2728. doi: 10.1021/bi400215w

Vitale, V., Dziedzic, J., Dubois, S. M.-M., Fangohr, H., and Skylaris, C.-K. (2015). Anharmonic infrared spectroscopy through the Fourier transform of time correlation function formalism in ONETEP. *J. Chem. Theory Comput.* 11, 3321–3332. doi: 10.1021/acs.jctc.5b00391

Warshel, A., and Levitt, M. (1976). Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103, 227–249. doi: 10.1016/0022-2836(76)90311-9

Wilson, E. B. Jr, Decius, J. C., and Cross, P. C. (1980). *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. New York, NY: Dover Publications, Inc..

Wu, C.-S., and Chai, J.-D. (2015). Electronic properties of zigzag graphene nanoribbons studied by TAO-DFT. *J. Chem. Theory Comput.* 11, 2003–2011. doi: 10.1021/ct500999m

Wu, C.-S., Lee, P.-Y., and Chai, J.-D. (2016). Electronic properties of cyclacenes from TAO-DFT. *Sci. Rep.* 6:37249. doi: 10.1038/srep37249

Xuan, F., Chai, J.-D., and Su, H. (2019). Local density approximation for the short-range exchange free energy functional. *ACS Omega* 4, 7675–7683. doi: 10.1021/acsomega.9b00303

Yeh, C.-N., and Chai, J.-D. (2016). Role of Kekulé and non-Kekulé structures in the radical character of alternant polycyclic aromatic hydrocarbons: a TAO-DFT study. *Sci. Rep.* 6:30562. doi: 10.1038/srep30562

Yeh, C.-N., Wu, C., Su, H., and Chai, J.-D. (2018). Electronic properties of the coronene series from thermally-assisted-occupation density functional theory. *RSC Adv.* 8, 34350–34358. doi: 10.1039/C8RA01336E

Yeh, S.-H., Manjanath, A., Cheng, Y.-C., Chai, J.-D., and Hsu, C.-P. (2020). Excitation energies from thermally-assisted-occupation density functional theory: theory and computational implementation. *J. Chem. Phys.* 153:084120. doi: 10.1063/1.5140243

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 10 November 2020
doi: 10.3389/fchem.2020.589992

Implementation of Laplace Transformed MP2 for Periodic Systems With Numerical Atomic Orbitals

Honghui Shang[1]* *and Jinlong Yang*[2]*

[1] State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, [2] Hefei National Laboratory for Physical Sciences at Microscale, Department of Chemical Physics, Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, China

We present an implementation of the canonical and Laplace-transformed formulation of the second-order Møller–Plesset perturbation theory under periodic boundary conditions using numerical atomic orbitals. To validate our approach, we show that our results of the Laplace-transformed MP2 correlation correction for the total energy and the band gap are in excellent agreement with the results of the canonical MP2 formulation. We have calculated the binding energy curve for the stacked trans-polyacetylene at the Hartree–Fock + MP2 level as a preliminary application.

Keywords: MP2, NAO, real-space, Hartree–Fock, periodic system

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Yi Yao,
Duke University, United States
Wei Li,
Nanjing University, China

***Correspondence:**
Honghui Shang
shanghui.ustc@gmail.com
Jinlong Yang
jlyang@ustc.edu.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *31 July 2020*
Accepted: *15 September 2020*
Published: *10 November 2020*

Citation:
Shang H and Yang J (2020) Implementation of Laplace Transformed MP2 for Periodic Systems With Numerical Atomic Orbitals. Front. Chem. 8:589992. doi: 10.3389/fchem.2020.589992

1. INTRODUCTION

The second-order Møller–Plesset perturbation theory (MP2) is a post-Hartree–Fock approach to take the electron correlation effect into account. Although it is very simple in form, it can capture around 90% of the correlation energy (Bartlett and Stanton, 2007); so the MP2 method is still of high interest in the quantum chemistry (Schütz et al., 1999; Kobayashi and Nakai, 2006; Bartlett and Stanton, 2007) and solid-state physics communities (Suhai, 1983, 1992; Sun and Bartlett, 1996; Pisani et al., 2008; Marsman et al., 2009; Schäfer et al., 2018).

However, the $O(N^5)$ calculation scaling of the original (canonical) MP2 method has limited the application of the MP2 method in large systems. A series of algorithms have been proposed to speed up the calculations, such as local MP2 method (Saebø and Pulay, 1993; Pisani et al., 2005, 2008; Maschio, 2011), Lapace-transformed MP2 method (Häser and Almlöf, 1992; Häser, 1993; Ayala and Scuseria, 1999; Ayala et al., 2001; Schäfer et al., 2018), or resolution of the identity (RI) MP2 method (Katouda and Nagase, 2010; Ren et al., 2012). The local MP2 method proposed by Pulay (1983) and Saebø and Pulay (1993) has been efficiently implemented (Schütz et al., 1999) in the MOLPRO code for molecules, then the periodic version of the local MP2 method has been implemented (Pisani et al., 2005, 2008; Maschio, 2011) in the CRYSCOR code and in the CP2K code (Usvyat et al., 2018) for extended systems. Since the spatially localized orbitals or Wannier functions are adopted, the computational scaling of the local MP2 method is $O(N)$. The Laplace-transformed MP2 method is originally proposed by Häser and Almlöf (1992) and Häser (1993), and have been implemented for both the molecule (Ayala and Scuseria, 1999) and extended systems (Ayala et al., 2001) in the GAUSSIAN suite of programs. The localized atomic orbitals have been employed and the computational scaling is also O(N). The Laplace-transformed MP2 method has been combined with the resolution of identity (RI) technique to further improve

the computational efficiency (Izmaylov and Scuseria, 2008). Further rigorous integral screening scheme (Lambrecht et al., 2005) has been introduced on top of the Laplace-transformed MP2 to perform the calculations for a system comprising 1,000 atoms (Doser et al., 2008). Recently, the Laplace-transformed MP2 method has also been implemented (Schäfer et al., 2018) in VASP using stochastic orbitals.

So far, most of the implementations of the MP2 are adopting the Gaussian-type orbital (GTO) as the basis set. However, in the calculation of the periodic system, too diffused GTO with a long tail will increase the number of cells in the auxiliary supercell, and therefore the computational cost will increase. Compared with GTO, the numerical atomic orbital (NAO) is strictly localized, which could naturally leads to lower order scaling of computational time vs. system size. Here in this work, we have implemented the canonical MP2 and Laplace-transformed MP2 for the extended systems using NAO, and the results obtained by these two approaches are consistent. Furthermore, we have investigated the MP2 correlation correction to the band structure with both the canonical and Laplace-transformed formulation; our implementation has been validated by comparing the MP2 correlation correction of the total energy and the band gap to the literature values.

The remainder of this paper is organized as follows. The fundamental theoretical framework and the implementation details for the canonical and Laplace transformed MP2 are presented in section 2. The benchmark calculations are presented in section 3. In section 4, we summarize our main achievement and highlight the possible future research direction related to this work.

2. METHOD

2.1. Numerical Atomic Orbitals

The numerical atomic orbital is defined by a product of a numerical radial function and a spherical harmonic

$$\chi_{Ilmn}(\mathbf{r}) = \varphi_{Iln}(r)\mathbf{Y}_{lm}(\hat{r})\ . \tag{1}$$

By solving the one-dimension radial Schrödinger equation

$$(-\frac{1}{2}\frac{\mathbf{1}}{\mathbf{r}}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{l(l+1)}{2r^2} + V(r) + V_{cut})\varphi_{Iln}(r) = \epsilon_l\varphi_{Iln}(r)\ , \tag{2}$$

we can get the radial part of the numerical atomic orbital $\varphi_{Iln}(r)$, where $V(r)$ denotes the electrostatic potential for orbital $\varphi_{Iln}(r)$, and V_{cut} ensures a smooth decay of each radial function, which is strictly zero outside a confining radius r_{cut}.

In order to perform the Hartree–Fock and MP2 calculation, the electron repulsion integrals (ERIs) are needed:

$$(\chi_\mu\chi_\nu|\chi_\lambda\chi_\sigma) = \int\int \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\chi_\lambda(\mathbf{r}')\chi_\sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathbf{dr}\mathbf{dr}' \tag{3}$$

we use NAO2GTO scheme described to calculate them as shown in the following section.

2.2. The NAO2GTO Scheme to Calculate ERIs

In the NAO2GTO scheme, we fit the NAO with GTOs, then we calculate the ERIs analytically; in this way, the strict cutoff of the atomic orbitals is satisfied with NAO and the construction of Hartree-Fock exchange (HFX) matrix can scale linearly with the system sizes (Shang et al., 2011). Since the angular part of the NAOs is spherical harmonic, while the GTOs are Cartesian harmonic function, a transformation between the Cartesian and spherical harmonic functions is performed within the NAO2GTO scheme.

2.3. Canonical MP2 Formulation

In extended systems, the normalized crystal orbital $\psi_i(\mathbf{k},\mathbf{r})$ is a linear combination of Bloch functions $\phi_\mu(\mathbf{k},\mathbf{r})$:

$$\psi_i(\mathbf{k},\mathbf{r}) = \sum_\mu C_{\mu,i}(\mathbf{k})\phi_\mu(\mathbf{k},\mathbf{r}) \tag{4}$$

$$\phi_\mu(\mathbf{k},\mathbf{r}) = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}}\chi_\mu^{\mathbf{R}}(\mathbf{r})\mathbf{e}^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{r}_\mu)} \tag{5}$$

in which N is the number of cells in extended systems, μ is the index of the atomic orbitals, i refers to the crystal orbital index, $\mathbf{R}$ denotes the cells in the extended systems (auxiliary supercell), $\chi_\mu^{\mathbf{R}}(\mathbf{r}) = \chi_\mu(\mathbf{r}-\mathbf{R}-\mathbf{r}_\mu)$ refers to the atomic orbital whose center is displaced from the cell $\mathbf{R}$ by $\mathbf{r}_\mu$, and $C_{\mu,i}(\mathbf{k})$ are the coefficients of the crystal orbitals.

The MP2 correlation correction for the total energy of the unit cell is

$$E_{\mathrm{mp2}} = -\frac{1}{N}\sum_i\sum_j\sum_a\sum_b\frac{1}{V_k^4}\int\mathbf{dk_i}\int\mathbf{dk_j}\int\mathbf{dk_a}\int\mathbf{dk_b}\frac{(IA|JB)[2(IA|JB)-(IB|JA)]^*}{\epsilon_a(\mathbf{k_a})+\epsilon_b(\mathbf{k_b})-\epsilon_i(\mathbf{k_i})-\epsilon_j(\mathbf{k_j})} \tag{6}$$

in which we use labeling i,j for occupied orbitals and a,b for unoccupied orbitals. I/J refer to the composite index (i,$\mathbf{k}_i$)/(j,$\mathbf{k}_j$), V_k is the volume of the Brillouin zone, and $\epsilon_i(\mathbf{k_i})$ is the Hartree–Fock eigenvalue for the eigenstate $\psi_i(\mathbf{k_i})$. It should be noted that by using the identity ($\sum_{\mathbf{R}}\exp i\mathbf{k}\cdot\mathbf{R} = N\delta_{\mathbf{k},0}$) derived with the Born–von Karman periodic boundary condition, we can remove one dimension integration over $\mathbf{k}_j$ since $\mathbf{k}_j = \mathbf{T}(-\mathbf{k}_i+\mathbf{k}_a+\mathbf{k}_b)$, where $\mathbf{T}$ is the translation operator. The formalism of summation over 3-fold $\mathbf{k}$ points and over 4-fold $\mathbf{k}$ points (Equation 6) give the same results.

Similarly, the MP2 correlation correction $(\epsilon_g(\mathbf{k}_g)^{(2)})$ to the Hartree–Fock eigenstate $\psi_g(\mathbf{k}_g)$ can be written as

$$\epsilon_g(\mathbf{k}_g)^{(2)} = \epsilon_g(\mathbf{k}_g)^{\mathrm{MP2}} - \epsilon_g(\mathbf{k}_g)^{\mathrm{HF}} = U(g) + V(g) \tag{7}$$

$$U(g) = -\sum_{i,a,b}\frac{1}{V_k^3}\int\mathbf{dk_i}\int\mathbf{dk_a}\int\mathbf{dk_b}\frac{(IA|GB)[2(IA|GB)-(IB|GA)]^*}{\epsilon_a(\mathbf{k_a})+\epsilon_b(\mathbf{k_b})-\epsilon_i(\mathbf{k_i})-\epsilon_g(\mathbf{k}_g)} \tag{8}$$

$$V(g) = \sum_{i,j,a}\frac{1}{V_k^3}\int\mathbf{dk_i}\int\mathbf{dk_j}\int\mathbf{dk_a}\frac{(IA|JG)[2(IA|JG)-(IG|JA)]^*}{\epsilon_a(\mathbf{k_a})+\epsilon_g(\mathbf{k}_g)-\epsilon_i(\mathbf{k_i})-\epsilon_j(\mathbf{k_j})} \tag{9}$$

When $\psi_g(\mathbf{k}_g)$ is the occupied orbital at the valance band maximum (VBM), we can see $U(g) < 0, V(g) > 0$ and $|U(g)| < |V(g)|$, then the MP2 renormalization of the VBM is positive and will move the VBM orbital upward. When $\psi_g(\mathbf{k}_g)$ is the unoccupied orbital at the conduction band minimum (CBM), we have $|U(g)| > |V(g)|$, so the MP2 renormaliztion of the CBM is negative, and will move the CBM downward. In total, the MP2 renormalization of the band gap is negative, and the MP2 band gap is smaller than the Hartree-Fock band gap.

2.4. Laplace-Transformed MP2 Formulation

The Laplace transform is defined as:

$$\frac{1}{x} = \int_0^\infty e^{-xt}\mathbf{d}t, x > 0 \tag{10}$$

which can be used to remove the denominator in the canonical MP2 formulation:

$$\frac{1}{\epsilon_a(k_a) + \epsilon_b(k_b) - \epsilon_i(k_i) - \epsilon_j(k_j)} = \int e^{[\epsilon_i(k_i)+\epsilon_j(k_j)-\epsilon_a(k_a)-\epsilon_b(k_b)]t}\mathbf{d}t \tag{11}$$

The integration in Equation (10) can either be done by using a least square fitting method (Häser and Almlöf, 1992; Häser, 1993) or by using a Jacobian transform (Ayala and Scuseria, 1999; Kobayashi and Nakai, 2006) of the Laplace integration variable in order to transform the integration range $[0, \infty)$ into the finite range $[0, 1]$. Here, we use the transform as follows:

$$\int_0^\infty e^{-xt}\mathbf{d}t = \int_0^1 e^{-xt}\frac{\mathbf{d}t}{\mathbf{d}r}\mathbf{d}r = \int_0^1 f(r)\mathbf{d}r \tag{12}$$

in which the Jacobian transform is

$$t = \frac{r^3 - 0.9r^4}{(1-r)^2} + r^2\tan(\frac{\pi r}{2}) \tag{13}$$

Then the final integration ($\int_0^1 f(r)\mathbf{d}r$) in Equation (12) in evaluated with Romberg quadrature method, which uses refinements of the extended trapezoidal rule to reduce error in definite integrals.

In this way, the E_{mp2} correlation correction energy can be written as a new integration form:

$$E_{\mathrm{mp2}} = -\int \mathbf{d}t \sum_{\mu\mathbf{0},\nu\mathbf{R}_\nu,\lambda\mathbf{R}_\lambda,\sigma\mathbf{R}_\sigma} T_{\mu,\lambda,\nu,\sigma}^{\mathbf{0}\mathbf{R}_\lambda\mathbf{R}_\nu\mathbf{R}_\sigma}(t)[2(\chi_\mu^{\mathbf{0}}\chi_\lambda^{\mathbf{R}_\lambda}|\chi_\nu^{\mathbf{R}_\nu}\chi_\sigma^{\mathbf{R}_\sigma}) - (\chi_\mu^{\mathbf{0}}\chi_\sigma^{\mathbf{R}_\sigma}|\chi_\nu^{\mathbf{R}_\nu}\chi_\lambda^{\mathbf{R}_\lambda})] \tag{14}$$

where μ,ν,σ,λ and the following $\lambda,\delta,\tau,\kappa$ refer to the indexes of the atomic orbitals. $(\chi_\mu^{\mathbf{0}}\chi_\sigma^{\mathbf{R}_\sigma}|\chi_\nu^{\mathbf{R}_\nu}\chi_\lambda^{\mathbf{R}_\lambda})$ is the electron repulsion integrals defined as

$$(\chi_\mu^{\mathbf{0}}\chi_\sigma^{\mathbf{R}_\sigma}|\chi_\nu^{\mathbf{R}_\nu}\chi_\lambda^{\mathbf{R}_\lambda}) = \int\int \frac{\chi_\mu^{\mathbf{0}}(\mathbf{r})\chi_\sigma^{\mathbf{R}_\sigma}(\mathbf{r})\chi_\nu^{\mathbf{R}_\nu}(\mathbf{r}')\chi_\lambda^{\mathbf{R}_\lambda}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}'. \tag{15}$$

and

$$T_{\mu,\lambda,\nu,\sigma}^{\mathbf{R}_\mu\mathbf{R}_\lambda\mathbf{R}_\nu\mathbf{R}_\sigma}(t) = \sum_{\gamma\mathbf{R}_\gamma,\delta\mathbf{R}_\delta,\tau\mathbf{R}_\tau,\kappa\mathbf{R}_\kappa} X_{\mu\gamma}^{\mathbf{R}_\mu\mathbf{R}_\gamma} X_{\nu\delta}^{\mathbf{R}_\nu\mathbf{R}_\delta} Y_{\lambda\kappa}^{\mathbf{R}_\lambda\mathbf{R}_\kappa} Y_{\sigma\tau}^{\mathbf{R}_\sigma\mathbf{R}_\tau}(\chi_\gamma^{\mathbf{R}_\gamma}\chi_\kappa^{\mathbf{R}_\kappa}|\chi_\delta^{\mathbf{R}_\delta}\chi_\tau^{\mathbf{R}_\tau}) \tag{16}$$

The 4-fold k points are treated independently within $X_{\gamma\mu}^{\mathbf{R}_\gamma\mathbf{R}_\mu}$ and $Y_{\lambda\kappa}^{\mathbf{R}_\lambda\mathbf{R}_\kappa}$:

$$X_{\gamma\mu}^{\mathbf{R}_\gamma\mathbf{R}_\mu} = \sum_i^{occ}\frac{1}{V_k}\int \mathbf{d}\mathbf{k_i} C_{\gamma i}^*(\mathbf{k}_\mathbf{i})C_{\mu i}(\mathbf{k}_\mathbf{i})e^{(\epsilon_i-\epsilon_f)t}e^{i\mathbf{k_i}(\mathbf{R}_\mu-\mathbf{R}_\gamma)} \tag{17}$$

$$Y_{\lambda\kappa}^{\mathbf{R}_\lambda\mathbf{R}_\kappa} = \sum_a^{unocc}\frac{1}{V_k}\int \mathbf{d}\mathbf{k_a} C_{\lambda a}^*(\mathbf{k}_\mathbf{a})C_{\kappa a}(\mathbf{k}_\mathbf{a})e^{-(\epsilon_a-\epsilon_f)t}e^{i\mathbf{k_a}(\mathbf{R}_\kappa-\mathbf{R}_\lambda)} \tag{18}$$

In this way, the 4-fold integration over $\mathbf{k}$-points in Equation (6) can be reduced to 1-dimensional $\mathbf{k}$-points integral as shown in Equations (17) and (18). Furthermore, the locality of the atomic basis function can be adopted for the calculation of $T_{\mu,\lambda,\nu,\sigma}^{\mathbf{R}_\mu\mathbf{R}_\lambda\mathbf{R}_\nu\mathbf{R}_\sigma}$ and electron repulsion integrals, and the total computational scaling could be $O(N \cdot N_k)$ if the distant screening between these ERIs are applied. Here in this work, such distance screening has not been used, so our implementation results in a $O(N^2 \cdot N_k)$ scaling. It is worth noting that in order to keep the exponential value ($e^{(\epsilon_i)t}$) in Equation (17)/Equation (18) to be smaller than unity, we have inserted the Fermi energy level into the exponential factor ($e^{(\epsilon_i-\epsilon_f)t}$) in order to make the calculation to be more stable.

Based on Equations (7) and (10), we have the Laplace-transformed MP2 correlation correction ($\epsilon_g(\mathbf{k}_g)^{(2)}$) for the eigenstate:

$$\epsilon_g(\mathbf{k}_g)^{(2)} = \int \mathbf{d}t \sum_{\mu\mathbf{0},\nu\mathbf{R}_\nu,\lambda\mathbf{R}_\lambda,\sigma\mathbf{R}_\sigma} G_{\mu,\lambda,\nu,\sigma}^{\mathbf{0}\mathbf{R}_\nu\mathbf{R}_\lambda\mathbf{R}_\sigma}(t)[2(\chi_\mu^{\mathbf{0}}\chi_\lambda^{\mathbf{R}_\lambda}|\chi_\nu^{\mathbf{R}_\nu}\chi_\sigma^{\mathbf{R}_\sigma}) - (\chi_\mu^{\mathbf{0}}\chi_\sigma^{\mathbf{R}_\sigma}|\chi_\nu^{\mathbf{R}_\nu}\chi_\lambda^{\mathbf{R}_\lambda})] \tag{19}$$

$$G_{\mu,\lambda,\nu,\sigma}^{\mathbf{R}_\mu\mathbf{R}_\lambda\mathbf{R}_\nu\mathbf{R}_\sigma}(t) = \sum_{\gamma\mathbf{R}_\gamma,\delta\mathbf{R}_\delta,\tau\mathbf{R}_\tau,\kappa\mathbf{R}_\kappa} X_{\mu\gamma}^{\mathbf{R}_\mu\mathbf{R}_\gamma} Y_{\lambda\kappa}^{\mathbf{R}_\lambda\mathbf{R}_\kappa}(\chi_\gamma^{\mathbf{R}_\gamma}\chi_\kappa^{\mathbf{R}_\kappa}|\chi_\delta^{\mathbf{R}_\delta}\chi_\tau^{\mathbf{R}_\tau}) \times(-W_{\nu\delta}^{\mathbf{R}_\nu\mathbf{R}_\delta} Y_{\sigma\tau}^{\mathbf{R}_\sigma\mathbf{R}_\tau} + X_{\nu\delta}^{\mathbf{R}_\nu\mathbf{R}_\delta} Z_{\sigma\tau}^{\mathbf{R}_\sigma\mathbf{R}_\tau}) \tag{20}$$

$$W_{\gamma\mu}^{\mathbf{R}_\gamma\mathbf{R}_\mu} = C_{\gamma g}^*(\mathbf{k}_g)C_{\mu g}(\mathbf{k}_g)e^{(\epsilon_g)t}e^{i\mathbf{k}_g(\mathbf{R}_\mu-\mathbf{R}_\gamma)} \tag{21}$$

$$Z_{\lambda\kappa}^{\mathbf{R}_\lambda\mathbf{R}_\kappa} = C_{\lambda g}^*(\mathbf{k}_g)C_{\kappa g}(\mathbf{k}_g)e^{-(\epsilon_g)t}e^{i\mathbf{k}_g(\mathbf{R}_\kappa-\mathbf{R}_\lambda)} \tag{22}$$

Similarly, in order to keep the exponential value to be smaller than unity and avoid computational divergence, we inserted the VBM/CBM value into the exponential factor when calculated the MP2 reformulation of the VBM/CBM.

The canonical and Laplace-transformed MP2 methods described above have been implemented in the Order-N performance HONPAS code (Qin et al., 2014).

3. RESULTS

In order to validate our implementation, we perform benchmark calculations for 1-dimensional systems. We use norm-conserving pseudopotentials generated with the Troullier–Martins scheme to represent the interaction between core ion and valence electrons. The single-zeta (SZ), double-zeta (DZ), and double-zeta polarized (DZP) basis sets are generated using SIESTA. Then the NAOs are fitted with GTOs to perform the Hartree–Fock calculation as discussed in Shang et al. (2011).

First, we use a 1-dimensional hydrogen chain as an example to make the comparison between the results of canonical MP2 and those of Laplace-transformed MP2. The lattice parameter for the 1-dimensional H chain is set to 2.6 Å, and the H-H bond length is set to 1.346 Å. The SZ basis set is adopted, so that there are only two atomic orbitals in the unit cell. The Brillouin zone is sampled by $1 \times 1 \times 6$ k-points. The unit cell is a $20 \times 20 \times 2.6$ Å box, and the real-space integration mesh is set to be 100 Ry. In the Laplace-transformed MP2 method, the Romberg method is adopted to perform the final integration. The accuracy of integration results depends on the order of Romberg integration (n), since the results of the Romberg integration are obtained in a recursive manner, $R(n,j) = R(n,j-1) + \frac{R(n,j-1)-R(n-1,j-1)}{4^{j-1}-1}$. As shown in **Table 1**, when the order of Romberg integration increases from 3 to 8, the MP2 correlation correction for the unit cell energy (E_{mp2}) as well as the MP2 correlation correction for the band structure ($\epsilon^{(2)}_{\mathrm{VBM}}$, $\epsilon^{(2)}_{\mathrm{CBM}}$, $\epsilon^{(2)}_{\mathrm{gap}}$) are converged to the results of canonical MP2. When using the media precision parameter (n = 5) in Laplace- transformed MP2, we get a absolute/relative error of 4×10^{-6} eV/0.0007% for the correlation of the unit cell energy (E_{mp2}), and we get a absolute/relative error of 3×10^{-5} eV/0.02% for the MP2 correlation correction for the band gap ($\epsilon^{(2)}_{\mathrm{gap}}$). Overall, we find an excellent agreement between the Laplace-transformed MP2 and the canonical MP2 benchmark results.

We also examine the relative error between the results of Laplace-transformed MP2 and those of canonical MP2 with respect to the basis set size (SZ, DZ, DZP), as shown in **Table 2** with ethylene molecule as an example. Again, we find an excellent agreement between the Laplace-transformed MP2 and the canonical MP2 benchmark results.

TABLE 1 | The comparison between the results of canonical MP2 and those of Laplace-transformed MP2.

H2-line		**Laplace**		**Canonical**
n	**3**	**5**	**8**	
E_{mp2} (eV)	−0.561442	−0.561190	−0.561194	−0.561194
$\epsilon^{(2)}_{\mathrm{VBM}}$ (eV)	0.16736223	0.150361892	0.15033577	0.15033577
$\epsilon^{(2)}_{\mathrm{CBM}}$ (eV)	−0.03044718	−0.01888047	−0.01887073	−0.01887073
$\epsilon^{(2)}_{\mathrm{gap}}$ (eV)	−0.19780941	−0.16924236	−0.16920650	−0.16920650

Here, the MP2 correlation correction for the unit cell energy (E_{mp2}) and the MP2 correlation correction for the band structure ($\epsilon^{(2)}_{VBM}$, $\epsilon^{(2)}_{CBM}$, $\epsilon^{(2)}_{gap}$) have been examined with the above two approaches. Here, we use a 1-dimensional hydrogen chain as an example.

Second, we perform the Laplace-transformed MP2 calculation for the 1D polymer trans-polyacetylene as shown in **Table 3**. The order of Romberg integration is set to be n = 5. The SZ basis set is adopted in our calculation. The Brillouin zone is sampled by $1 \times 1 \times 30$ k-points. The real-space integration mesh is set to be 200 Ry. We compare our calculated MP2 correlation correction for the total energy per unit cell with the one obtained in Sun and Bartlett (1996). The G3 geometry parameters as listed in Sun and Bartlett (1996) are adopted to keep the geometry to be the same for comparison. We get an absolute/relative error of 0.26 eV/8% for the correlation correction of the unit cell energy (E_{mp2}). The difference comes from the usage of the different basis set, since in Sun and Bartlett (1996), the STO-3G basis set is used, whose shape is different from the SZ basis that we are using. For a similar reason, when using the same G6 geometry parameter of trans-polyacetylene (Sun and Bartlett, 1996), we get an absolute/relative error of 0.09 eV/7% for the correlation correction of the band gap ($\epsilon^{(2)}_{\mathrm{gap}}$) when compared with Sun's result.

We then investigate the performance and scaling of our implementation, and we show timings for the trans-polyacetylene molecules with variable number of atoms in **Figure 1**. We find a linear scaling for the calculation of ERIs and an $O(N^2)$ scaling for the calculation of the Laplace-transformed MP2. This is not too surprising, since we can see from Equation (14) that there are two loops over the ERIs for the calculation of Laplace-transformed MP2, so we get the $O(N^2)$ scaling.

Finally, we show the calculated binding-energy curves as functions of the distance between two trans-polyacetylene chains with the PBC-MP2 method. Although the MP2 theory gives overestimation of the dispersion interaction energy (Tkatchenko et al., 2009), it is still a superior starting point for the dispersion correction compared to Hartree–Fock and semi-local density functional theory (DFT). As shown in **Figure 2**, the MP2 method results in a binding states. On the contrary, Hartree–Fock and PBE functional fail to identify any binding between the two chains. We can see in **Figure 2** that the energy profile

TABLE 2 | The relative error between the results of canonical MP2 and those of Laplace-transformed MP2 with different basis set.

Ethylene (C_2H_4)	**SZ (%)**	**DZ (%)**	**DZP (%)**
Relative error	0.0007	0.003	0.001

Here, we use ethylene molecule as an example. The order of Romberg integration (n) in the Laplace transformed MP2 is set to n = 5.

TABLE 3 | The comparison between our results and those from the literature (Sun and Bartlett, 1996).

Trans-polyacetylene	**Sun and Bartlett (1996) (eV)**	**Our results (eV)**
E_{mp2}	−3.22	−2.96
$\epsilon^{(2)}_{\mathrm{gap}}$	−1.18	−1.09

Here, the MP2 correlation correction for the unit cell energy (E_{mp2}) and the MP2 correlation correction for the band gap ($\epsilon^{(2)}_{gap}$) have been examined.

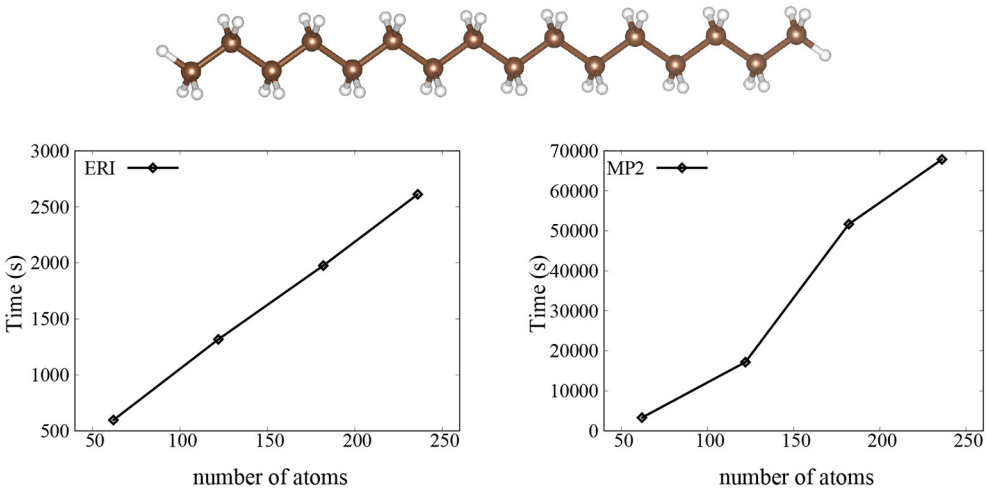

FIGURE 1 | The CPU time for the ERI and MP2 calculation using SZ basis set. Here, the trans-polyacetylene molecules are used as the test systems.

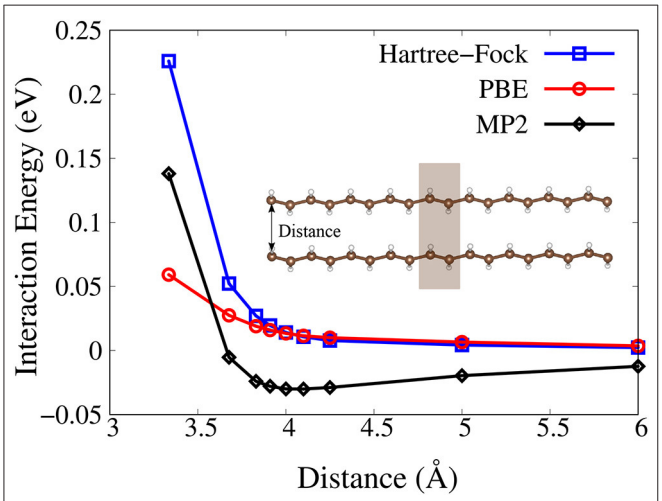

FIGURE 2 | Interaction energy as functions of the distance between two trans-polyacetylene chains as predicted by the Hartree-Fock (blue), PBE (red), and MP2 (black) method. The unit cell is marked with shaded box.

calculated with Hartree–Fock and PBE functional shows a repulsive behavior as the two chains are brought closer together.

4. CONCLUSIONS

We have implemented the canonical and Laplace-transformed algorithms to calculate the MP2 correlation correction for the total energy and the band gap of periodic systems in HONPAS code with numerical atomic orbitals. The results obtained by the canonical MP2 and Laplace-transformed MP2 are consistent with each other. We have also validated the implementation by comparing the results with the literature data. We have studied the binding-energy curves for the two stacked trans-polyacetylene chains, which shows the MP2 method can well describe the correlation energy and the long-range van der Waals interactions. Future work will address the application of the Laplace-transformed MP2 method to 3-dimensional periodic systems in the HONPAS code.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

JY designed and directed the project. HS developed the theoretical formalism, carried out the implementation, and performed the numerical simulations. Both JY and HS contributed to the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by CARCH4205, CARCH4411, and by the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDC01040100) and by National Natural Science Foundation of China (Grant No. 22003073).

ACKNOWLEDGMENTS

HS acknowledges Georg Kresse for inspiring discussions. HS thanks the Tianhe-2 Supercomputer Center for computational resources.

REFERENCES

Ayala, P. Y., Konstantin, K. N., Kudin, N., and Scuseria, G. E. (2001). Atomic orbital laplace-transformed second-order Møller-plesset theory for periodic systems. *J. Chem. Phys.* 115, 9698–9707. doi: 10.1063/1.1414369

Ayala, P. Y., and Scuseria, G. E. (1999). Linear scaling second-order Moller–Plesset theory in the atomic orbital basis for large molecular systems. *J. Chem. Phys.* 110, 3660–3671. doi: 10.1063/1.478256

Bartlett, R. J. and Stanton, J. F. (2007). "Applications of post-hartree-fock methods: a tutorial," *in Reviews in Computational Chemistry,* eds K.B. Lipkowitz and D.B. Boyd. doi: 10.1002/9780470125823.ch2

Doser, B., Lambrecht, D. S., and Ochsenfeld, C. (2008). Tighter multipole-based integral estimates and parallel implementation of linear-scaling AO–MP2 theory. *Phys. Chem. Chem. Phys*. 10, 3335–3344. doi: 10.1039/b804110e

Häser, M. (1993). Møller-Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* 87, 147–173. doi: 10.1007/BF01113535

Häser, M., and Almlöf, J. (1992). Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys*. 96, 489–494. doi: 10.1063/1.462485

Izmaylov, A. F., and Scuseria, G. E. (2008). Resolution of the identity atomic orbital Laplace transformed second order Møller–Plesset theory for nonconducting periodic systems. *Phys. Chem. Chem. Phys*. 10, 3421–3429. doi: 10.1039/b803274m

Katouda, M., and Nagase, S. (2010). Application of second-order Møller-Plesset perturbation theory with resolution-of-identity approximation to periodic systems. *J. Chem. Phys*. 133, 1–9. doi: 10.1063/1.3503153

Kobayashi, M., and Nakai, H. (2006). Implementation of Surján's density matrix formulae for calculating second-order Møller–Plesset energy. *Chem. Phys. Lett*. 420, 250–255. doi: 10.1016/j.cplett.2005.12.088

Lambrecht, D. S., Doser, B., and Ochsenfeld, C. (2005). Rigorous integral screening for electron correlation methods. *J. Chem. Phys*. 123:184102. doi: 10.1063/1.2079987

Marsman, M., Grüneis, A., Paier, J., and Kresse, G. (2009). Second-order Møller-Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys*. 130, 1–10. doi: 10.1063/1.3126249

Maschio, L. (2011). Local MP2 with density fitting for periodic systems: a parallel implementation. *J. Chem. Theory Comput*. 7, 2818–2830. doi: 10.1021/ct200352g

Pisani, C., Busso, M., Capecchi, G., Casassa, S., Dovesi, R., Maschio, L., et al. (2005). Local-MP2 electron correlation method for nonconducting crystals. *J. Chem. Phys*. 122:94113. doi: 10.1063/1.1857479

Pisani, C., Maschio, L., Casassa, S., Halo, M., Schütz, M., and Usvyat, D. (2008). Periodic local MP2 method for the study of electronic correlation in crystals: Theory and preliminary applications. *J. Comput. Chem*. 29, 2113–2124. doi: 10.1002/jcc.20975

Pulay, P. (1983). Localizability of dynamic electron correlation. *Chem. Phys. Lett*. 100, 151–154. doi: 10.1016/0009-2614(83)80703-9

Qin, X., Shang, H., Xiang, H., Li, Z., and Yang, J. (2014). HONPAS: a linear scaling open-source solution for large system simulations. *Int. J. Quantum Chem*. 115, 647–655. doi: 10.1002/qua.24837

Ren, X., Rinke, P., Blum, V., Wieferink, J., Tkatchenko, A., Sanfilippo, A., et al. (2012). Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys*. 14:053020. doi: 10.1088/1367-2630/14/5/053020

Saebø, S., and Pulay, P. (1993). Local treatment of electron correlation. *Annu. Rev. Phys. Chem*. 44, 213–236. doi: 10.1146/annurev.pc.44.100193.001241

Schäfer, T., Ramberger, B., and Kresse, G. (2018). Laplace transformed MP2 for three dimensional periodic materials using stochastic orbitals in the plane wave basis and correlated sampling. *J. Chem. Phys*. 148:064103. doi: 10.1063/1.5016100

Schütz, M., Hetzer, G., and Werner, H.-J. (1999). Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys*. 111, 5691–5705. doi: 10.1063/1.479957

Shang, H., Li, Z., and Yang, J. (2011). Implementation of screened hybrid density functional for periodic systems with numerical atomic orbitals: basis function fitting and integral screening. *J. Chem. Phys*. 135:034110. doi: 10.1063/1.3610379

Suhai, S. (1983). Perturbation theoretical investigation of electron correlation effects in infinite metallic and semiconducting polymers. *Int. J. Quantum Chem*. 23, 1239–1256. doi: 10.1002/qua.560230414

Suhai, S. (1992). Structural and electronic properties of infinitecis andtrans polyenes: perturbation theory of electron correlation effects. *Int. J. Quantum Chem*. 42, 193–216. doi: 10.1002/qua.560420112

Sun, J., and Bartlett, R. J. (1996). Second-order many-body perturbation-theory calculations in extended systems. *J. Chem. Phys*. 104, 8553–8565. doi: 10.1063/1.471545

Tkatchenko, A., DiStasio, R. A., Head-Gordon, M., and Scheffler, M. (2009). Dispersion-corrected Moller–Plesset second-order perturbation theory. *J. Chem. Phys*. 131:094106. doi: 10.1063/1.3213194

Usvyat, D., Maschio, L., and Schütz, M. (2018). Periodic and fragment models based on the local correlation approach. *WIREs Comput. Mol. Sci*. 8:e1357. doi: 10.1002/wcms.1357

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fchem.2020.589910

Parallel Implementation of Large-Scale Linear Scaling Density Functional Theory Calculations With Numerical Atomic Orbitals in HONPAS

Zhaolong Luo, Xinming Qin, Lingyun Wan, Wei Hu* and Jinlong Yang*

Hefei National Laboratory for Physical Sciences at the Microscale, Department of Chemical Physics, and Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, China

OPEN ACCESS

Edited by:
Hans Martin Senn,
University of Glasgow,
United Kingdom

Reviewed by:
Ma Yingjin,
Computer Network Information Center (CAS), China
C. Y. Yam,
Beijing Computational Science Research Center, China

****Correspondence:***
Xinming Qin
xmqin03@ustc.edu.cn
Wei Hu
whuustc@ustc.edu.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *31 July 2020*
Accepted: *08 September 2020*
Published: *26 November 2020*

Citation:
Luo Z, Qin X, Wan L, Hu W and Yang J (2020) Parallel Implementation of Large-Scale Linear Scaling Density Functional Theory Calculations With Numerical Atomic Orbitals in HONPAS. Front. Chem. 8:589910. doi: 10.3389/fchem.2020.589910

Linear-scaling density functional theory (DFT) is an efficient method to describe the electronic structures of molecules, semiconductors, and insulators to avoid the high cubic-scaling cost in conventional DFT calculations. Here, we present a parallel implementation of linear-scaling density matrix trace correcting (TC) purification algorithm to solve the Kohn–Sham (KS) equations with the numerical atomic orbitals in the HONPAS package. Such a linear-scaling density matrix purification algorithm is based on the Kohn's nearsightedness principle, resulting in a sparse Hamiltonian matrix with localized basis sets in the DFT calculations. Therefore, sparse matrix multiplication is the most time-consuming step in the density matrix purification algorithm for linear-scaling DFT calculations. We propose to use the MPI_Allgather function for parallel programming to deal with the sparse matrix multiplication within the compressed sparse row (CSR) format, which can scale up to hundreds of processing cores on modern heterogeneous supercomputers. We demonstrate the computational accuracy and efficiency of this parallel density matrix purification algorithm by performing large-scale DFT calculations on boron nitrogen nanotubes containing tens of thousands of atoms.

Keywords: linear-scaling density functional theory, density matrix purification algorithm, sparse matrix multiplication, parallel implementation, tens of thousands of atoms

1. INTRODUCTION

The Kohn–Sham density functional theory (DFT) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965) has been successfully applied to perform first-principles calculations for describing the electronic structures of both molecules and solids. However, conventional DFT calculations based on direct diagonalization methods for solving the KS equations have a high cubic-scaling cost (Goedecker, 1999), which can usually be used to study medium-scale systems containing up to hundreds of atoms. Therefore, it is difficult to achieve massive parallelism for these conventional cubic-scaling methods due to complex communication issues. To avoid the bottleneck arising from the computational cost and memory usage of directly diagonalizing the Hamiltonian matrix in conventional DFT calculations, linear-scaling methods using local basis functions

have been proposed (Goedecker, 1999), strongly promoting the applications of DFT calculations in large-scale systems containing thousands of atoms.

In general, linear-scaling methods include direct, variational, and purification methods (Bowler and Miyazaki, 2012). The direct methods are featured by direct evaluation of density matrix using various approximations, including divide and conquer (Yang, 1991; Yang and Lee, 1995) and Fermi operator expansion (Goedecker and Colombo, 1994; Goedecker and Teter, 1995; Liang et al., 2003). The variational methods minimize the total energy with respect to the auxiliary density matrix or Wannier-like orbitals, covering density matrix minimization method (Daw, 1993; Li et al., 1993; Nunes and Vanderbilt, 1994) and orbital minimization method (OMM) (Galli and Parrinello, 1992; Mauri and Galli, 1994; Kim et al., 1995; Ordejón et al., 1995). The third scheme exploits the purification polynomial and iterative solution, which is known as density matrix purification method (Palser and Manolopoulos, 1998; Niklasson, 2002; Niklasson et al., 2003). Nearly all of the linear-scaling methods are based on the Kohn's nearsightedness principle with localized basis sets, such as Gaussian type orbitals (GTOs) (Frisch et al., 1984) and numerical atomic orbitals (Shang et al., 2010) (NAOs), resulting in the sparsity of density matrix with a number of non-zero entries that increase linearly with the system size, so the linear-scaling matrix-matrix multiplication can be achieved (VandeVondele et al., 2012; Kim and Jung, 2016). In particular, the density matrix purification algorithms without prior knowledge of the chemical potential, including the trace-preserving canonical purification scheme of Palser and Manolopoulos (PM) (Palser and Manolopoulos, 1998; Daniels and Scuseria, 1999), the trace-correcting purification (TC) (Niklasson, 2002), and the trace resetting density matrix purification (TRS) (Niklasson et al., 2003), have been demonstrated as accurate and efficient linear-scaling methods to describe the electronic structures of molecules, semiconductors, and insulators. However, almost all of the developed linear-scaling techniques (direct, variational, and purification methods) assume the presence of a non-zero gap in the electronic structure, which prevents them from treating metallic systems. Recently, Suryanarayana (2017) have employed the $O(N)$ Spectral Quadrature (SQ) method (Suryanarayana, 2013; Pratapa et al., 2016) to study the locality of electronic interactions in aluminum (a prototypical metallic system) as a function of smearing/electronic temperature. They have found exponential convergence accompanied by a rate that increases sub-linearly with smearing. It is also worth mentioning that all these linear-scaling methods based on Kohn's nearsightedness principle are limited to the localization of density matrix (Bowler and Miyazaki, 2012). A recently published innovative version of PEXSI scheme named iPEXSI (Etter, 2020), which does not rely on the nearsightedness principle, can scale provably better than cubically even in the absence of localization of density matrix. The iPEXSI algorithm utilizes a localization property of triangular factorization, which could extend the usable range of linear-scaling method to metallic system without the constraint of finite electronic temperature.

Nowadays, with the rapid development of modern heterogeneous supercomputers, the high-performance computing (HPC) has become a powerful tool for accelerating the DFT calculations to deal with large-scale systems. Several highly efficient DFT software based on low-scaling methods have been developed, such as SIESTA (Soler et al., 2002), OPENMX (Ozaki and Kino, 2005), CP2K (Kühne et al., 2020), CONQUEST (Gillan et al., 2007), PROFESS (Ho et al., 2008), FREEON (Challacombe, 2014), ONETEP (Skylaris et al., 2005), BigDFT (Genovese et al., 2008; Mohr et al., 2014), FHI-aims (Blum et al., 2009), ABACUS (Chen et al., 2010, 2011), HONPAS (Qin et al., 2015), and DGDFT (Lin et al., 2012; Hu et al., 2015a,b; Banerjee et al., 2016; Zhang et al., 2017), which are capable to make full advantage of the massive parallelism available on HPC architectures beneting from the local data communication of sparse Hamiltonian matrix generated with local basis sets. In linear-scaling DFT calculations, the kernel for HPC is to parallel sparse matrix–matrix multiplication. In order to realize the HPC parallelism, two massively parallel libraries of BCSR (Borštnik et al., 2014) and NTPOLY (Dawson and Nakajima, 2018) have been developed, which have shown a high performance for the density matrix purification algorithms implemented in the CP2K (Kühne et al., 2020) and CONQUEST (Gillan et al., 2007) packages.

In this work, we present a parallel implementation of linear-scaling density matrix second-order trace-correcting purification (TC2) algorithm (Niklasson, 2002) to solve the KS equations with the NAOs in the HONPAS package (Qin et al., 2015). We propose to use the MPI_Allgather function for parallel programming to deal with such sparse matrix multiplication within the CSR format, which can be scaled linearly up to hundreds of processing cores on modern heterogeneous supercomputers. We demonstrate the computational accuracy and efficiency of this linear-scaling density matrix purification method by performing large-scale DFT calculations on boron nitrogen nanotubes containing thousands of atoms.

2. METHODOLOGY

2.1. Density Functional Theory

We first give a brief review of Kohn–Sham density functional theory (KS-DFT). The key spirit of KS-DFT is to solve the KS equations defined as

$$\hat{H}\psi_i(\mathbf{r}) = (\hat{T} + \hat{V}_{\text{ion}} + \hat{V}_{\text{H}} + \hat{V}_{\text{xc}})\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{1}$$

where $\hat{H}$ is the Hamiltonian operator, ψ_i is the ith molecular orbital, and ϵ_i is the corresponding orbital energy. $\hat{T}$ is the kinetic operator, $\hat{V}_{\text{ion}}$ is the ionic potential operator, and $\hat{V}_{\text{H}}$ is the Hartree potential operator defined as

$$\hat{V}_{\text{H}}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{2}$$

where the electron density is given by

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_e} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{3}$$

In the approximation of linear combination of atomic orbitals (LCAO) (Mulliken, 1955), the ψ_i is expanded on a set of NAOs $\{\phi_\mu(r)\}_{\mu=1}^{N_\mathrm{b}}$

$$\psi_i(\mathbf{r}) = \sum_{\mu}^{N_\mathrm{b}} c_{\mu i}\phi_\mu(\mathbf{r}) \tag{4}$$

where $c_{\mu i}$ is the expansion coefficient at the μth atomic orbital and N_b is the number of NAOs. Then, the KS equations can be rewritten into matrix notations as

$$HC = SCE \tag{5}$$

where C is coefficient matrix and E is the corresponding orbital energy. H and S are the Hamiltonian and overlap matrices over the NAOs

$$\begin{aligned} H_{\mu\nu} &= \int \phi_\mu^*(\mathbf{r})\hat{H}\phi_\nu(\mathbf{r})d\mathbf{r} \\ S_{\mu\nu} &= \int \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r})d\mathbf{r} \end{aligned} \tag{6}$$

The default choice in the SIESTA package is to use the direct diagonalization method though the LAPACK and ScaLAPACK libraries to solve this eigenvalue problem with a high cubic-scaling cost. Therefore, the computational cost and memory usage of such DFT calculations increase rapidly as the system size, which are only limited to small systems containing hundreds of atoms. In order to overcome this limitation, several linear-scaling methods have been implemented in the SIESTA package, such as the Kim–Mauri–Galli (KMG) orbital minimization (OMM) method (Galli and Parrinello, 1992; Mauri and Galli, 1994; Kim et al., 1995; Corsetti, 2014) and divide and conquer method (Cankurtaran et al., 2008). The KMG requires a initial approximate Wannier functions and a prior knowledge of the chemical potential. In the HONPAS-SIESTA package (Qin et al., 2015), we implement the density matrix purification algorithms, including the trace-preserving canonical purification scheme of PM (Palser and Manolopoulos, 1998; Daniels and Scuseria, 1999), the trace-correcting purification (TC) (Niklasson, 2002), and the trace resetting density matrix purification (TRS) (Niklasson et al., 2003).

2.2. Linear-Scaling Density Matrix Purification Algorithms

After constructing the Hamiltonian matrix, the density matrix can be obtained by directly diagonalizing the Hamiltonian matrix with cubic-scaling cost. In order to avoid the high cost of explicit diagonalization, we implement three density matrix purification algorithms, without prior knowledge of the chemical potential for linear-scaling DFT calculations, including the trace-preserving canonical purification scheme of PM, the trace-correcting purification (TC) (Niklasson, 2002), and the trace resetting density matrix purification (TRS) (Niklasson et al., 2003), in the HONPAS package (Qin et al., 2015). In this work, we use the second-order trace-correcting purification (TC2) (Niklasson, 2002) algorithm with orthogonal basis sets to illustrate our parallel algorithms. In the coordinate presentation, the general form of density matrix can be given by

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{i=1}^{N_b} f(\varepsilon_i)\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') \tag{7}$$

where $f(\varepsilon_i)$ is the Fermi distribution function of energy level ε_i at finite electronic temperature

$$f(\varepsilon_i) = \frac{1}{1+e^{\beta(\varepsilon_i-\mu)}} \tag{8}$$

with the chemical potential μ and the inverse temperature $\beta = 1/k_BT$. Within the LCAO method, we can transform the density matrix from coordinate presentation to the basis presentation, then the density matrix element $P_{\mu\nu}$ becomes:

$$\begin{aligned} P_{\mu\nu} &= \int \phi_\mu^*(\mathbf{r})\rho(\mathbf{r},\mathbf{r}')\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \int \phi_\mu^*(\mathbf{r})\sum_{i=1}^{N_\mathrm{b}} f(\varepsilon_i)\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \sum_{i=1}^{N_b} f(\varepsilon_i)\sum_{\lambda}^{N_\mathrm{b}} c_{\lambda i}\sum_{\kappa}^{N_\mathrm{b}} c_{\kappa i}^*\int \phi_\mu^*(\mathbf{r})\phi_\lambda(\mathbf{r})d\mathbf{r}\int \phi_\kappa^*(\mathbf{r}')\phi_\nu(\mathbf{r}')d\mathbf{r}' \\ &= \sum_{i=1}^{N_b} f(\varepsilon_i)\sum_{\lambda}^{N_\mathrm{b}} c_{\lambda i}\sum_{\kappa}^{N_\mathrm{b}} c_{\kappa i}^* S_{\mu\lambda}S_{\nu\kappa} \end{aligned} \tag{9}$$

If NAOs are orthogonal, the density matrix element $P_{\mu\nu}$ can be written as

$$P_{\mu\nu} = \sum_{i}^{N_b} f(\varepsilon_i)c_{\mu i}c_{\nu i}^* \tag{10}$$

Note that ε_i is relative to the eigenvalue of $\hat{H}\psi_i = \varepsilon_i\psi_i$, so $P_{\mu\nu}$ can be rewritten as

$$\begin{aligned} P_{\mu\nu} &= \int \phi_\mu^*(\mathbf{r})\rho(\mathbf{r},\mathbf{r}')\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \sum_{i}^{N_b}\int \phi_\mu^*(\mathbf{r})f(\varepsilon_i)\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \int \phi_\mu^*(\mathbf{r})f(\hat{H})\sum_{i}^{N_b}(\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}'))\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \int \phi_\mu^*(\mathbf{r})f(\hat{H})\phi_\nu(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \end{aligned} \tag{11}$$

which implies that P is commutative with H, namely $[H,P] = 0$. Another substantial property of the appropriate density matrix is particle conservation, $\mathrm{Tr}(P) = N_e/2$.

When the electronic temperature is zero, $f(\varepsilon_i) = 1$ and the density matrix of insulator can be written as

$$P_{\mu\nu} = \sum_{i}^{N_e} c_{\mu i}c_{\nu i}^* \tag{12}$$

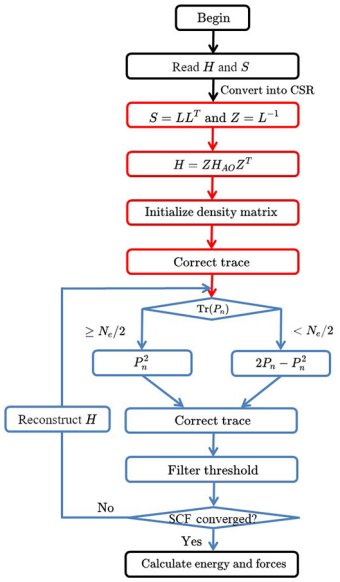

FIGURE 1 | The flowchart of density matrix purification TC2 method. There are four time-consuming parts in the TC2 method, including constructing the Hamiltonian matrix, initializing the density matrix P from the Hamiltonian matrix with Cholesky and Lanzos methods, updating the density matrix with parallel sparse matrix–matrix multiplication, and computing total energy and atomic forces after SCF iterations.

which must satisfy the so-called idempotency $PP = P$ with an orthogonal basis.

The solution of eigenvalue problem is under three restricted conditions of the density matrix mentioned above, which is known as the purification method. The trace-preserving canonical purification scheme of PM (Palser and Manolopoulos, 1998), which imposes commutation relation and trace-preserving condition, works with a predefined occupation and does not need the input or adjustment of the chemical potential. Trace-conserving spectral projections are performed during each iteration, until the density matrix P_n converges to the correct one that satisfies the idempotency condition. This method is inefficient at low and high partial occupancies (Palser and Manolopoulos, 1998; Daniels and Scuseria, 1999). A subsequent strategy proposed by Niklasson named TC algorithm (Niklasson, 2002). Its second-order form is called the second-order trace-correcting purification (TC2) method (Niklasson et al., 2003). The higher order TC2 requires additional matrix multiplications, which pursues a more rapid reduction of errors and a less step of purification iterations (Kim and Jung, 2016).

The TC2 purification algorithm is simple, robust, and rapidly convergent for closed-shell systems, and more efficient in orthogonal basis sets (Xiang et al., 2005). In this work, H denotes the Hamiltonian matrix under the presentation of orthogonal basis sets. Reasonably in the preparatory step, a transformation $H = Z^{\mathrm{T}}H_{\mathrm{AO}}Z$ is required, here the matrix Z is obtained by solving out the inverse square root of overlap matrix S by the Cholesky factorization (Cholesky, 2005). The idempotency and commutativity are satisfied naturally since the initial guess P_0 is obtained by the Lanczos method (Lanczos, 1950).

Algorithm 1 The pseudocode of TC2 algorithm, where N_e is the number of electrons, E is the energy-density matrix, $\varepsilon_{\min}(H)$ and $\varepsilon_{\max}(H)$ denote the minimal and maximum eigenvalue of the Hamiltonian matrix H, respectively.

subroutine TC2 (H, P, N_e)
$S = LL^T$
$Z = L^{-1}$
$H = ZH_{AO}Z^T$
$P_0 = (\varepsilon_{\max}I - H) / (\varepsilon_{\max} - \varepsilon_{\min})$
do iter = 1, niter
  if $\mathrm{Tr}(P_n) \leq N_e/2$ **then**
    $P_{n+1} = P_n^2$
  else
    $P_{n+1} = 2P_n - P_n^2$
    $\delta = (\mathrm{Tr}(P_{n+1}H) - \mathrm{Tr}(P_nH)) / \mathrm{Tr}(P_nH)$
enddo
if (Converged) **then**
$P_{AO} = Z^TPZ$
$E = P_{AO}H_{AO}S^{-1} = P_{AO}H_{AO}Z^TZ$
endif
end subrouine

During each iteration step, the trace of P_{n+1} is corrected by P_n^2 ($\mathrm{Tr}(P_n) \leq N_e/2$) or $2P_n - P_n^2$ ($\mathrm{Tr}(P_n) > N_e/2$). Then, matrix elements less than a numerical threshold δ_{filter} (10^{-4} or 10^{-6}) are dropped to zero, thus maintaining the sparsity of density matrix. The pseudocode of the TC2 algorithm is given in Algorithm 1 and its corresponding owchart is shown in **Figure 1**.

The density matrix purification method in HONPAS is based on the fact that both the density matrix and Hamiltonian matrix are sparse with NAOs. Therefore, sparse matrix multiplication is the most expensive step in the density matrix purification method. **Figure 2** shows the sparsity of the density matrix P saved as CSR format for the BN nanotubes consisting of 100 and 1,000 atoms (BNNT100 and BNNT1000) with different basis sets [single-ζ (SZ), double-ζ (DZ), and double-ζ plus polarization (DZP)] and thresholds ($\delta_{\mathrm{filter}} = 10^{-4}$ and 10^{-6}) of non-zero elements in the density matrix P.

When the selected basis sets are strictly localized, the coefficient matrix C for BNNT100 and BNNT1000 systems formed in a similar block diagonal matrix and arbitrary row of the resultant P is occupied by same number of non-zero elements since $P = CC^T$. Therefore, the total number of non-zero elements grows linearly with the system size under tight binding approximation, which is the substantial precondition of almost all linear scaling algorithms. If we have an observation onto the first two columns of **Figure 2**, matrices show block-multidiagonal patterns and sparse degree of BNNT 1000 with the SZ basis set under $\delta_{\mathrm{filter}} = 10^{-4}$ is 11.6%, which is obviously less than that of BNNT100 (87.7%). Variation trend of sparse degree is consistent with the character mentioned above from a perspective of image. Moreover, the cutoff radius r_c of the DZP basis set is higher, so non-zero elements per row distribute more intensively

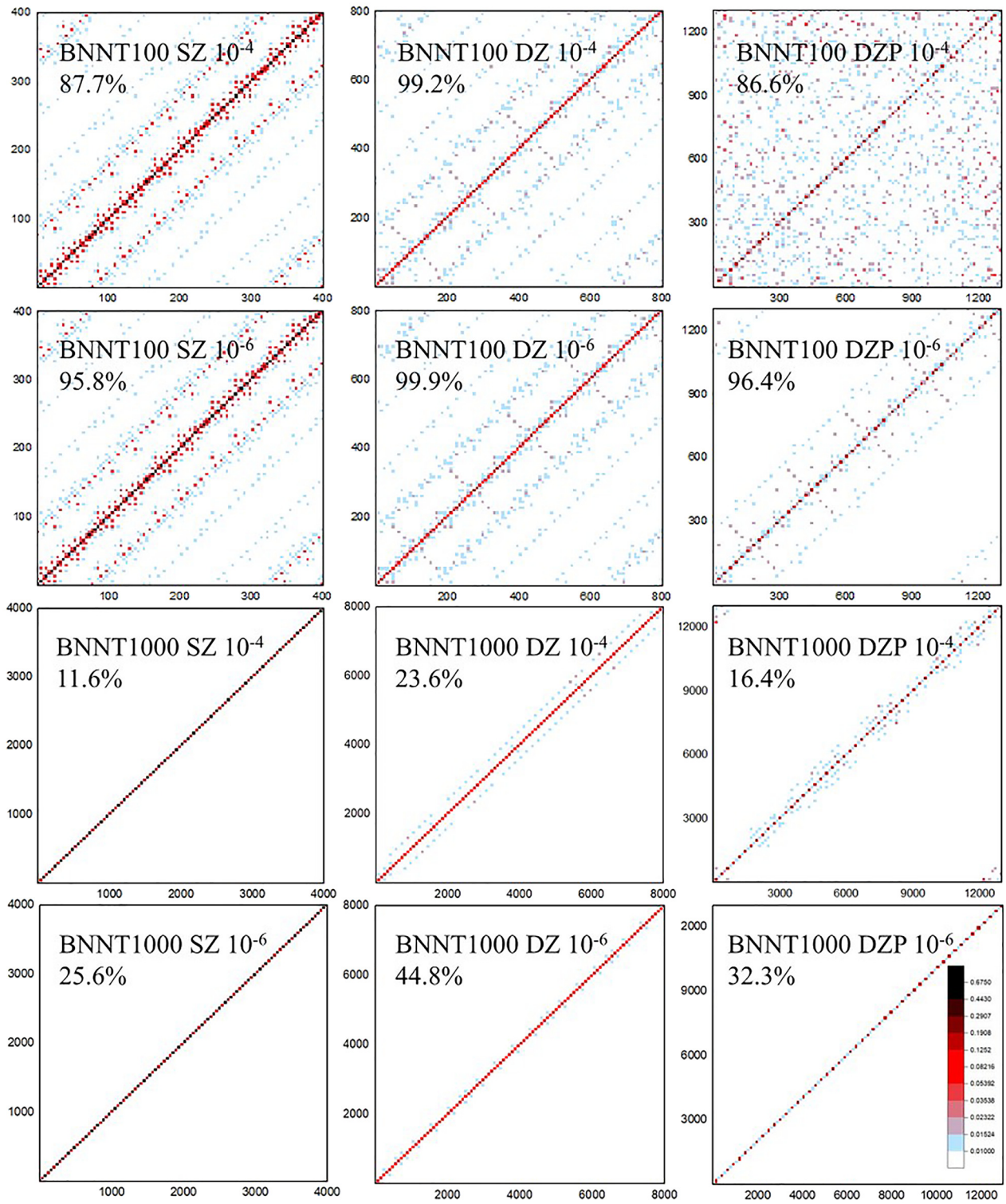

FIGURE 2 | The sparsity of the density matrix P in the compressed sparse row (CSR) format of BN nanotubes (BNNTs) consisting of 100 and 1,000 atoms (BNNT100 and BNNT1000). Percentage under the system name indicates sparse degree of P. The matrix elements exceeding 10^{-2} are labeled as light blue and those exceeding 10^{-1} are marked as red. White area is remained for elements with tiny values or zero. Density matrix becomes less occupied with its elements gathering close to diagonal when the threshold after multiplication 100× tighter (10^{-6}) than that of higher occupied matrix (10^{-4}). When the basis sets become larger from SZ to DZP, sparse degree varies and δ_{filter} strongly affects the sparsity pattern.

but their value is smaller than that of SZ and DZ. Sparse degree of BNNT1000 also decreases significantly from that of BNNT100 in the case of DZP basis set, such as 86.6% of BNNT100 and 16.4% of BNNT1000 at the same threshold, respectively.

On the perspective of threshold as shown in **Figure 2**, we observe that, as the system size increases, the influence of δ_{filter} is more significant, either the patterns or the sparse degree display an obvious difference under $\delta_{\text{filter}} = 10^{-4}$ and 10^{-6}. For example, the density matrix of BNNT1000 with the DZP basis set is 32.3% occupied under $\delta_{\text{filter}} = 10^{-6}$ and 16.4% occupied under $\delta_{\text{filter}} = 10^{-4}$. On the other hand, δ_{filter} has a slight influence on the sparsity of BNNT100 compared with BNNT1000, which implies that the distribution of numerical value of matrix elements is shifted to higher level than that of BNNT1000. Just as elements dotted with deep color in BNNT100 are much more intensive than those of BNNT1000. It should be noted that dropping matrix elements less than δ_{filter} and using strictly truncated NAOs both contribute to the sparsity of P.

2.3. Parallel Implementation of Sparse Matrix Multiplication

The time required to process matrix–matrix multiplications during each iteration step accounts for a major part of total time. Note that there are some additional steps such as data communication and matrix addition. Fortunately, all matrices we need to deal with are sparse, so that the number of dot products reduces. The linear scaling cost arises from the fact that all matrix operations are performed on sparse matrices, which has a number of non-zero entries that increase linearly with the system size (VandeVondele et al., 2012).

The sparsity of matrix also causes unexpected drawbacks. An apparent disadvantage is, the matrix multiplication step would change the sparsity pattern during the self-consistent field (SCF) iterations, resulting in a load imbalance between matrix computation and data commutation among different processing cores. Since each matrix is distributed on a series of processes in advance, the instability of sparsity pattern will occur at each iteration, thus we also need to modify the data distribution after each iteration or exploit a block-cyclic distribution scheme. Apart from those, dropping matrix elements with the numerical value less than a threshold can reduce the number of dot products. But the computational accuracy of total energy and atomic forces is sacrificed inevitably under a loose threshold. The parallel version of TC2 algorithm in HONPAS is based on CSR data format and message-passing interface (MPI), which is capable of performing massive parallelism on modern heterogeneous supercomputers. We employ the SPARSEKIT library to manipulate and deal with sparse matrices, which provides programs for converting data structures, filtering out elements, and performing basic linear algebra operations with sparse matrix (Saad, 1994).

In the parallel TC2 module, there is a single hierarchical structure of parallelization that consists of single type of data distribution and communication scheme. The TC2 module utilizes the MPI parallel programming to deal with data communications between different MPI processes. In our work, the MPI processes are organized in 1D row MPI grids. The density matrix is distributed by 1D row blocks across MPI processes, and each process saves N_b/N_p rows of global matrix. Thus, such local and global sparse matrix–matrix multiplication does not require additional data communication. Individual process computes its part of the multiplication, processing a row block of np ($n = 1, 2, \ldots, N_p$) at a time. After the local multiplication has been processed, each processor just gathers a local subset of global density matrix. We use the MPI_Allgather function to gather local matrices into global density matrix in each MPI process, similar to the case of MPI_Gather and then MPI_Bcast, then performing local sparse matrix–matrix multiplication at the next iteration step. **Figure 3** illustrates the schematic diagram of MPI communication on CSR data format, in which we set $N_p = 4$ to simplify the discussion.

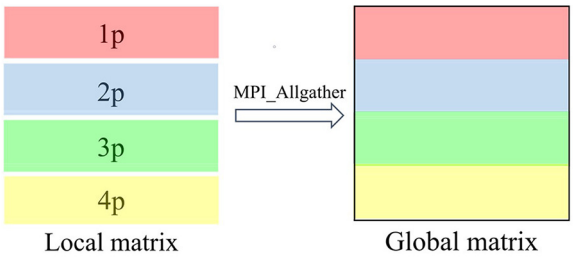

FIGURE 3 | Parallel data distribution and communication of sparse matrix multiplication when $N_p = 4$. The density matrix is partitioned into four row block local matrices with 1D row BN nanotube (BNNT) grid parallelism (1p, 2p, 3p, and 4p). Each local matrix is stored in the compressed sparse row (CSR) format. MPI_Allgather is invoked to integrate these four row block local matrices into a global matrix in the CSR format.

3. RESULTS AND DISCUSSION

In this section, we demonstrate the computational accuracy and efficiency of our parallel TC2 algorithm. We implement this method in the HONPAS package (Qin et al., 2015), which has been written in the Fortran programming language with the MPI for parallelism. We use the norm-conserving Troullier-Martins pseudopotentials (Troullier and Martins, 1991) to represent interaction between core and valence electrons. We use the exchange-correlation functional of local density approximation of Goedecker-Teter-Hutter (LDA-PZ) (Goedecker et al., 1996) to describe the electronic structures of these BNNTs with a grid cutoff of 100 Ry. In our calculations, the NAOs are generated by default parameters in SIESTA. We utilize the linear-scaling density matrix TC2 purification algorithm to calculate the electronic structures of a series of boron nitride nanotubes (BNNTs), containing 100–18,000 atoms (labeled by BNNT100-BNNT18000). The details of the input parameters and atomic structures of BNNTs used in this work as well as the performance data are shown in the **Supplementary Materials**.

TABLE 1 | Absolute error of total energy ΔE_{tot} (eV/atom) and the maximum of root mean square error of atomic forces ΔF_{max} (eV/Å) of the TC2 method with varying thresholds of $\delta_{filter} = 10^{-4}$ and 10^{-6} for the BNNT100 system with SZ, DZ, and DZP basis sets.

Basis sets	δ_{filter}	ΔE_{tot}	ΔF_{max}
SZ	10^{-4}	3.36×10^{-4}	8.48×10^{-3}
DZ	10^{-4}	2.26×10^{-3}	9.80×10^{-2}
DZP	10^{-4}	8.80×10^{-3}	3.07×10^{-1}
SZ	10^{-6}	3.70×10^{-7}	3.64×10^{-5}
DZ	10^{-6}	4.72×10^{-5}	1.73×10^{-4}
DZP	10^{-6}	1.06×10^{-5}	4.30×10^{-4}

The reference results are computed by the direct diagonalization method.

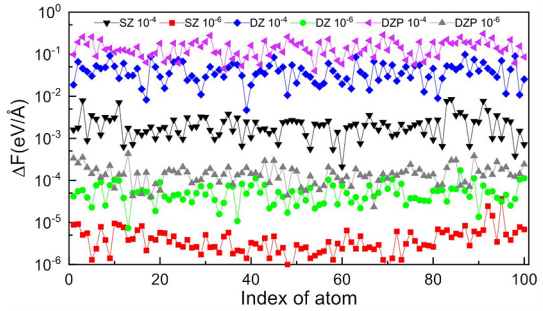

FIGURE 4 | Variation of root mean square error of atomic force on each atom computed with the TC2 and diagonalization methods with different basis sets (SZ, DZ, and DZP) and thresholds ($\delta_{filter} = 10^{-4}$ and 10^{-6}) for BNNT100.

3.1. Accuracy

We benchmark the computational accuracy of parallel TC2 method by comparing the results with those obtained from the diagnonalization method. We consider the effects of both the size of basis sets (SZ, DZ, and DZP) and different values of thresholds ($\delta_{filter} = 10^{-4}$ to 10^{-6}) on the computational accuracy of density matrix TC2 purification algorithm. We define the errors of total energy and atomic forces, respectively, as

$$\Delta E_{tot} = \mid \Delta E^{TC2} - E^{DIAG} \mid / N_A$$
$$\Delta F_I = \mid F_I^{TC2} - F_I^{DIAG} \mid$$

where N_A is the total number of atoms and I is the atom index.

In the HONPAS calculations, the default convergence accuracy for total energy and atomic forces are 10^{-4} eV/atom and 0.02 eV/Å, respectively. **Table 1** shows that the TC2 purification calculation for total energy is performed very well when choosing a tight dropping threshold, and $\delta_{filter} = 10^{-6}$ can yield a total energy accuracy of 10^{-5} eV/atom at least. On the other hand, strictly truncated NAOs can yield the sparsity without loss of accuracy simultaneously (Shang et al., 2010). We compute the total energy and atomic forces under different basis sets using a variable threshold. As shown in **Figure 4**, the errors of atomic forces from TC2 and those obtained from direct diagonalization method are indistinguishable. For all tested systems, the accuracy of the TC2 method can be obviously improved by tightening the threshold (10^{-4} to 10^{-6}). In particular, when the threshold is set to 10^{-4}, ΔF_{max} with the most general case of DZP basis set reaches 10^{-1} eV/Å, which is already comparable to the magnitude of atomic force itself. In contrast, $\Delta F_{max} = 4.30 \times 10^{-4}$ eV/Å under $\delta_{filter} = 10^{-6}$ with the same basis set. The noticeable error arises from the lack of information in density matrix when too many elements are neglected after each iteration step and the information of Hamiltonian matrix just included in the initial step. However, the relative error of energy per atom is less than 10^{-7} when threshold is set to 10^{-6} in the case of SZ basis set, which indicates that the computational accuracy of TC2 method is still guaranteed. On the perspective of basis sets, high accuracy is ensured when we employ rigorously localized basis sets (SZ). Note that systems with the DZP basis set have relatively larger errors, since the information of polarization orbital is partly omitted by dropping matrix elements. For instance, when δ_{filter} is set to 10^{-4}, the energy error for SZ is 3.36×10^{-4} but that for DZP is 8.80×10^{-3} (still can achieve the converged accuracy). As we have mentioned in section 2, non-zero elements those hold relatively small value distribute more intensively in the case of DZP basis set, and the physical information can be seriously lacking under relatively large δ_{filter}.

3.2. Efficiency

We demonstrate the computational efficiency and parallel scalability of linear-scaling TC2 method by checking the weak and strong scaling performance on BNNT systems with the SZ basis set and a threshold of $\delta_{filter} = 10^{-4}$. We illustrate the total time of the main time-consuming parts as shown in **Figure 1**: (a) Construction of Hamiltonian matrix, (b) evaluation of density matrix P from Hamiltonian matrix by Cholesky factorization following Lanczos method, and (c) purification with matrix multiplication and addition. It should be noticed that the data communication via MPI interface also occupies numerous time resource while performing massive parallelization over plenty of processing cores. Practical tests on the computational efficiency and parallel scalability are performed in the case of BNNT systems with MPI parallelism on modern heterogeneous supercomputers, including comparison of TC2 and diagonalization methods with respect to different system sizes and process counts, as shown in **Figures 5**, **6**, respectively. There are some additional steps such as computing total energy and atomic forces, which are all included in the total wall clock time of outer SCF iterations in the TC2 method.

Since the computational cost of linear-scaling TC2 method grows linearly with respect to the system size, a noteworthy speed up is supposed to be observed. We choose all tested systems with the SZ basis set to illustrate strong and weak scaling behaviors, since it is more strictly localized, resulting in a relatively small change of sparsity pattern after each iteration step. The variation of total time with respect to the system size is plotted in **Figure 5**. We can see that the scaling of TC2 is fitted to $O(N)$ due to the linearly growing sparse degree of P, and the number required to perform multiplication has the same trend. Linear scaling behavior is obtained with various systems containing 2,000–9,000

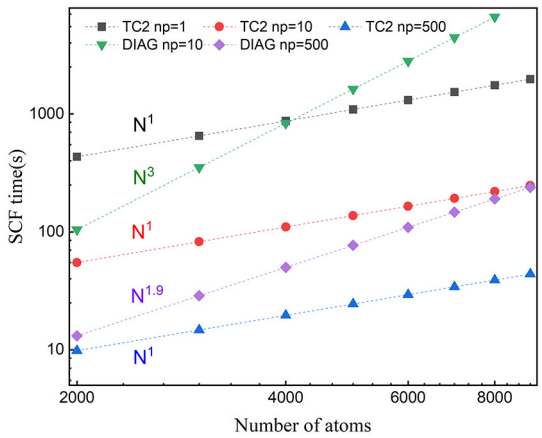

FIGURE 5 | Weak scaling of wall clock time per SCF iteration with respect to the number of atoms with the message-passing interface (MPI) parallelism for BN nanotubes (BNNTs) with 2,000 and 9,000 atoms (BNNT2000 and BNNT9000) computed with the TC2 and diagonalization methods.

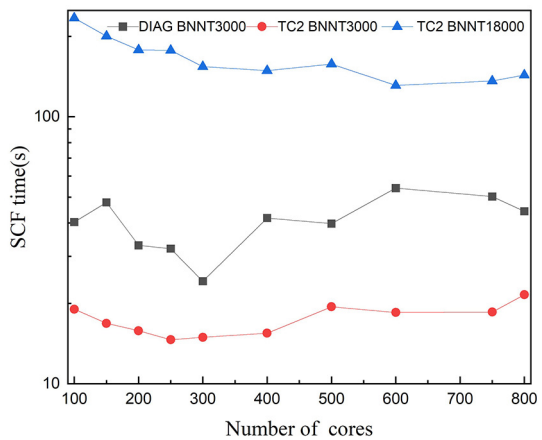

FIGURE 6 | Strong scaling of wall clock time per SCF iteration with respect to the number of cores with the message-passing interface (MPI) parallelism for BN nanotubes (BNNTs) with 3,000 and 18,000 atoms (BNNT3000 and BNNT18000) computed with the TC2 and diagonalization methods.

atoms under serial mode ($N_p = 1$), and it continues to scale further up to 500 cores at least, which benefits from the efficient parallel implementation of matrix multiplications based on the CSR formatted sparse density matrix. A speed-up of 4.7 can be achieved for 9,000 atoms (500 cores) and could be larger for more atoms. The fitted scaling for explicit diagonalization is just $O(N^{1.9})$ with number of atoms fewer than 5,000 when the number of processors is relatively large, arising from the load imbalance that problem size (number of computational tasks) distributed on each process is not adequate and some cores remain idle. If the processors keep increasing, low efficiency of parallelization is going to happen. When the size of system grows sufficiently or processing cores have a relevant scale, fitted scaling turns back to $O(N^3)$ due to the cubic scaling of conventional diagonalization step. As a conclusion, linear-scaling TC2 method outperforms explicit diagonalization in terms of expansibility to large systems and massive parallel implementation.

Figure 6 compares the parallel scalability of TC2 to diagonalization methods. As it can be seen, the parallel scalability of both methods is unsatisfactory, especially with the smaller system size. This issue arises in the load imbalance caused by idle processors since computational tasks are inadequate compared with hundreds of cores. Test for 18,000 atoms with diagonalization is not represented due to a memory overflow problem (the dimension of matrix is 72,000). Unlike the diagonalization method, test for TC2 has been performed since the utilization of CSR data format reduces the memory requirement. TC2 method demonstrates just scaling up to 600 cores, since the 1D processes layout prevents it from massive parallelization. The performance of global MPI communications such as MPI_Allgather is strongly impacted by the physical distance of remote processing cores, which prompts us to utilize BCSR storage format and 2D block-cyclic processor layout.

4. CONCLUSION AND OUTLOOK

In summary, we present a parallel implementation of linear-scaling density matrix trace correcting (TC) purification algorithm to solve KS equations with numerical atomic orbitals in the HONPAS package. We use the MPI_Allgather function for parallel programming to deal with the sparse matrix multiplication within the CSR format, which can scale up to hundreds of processing cores on modern heterogeneous supercomputers. We demonstrate the computational accuracy and efficiency of this linear-scaling density matrix purification algorithm by performing large-scale DFT calculations on boron nitrogen nanotubes containing tens of thousands of atoms.

However, our parallel implementation of TC2 method in HONPAS is inferior to that of BigDFT (Genovese et al., 2008; Mohr et al., 2014), ONETEP (Skylaris et al., 2005), and CONQUEST (Gillan et al., 2007). They exploit more than one level of organization and data distribution schemes resembling the BCSR format to handle the groups of atoms, which achieve high flexibility in load balancing (Bowler et al., 2002) with high performance on modern heterogeneous supercomputers. In the future, We plan to implement a massively parallel algorithm based on the NTPoly library (Dawson and Nakajima, 2018) in HONPAS. The NTPoly library utilizes the 3D sparse matrix multiplication algorithm, that is, the processors are organized into a three dimensional, cube-shaped virtual topology. In this case, density matrix purification algorithms can scale up to thousands of processing cores on modern heterogeneous supercomputers.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

ZL has tested the computational accuracy and efficiency of the TC2 method in HONPAS package, written manuscript and drawn figures. WH and XQ provide the research direction and revise the manuscript. WH has provided codes of the parallel version of TC2 method in HONPAS and directive methods for tests of computational accuracy and efficiency. LW has offered help for the modification of formulas. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partly supported by the National Natural Science Foundation of China (21688102, 21803066), the Chinese Academy of Sciences Pioneer Hundred Talents Program (KJ2340000031), the National Key Research and Development Program of China (2016YFA0200604), the Anhui Initiative in Quantum Information Technologies (AHY090400), the Strategic Priority Research Program of Chinese Academy of Sciences (XDC01040100), the Fundamental Research Funds for the Central Universities (WK2340000091), the Research Start-Up Grants (KY2340000094), and the Academic Leading Talents Training Program (KY2340000103) from University of Science and Technology of China. The authors thank the Supercomputing Center of Chinese Academy of Sciences, the Supercomputing Center of USTC, the National Supercomputing Center in Kunshan, Wuxi, Shanghai, and Guangzhou for the computational resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.589910/full#supplementary-material

REFERENCES

Banerjee, A. S., Lin, L., Hu, W., Yang, C., and Pask, J. E. (2016). Chebyshev polynomial filtered subspace iteration in the discontinuous Galerkin method for large-scale electronic structure calculations. *J. Chem. Phys.* 145:154101. doi: 10.1063/1.4964861

Blum, V., Gehrke, R., Hanke, F., Havu, P., Havu, V., Ren, X., et al. (2009). *Ab initio* molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* 180, 2175–2196. doi: 10.1016/j.cpc.2009.06.022

Borštnik, U., VandeVondele, J., Weber, V., and Hutter, J. (2014). Sparse matrix multiplication: The distributed block-compressed sparse row library. *Parallel Comput.* 40, 47–58. doi: 10.1016/j.parco.2014.03.012

Bowler, D., Miyazaki, T., and Gillan, M. (2002). Recent progress in linear scaling *ab initio* electronic structure techniques. *J. Phys. Condens. Matter* 14:2781. doi: 10.1088/0953-8984/14/11/303

Bowler, D. R., and Miyazaki, T. (2012). O(N) methods in electronic structure calculations. *Rep. Prog. Phys.* 75:036503. doi: 10.1088/0034-4885/75/3/036503

Cankurtaran, B. O., Gale, J., and Ford, M. (2008). First principles calculations using density matrix divide-and-conquer within the SIESTA methodology. *J. Phys. Condens. Matter* 20:294208. doi: 10.1088/0953-8984/20/29/294208

Challacombe, M. (2014). Linear scaling solution of the time-dependent self-consistent-field equations. *Computation* 2, 1–11. doi: 10.3390/computation2010001

Chen, M., Guo, G., and He, L. (2010). Systematically improvable optimized atomic basis sets for *ab initio* calculations. *J. Phys. Condens. Matter* 22:445501. doi: 10.1088/0953-8984/22/44/445501

Chen, M., Guo, G., and He, L. (2011). Electronic structure interpolation via atomic orbitals. *J. Phys. Condens. Matter* 23:325501. doi: 10.1088/0953-8984/23/32/325501

Cholesky, A.-L. (2005). Sur la résolution numérique des systémes d'équations linéaires. *Bulletin de la Sabix. Société des amis de la Bibliothèque et de l'Histoire de l'École polytechnique* 39, 81–95. doi: 10.4000/sabix.529

Corsetti, F. (2014). The orbital minimization method for electronic structure calculations with finite-range atomic basis sets. *Comput. Phys. Commun.* 185, 873–883. doi: 10.1016/j.cpc.2013.12.008

Daniels, A. D., and Scuseria, G. E. (1999). What is the best alternative to diagonalization of the Hamiltonian in large scale semiempirical calculations? *J. Chem. Phys.* 110, 1321–1328. doi: 10.1063/1.478008

Daw, M. S. (1993). Model for energetics of solids based on the density matrix. *Phys. Rev. B* 47:10895. doi: 10.1103/PhysRevB.47.10895

Dawson, W., and Nakajima, T. (2018). Massively parallel sparse matrix function calculations with NTPoly. *Comput. Phys. Commun.* 225, 154–165. doi: 10.1016/j.cpc.2017.12.010

Etter, S. (2020). Incomplete selected inversion for linear-scaling electronic structure calculations. *arXiv [Preprint]. arXiv:2001.06211*. Available online at: https://arxiv.org/abs/2001.06211

Frisch, M. J., Pople, J. A., and Binkley, J. S. (1984). Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* 80, 3265–3269. doi: 10.1063/1.447079

Galli, G., and Parrinello, M. (1992). Large scale electronic structure calculations. *Phys. Rev. Lett.* 69:3547. doi: 10.1103/PhysRevLett.69.3547

Genovese, L., Neelov, A., Goedecker, S., Deutsch, T., Ghasemi, S. A., Willand, A., et al. (2008). Daubechies wavelets as a basis set for density functional pseudopotential calculations. *J. Chem. Phys.* 129:014109. doi: 10.1063/1.2949547

Gillan, M. J., Bowler, D. R., Torralba, A. S., and Miyazaki, T. (2007). Order-N first-principles calculations with the CONQUEST code. *Comput. Phys. Commun.* 177, 14–18. doi: 10.1016/j.cpc.2007.s02.075

Goedecker, S. (1999). Linear scaling electronic structure methods. *Rev. Mod. Phys.* 71:1085. doi: 10.1103/RevModPhys.71.1085

Goedecker, S., and Colombo, L. (1994). Efficient linear scaling algorithm for tight-binding molecular dynamics. *Phys. Rev. Lett.* 73:122. doi: 10.1103/PhysRevLett.73.122

Goedecker, S., and Teter, M. (1995). Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals. *Phys. Rev. B* 51:9455. doi: 10.1103/PhysRevB.51.9455

Goedecker, S., Teter, M., and Hutter, J. (1996). Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* 54:1703. doi: 10.1103/PhysRevB.54.1703

Ho, G. S., Lignères, V. L., and Carter, E. A. (2008). Introducing PROFESS: a new program for orbital-free density functional theory calculations. *Comput. Phys. Commun.* 179, 839–854. doi: 10.1016/j.cpc.2008.07.002

Hohenberg, P., and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.* 136, B864–B871. doi: 10.1103/PhysRev.136.B864

Hu, W., Lin, L., and Yang, C. (2015a). DGDFT: a massively parallel method for large scale density functional theory calculations. *J. Chem. Phys.* 143:124110. doi: 10.1063/1.4931732

Hu, W., Lin, L., and Yang, C. (2015b). Edge reconstruction in armchair phosphorene nanoribbons revealed by discontinuous Galerkin density functional theory. *Phys. Chem. Chem. Phys.* 17, 31397–31404. doi: 10.1039/C5CP00333D

Kim, J., and Jung, Y. (2016). A perspective on the density matrix purification for linear scaling electronic structure calculations. *Int. J. Quant. Chem.* 116, 563–568. doi: 10.1002/qua.25048

Kim, J., Mauri, F., and Galli, G. (1995). Total-energy global optimizations using nonorthogonal localized orbitals. *Phys. Rev. B* 52:1640. doi: 10.1103/PhysRevB.52.1640

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev*. 140, A1133–A1138. doi: 10.1103/PhysRev.140.A1133

Kühne, T. D., Iannuzzi, M., Del Ben, M., Rybkin, V. V., Seewald, P., Stein, F., et al. (2020). CP2K: an electronic structure and molecular dynamics software package-Quickstep: efficient and accurate electronic structure calculations. *J. Chem. Phys*. 152:194103. doi: 10.1063/5.0007045

Lanczos, C. (1950). *An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators*. United States Government Press Office, Los Angeles, CA. doi: 10.6028/jres.045.026

Li, X.-P., Nunes, R. W., and Vanderbilt, D. (1993). Density-matrix electronic-structure method with linear system-size scaling. *Phys. Rev. B* 47:10891. doi: 10.1103/PhysRevB.47.10891

Liang, W., Saravanan, C., Shao, Y., Baer, R., Bell, A. T., and Head-Gordon, M. (2003). Improved Fermi operator expansion methods for fast electronic structure calculations. *J. Chem. Phys*. 119, 4117–4125. doi: 10.1063/1.1590632

Lin, L., Lu, J., Ying, L., and Weinan, E. (2012). Adaptive local basis set for Kohn-Sham density functional theory in a discontinuous Galerkin framework I: total energy calculation. *J. Com. Phys*. 231, 2140–2154. doi: 10.1016/j.jcp.2011.11.032

Mauri, F., and Galli, G. (1994). Electronic-structure calculations and molecular-dynamics simulations with linear system-size scaling. *Phys. Rev. B* 50:4316. doi: 10.1103/PhysRevB.50.4316

Mohr, S., Ratcliff, L. E., Boulanger, P., Genovese, L., Caliste, D., Deutsch, T., et al. (2014). Daubechies wavelets for linear scaling density functional theory. *J. Chem. Phys*. 140:204110. doi: 10.1063/1.4871876

Mulliken, R. S. (1955). Electronic population analysis on LCAO-MO molecular wave functions. I. *J. Chem. Phys*. 23, 1833–1840. doi: 10.1063/1.1740588

Niklasson, A. M. (2002). Expansion algorithm for the density matrix. *Phys. Rev. B* 66:155115. doi: 10.1103/PhysRevB.66.155115

Niklasson, A. M., Tymczak, C., and Challacombe, M. (2003). Trace resetting density matrix purification in O(N) self-consistent-field theory. *J. Chem. Phys*. 118, 8611–8620. doi: 10.1063/1.1559913

Nunes, R., and Vanderbilt, D. (1994). Generalization of the density-matrix method to a nonorthogonal basis. *Phys. Rev. B* 50:17611. doi: 10.1103/PhysRevB.50.17611

Ordejón, P., Drabold, D. A., Martin, R. M., and Grumbach, M. P. (1995). Linear system-size scaling methods for electronic-structure calculations. *Phys. Rev. B* 51:1456. doi: 10.1103/PhysRevB.51.1456

Ozaki, T., and Kino, H. (2005). Efficient projector expansion for the *ab initio* LCAO method. *Phys. Rev. B* 72:045121. doi: 10.1103/PhysRevB.72.045121

Palser, A. H., and Manolopoulos, D. E. (1998). Canonical purification of the density matrix in electronic-structure theory. *Phys. Rev. B* 58:12704. doi: 10.1103/PhysRevB.58.12704

Pratapa, P. P., Suryanarayana, P., and Pask, J. E. (2016). Spectral quadrature method for accurate O(N) electronic structure calculations of metals and insulators. *Comput. Phys. Commun*. 200, 96–107. doi: 10.1016/j.cpc.2015.11.005

Qin, X., Shang, H., Xiang, H., Li, Z., and Yang, J. (2015). HONPAS: a linear scaling open-source solution for large system simulations. *Int. J. Quant. Chem*. 115, 647–655. doi: 10.1002/qua.24837

Saad, Y. (1994). *PARSEKIT: A Basic Tool Kit for Sparse Matrix Computation (version2)*. Urbana, IL: University of Illinois.

Shang, H., Xiang, H., Li, Z., and Yang, J. (2010). Linear scaling electronic structure calculations with numerical atomic basis set. *Int. Rev. Phys. Chem*. 29, 665–691. doi: 10.1080/0144235X.2010.520454

Skylaris, C.-K., Haynes, P. D., Mostofi, A. A., and Payne, M. C. (2005). Introducing ONETEP: linear-scaling density functional simulations on parallel computers. *J. Chem. Phys*. 122:084119. doi: 10.1063/1.1839852

Soler, J. M., Artacho, E., Gale, J. D., García, A., Junquera, J., Ordejón, P., et al. (2002). The SIESTA method for *ab initio* order-N materials simulation. *J. Phys. Condens. Matter* 14:2745. doi: 10.1088/0953-8984/14/11/302

Suryanarayana, P. (2013). On spectral quadrature for linear-scaling density functional theory. *Chem. Phys. Lett*. 584, 182–187. doi: 10.1016/j.cplett.2013.08.035

Suryanarayana, P. (2017). On nearsightedness in metallic systems for O(N) density functional theory calculations: a case study on aluminum. *Chem. Phys. Lett*. 679, 146–151. doi: 10.1016/j.cplett.2017.04.095

Troullier, N., and Martins, J. L. (1991). Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* 43:1993. doi: 10.1103/PhysRevB.43.1993

VandeVondele, J., Borstnik, U., and Hutter, J. (2012). Linear scaling self-consistent field calculations with millions of atoms in the condensed phase. *J. Chem. Theory Comput*. 8, 3565–3573. doi: 10.1021/ct200897x

Xiang, H., Liang, W., Yang, J., Hou, J., and Zhu, Q. (2005). Spin-unrestricted linear-scaling electronic structure theory and its application to magnetic carbon-doped boron nitride nanotubes. *J. Chem. Phys*. 123:124105. doi: 10.1063/1.2034448

Yang, W. (1991). Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett*. 66:1438. doi: 10.1103/PhysRevLett.66.1438

Yang, W., and Lee, T.-S. (1995). A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules. *J. Chem. Phys*. 103, 5674–5678. doi: 10.1063/1.470549

Zhang, G., Lin, L., Hu, W., Yang, C., and Pask, J. E. (2017). Adaptive local basis set for Kohn-Sham density functional theory in a discontinuous Galerkin framework II: force, vibration, and molecular dynamics calculations. *J. Com. Phys*. 335, 426–443. doi: 10.1016/j.jcp.2016.12.052

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fchem.2020.589795

Crystal Structure Prediction of Binary Alloys via Deep Potential

Haidi Wang[1], *Yuzhi Zhang*[2,3], *Linfeng Zhang*[4]* *and Han Wang*[5]*

[1] School of Electronic Science and Applied Physics, Hefei University of Technology, Hefei, China, [2] Yuanpei College of Peking University, Beijing, China, [3] Beijing Institute of Big Data Research, Peking University, Beijing, China, [4] Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, United States, [5] Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing, China

Predicting crystal structure has been a challenging problem in physics and materials science for a long time. A reliable energy calculation engine combined with an efficient global search algorithm, such as particle swarm optimization algorithm or genetic algorithm, is needed to conduct crystal structure prediction. In recent years, machine learning-based interatomic potential energy surface models have been proposed, potentially allowing us to perform crystal structure prediction for systems with the accuracy of density functional theory (DFT) and the speed of empirical force fields. In this paper, we employ a previously developed Deep Potential model to predict the intermetallic compound of the aluminum–magnesium system, and find six meta-stable phases with negative or nearly zero formation energy. In particular, $Mg_{12}Al_8$ shows excellent ductility and Mg_5Al_{27} has a high Young's modulus. Based on our benchmark results, we propose a relatively robust structure screening criterion that selects potentially stable structures from the Deep Potential-based convex hull and performs DFT refinement. By using this criterion, the computational cost needed to construct the convex hull with *ab initio* accuracy can be dramatically reduced.

Keywords: many-body potential energy, deep learning, crystal structure prediction, Al-Mg, alloy

OPEN ACCESS

Edited by:
Wei Hu,
Lawrence Berkeley National Laboratory, United States

Reviewed by:
Honghui Shang,
Chinese Academy of Sciences (CAS), China
Zp Hu,
Nankai University, China

***Correspondence:**
Linfeng Zhang
linfengz@princeton.edu
Han Wang
wang_han@iapcm.ac.cn

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *31 July 2020*
Accepted: *28 August 2020*
Published: *26 November 2020*

Citation:
Wang H, Zhang Y, Zhang L and Wang H (2020) Crystal Structure Prediction of Binary Alloys via Deep Potential. Front. Chem. 8:589795. doi: 10.3389/fchem.2020.589795

INTRODUCTION

In recent years, crystal structure prediction has played an increasingly important role, not only for understanding the ground-state structure of matter, but also for designing materials and drug molecules with desired functionality (Oganov, 2018; He et al., 2019; Zhao et al., 2019; Xie et al., 2020). Generally speaking, a ground-state crystal structure prediction method involves three components: a model that generates the interatomic potential energy surface (PES) and forces, a sampling technique for exploring different conformations in the configuration space, and a relaxation procedure to find the local minima on the PES (Podryabinkin et al., 2019). While the relaxation procedure is relatively standard, different sampling techniques have been championed by different software packages. For instance, the genetic evolutionary algorithm, the particle swarm algorithm, and the firefly algorithm have been widely used in the USPEX software (Glass et al., 2006), the CALYPSO package (Wang et al., 2010), and the PyChemia library (Avendaño-Franco and Romero, 2016), respectively. Due to the high accuracy required by both sampling and relaxation, density functional theory (DFT) (Kohn and Sham, 1965) is typically used for generating the PES. Despite its widespread success, DFT has a high computational cost that typically scales cubically with the system size, which, to some extent, hinders routine applications to large and complex systems.

Many empirical PES models for popular solid-state systems have been proposed (Mendelev et al., 2009; Jelinek et al., 2012; Dickel et al., 2018) to address the efficiency issue of DFT. Due to the relatively simple and analytical expressions adopted by these empirical models, an acceleration of many orders of magnitude in terms of computational cost can be gained, but presumably at the price of accuracy and transferability. As such, trail-and-error processes are typically required for developing such models, yet challenges have remained for systems involving multiple elements, complex and exotic phases, or bond breaking and formation events.

In recent years, a few machine learning (ML) techniques have been proposed for representing the PES (Behler and Parrinello, 2007; Bartók et al., 2010; Artrith and Urban, 2016; Khorshidi and Peterson, 2016; Shapeev, 2016; Han et al., 2018; Zhang et al., 2018a,b). Unlike typically empirical PES models, representations coming from ML tasks, such as kernel functions and neural networks (NNs), have shown great promise to fit high-dimensional functions. When trained on a suitably generated dataset of atomic configurations and corresponding potential energies and forces, a good ML-based PES model can be used with an accuracy of the reference DFT model, and an efficiency comparable to that of empirical PES models. Not surprisingly, ML-based PES models have been employed in recent work for structure search tasks. For instance, boron has been studied by several groups: Podryabinkin et al. (2019) adopted the moment tensor potential (Shapeev, 2016) and the USPEX evolutionary algorithm; Huang et al. (2018) used the Behler–Parrinello potential (Behler and Parrinello, 2007) and the stochastic surface walking global optimization method; Tong et al. (2018) used the Gaussian Approximation Potential (Bartók et al., 2010) and the CALYPSO approach.

In this work, we target at using ML-based PES models for crystal structure prediction of alloys. We adopt the smooth version of the Deep Potential (DP) model (Zhang et al., 2018b), which employs NN architectures to parameterize two networks, the embedding network that defines a list of symmetry-preserving descriptors, and the fitting network that maps these descriptors to local energy contributions. The versatile architecture of DP makes it particularly suitable for multicomponent systems and those involving bond breaking and formation, for which most methodologies are hard to handle. The aluminum–magnesium (Al–Mg) binary alloy system is selected as an example based on the following reasons: First, Al–Mg binary alloys are important in real-life applications. They are widely used in automotive, aerospace, and electronic device industries (Gupta and Ling, 2011) due to their lightweight nature and excellent mechanical properties. However, only a limited number of intermetallic compounds of the Al–Mg binary system have been documented in well-known databases, such as the American Society for Metals (ASM) Alloy Phase Diagram Database[1], the Inorganic Crystal Structure Database (ICSD)[2], the Open Quantum Materials Database (OQMD)[3], and the Material Project database (MP)[4]. Second, our previous study has established an Al–Mg DP model (Zhang et al., 2019), which describes well the basic physicochemical properties and has been carefully tested. As such, this DP model can be readily used for crystal structure prediction and can be download online[5].

Combining the particle swarm optimization (PSO) method and the DP model, potential intermetallic compounds of the Al–Mg system are systematically explored. Compared with a previous study (Zhuang et al., 2017), which only explored the Mg-rich phases, our simulation covers a much wider concentration range. Six new Al–Mg intermetallic compounds ($Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27}) are found to be meta-stable. The mechanical properties of these new compounds are further investigated. To facilitate future investigations of more complicated tasks, special attention is given to the whole simulation protocol and the selection criterion for further DFT validations. Direct comparisons with popular empirical PES models and DFT show the advantage of DP in terms of both accuracy and efficiency.

COMPUTATIONAL METHODS

We adopt the PSO method, as implemented in the CALYPSO package (Wang et al., 2010), to search potentially stable and meta-stable Mg–Al intermetallic structures. PSO is inspired by the choreography of a bird flock and can be seen as a distributed behavior algorithm that performs multidimensional search. In the CALYPSO package, there are three steps for a global structure prediction task. First, a group of structures called population is generated randomly with symmetric constraints to allow a diverse sampling of the PES. The number of structures employed here is defined by a parameter called population size (`PopSize`). Second, a local relaxation of the population is performed based on a PES engine, which is typically a DFT model, and here we replace it with a DP model. A procedure that eliminates similar structures by using the so-called bond characterization matrix is followed up to enhance the search efficiency. Third, a certain number of new structures (the best 60% of the population size) are generated by PSO. Within the PSO scheme, a velocity vector associated with each structure is updated using the information of the previously proposed and optimized structure, as well as the globally best structure, that is, the structure with the lowest enthalpy, at the current generation. The new structures are generated based on the current structures and the velocity vectors. The last two steps continue iteratively until the predefined largest number of generations (`GenNumb`) is reached. The parameter `GenNumb` is typically selected to be large enough so that the structure with the lowest energy can sustain for several generations. Generally speaking, the more atoms (`Natom`) in a structure, the larger `PopSize` and `GenNumb` are required. We refer to Yanchao Wang and Ma (2012) for more details of the CALYPSO

[1]ASM https://www.asminternational.org/phase-diagrams
[2]ICSD https://icsd.fiz-karlsruhe.de/search/basic.xhtml
[3]OQMD http://www.oqmd.org
[4]MP https://www.materialsproject.org
[5]deepmd http://www.deepmd.org

code. We set the CALYPSO parameters according to the following criteria:

$$\begin{cases} \texttt{PopSize} = 30, & \texttt{GenNumb} = 60; & \text{if} \quad \texttt{Natom} \leq 10, \\ \texttt{PopSize} = 40, & \texttt{GenNumb} = 80; & \text{if } 10 < \texttt{Natom} \leq 20, \\ \texttt{PopSize} = 50, & \texttt{GenNumb} = 100; & \text{if } 20 < \texttt{Natom} \leq 32. \end{cases} \tag{1}$$

The DP model (Zhang et al., 2018b) used here employs NN functions to represent the PES. In short, the total energy of a system is described as the sum of atomic energies,

$$E = \sum_{n=1}^{N} \epsilon_i, \tag{2}$$

where ϵ_i is the ith atomic energy. The atomic energy is represented as

$$\epsilon_i = N_{\omega(i)}(\{R_{ij}\}_{j \in \mathcal{N}(i)}), \tag{3}$$

where $N_{\omega(i)}$ is called a sub-network that computes the atomic contribution to the total energy, and $\omega(i)$, which depends on the chemical species of atom i, denotes the weights used to parameterize the sub-network. The neighbors of atom i within the cut-off radius R_c are denoted by $\mathcal{N}(i)$. R_{ij} is the position of atom j relative to i used to describe the local environment of atom i. To generate uniformly accurate DP models in a way that minimizes both human intervention and the computational cost for data generation and model training, a concurrent learning strategy called the Deep Potential GENerator (DP-GEN) (Zhang et al., 2019) has been proposed. In this strategy, an initial dataset (random Al–Mg alloy structures) labeled by DFT calculations is used to train an ensemble of DP models, and molecular dynamics is driven by one of the DP models to sample the configuration space. An error indicator serves to select a small fraction out of the new samples as candidates, which are labeled with *ab initio* energies and forces and added to the database. Such iterations are repeated until the configuration space has been explored sufficiently, and a decent DP model has been obtained with high accuracy and transferability. The training is performed using the DeePMD-kit package (Wang et al., 2018) and the concurrent learning strategy is realized by the DP-GEN software package (Zhang et al., 2020). In details, the DeepPot-SE model is used with a cutoff radius of 8.0 Å. The size of the embedding and fitting NNs are $25 \times 50 \times 100$ and $240 \times 240 \times 240$, respectively. During the training, the learning rate decreases exponentially with respect to the starting value of 0.0005. The decay rate and decay step are set to 0.95 and 128,000, respectively. In addition, the prefactors of loss functions are set to $p_e^{start} = 0.02, p_e^{limit} = 2, p_f^{start} = 1000, p_f^{limit} = 1, p_v^{start} = 0.0, p_v^{limit} = 0.0$. Both DeePMD-kit and DP-GEN are publicly available online[6]. For more details, we refer the reader to Wang et al. (2018) and Zhang et al. (2019, 2020).

All DFT calculations are carried out with the Vienna Ab-Initio Simulation Package (VASP, version 5.4.4) (Kresse and Furthmüller, 1996). The generalized gradient approximation within the Perdew–Burke–Ernzerhof (Perdew et al., 1996) (PBE) functional is used to model the exchange-correlation energy. The plane wave basis sets with kinetic energy cutoff of 600 eV are used to expand the valence electron wave functions. For all structural relaxations, the convergence criterion for the energy in electronic SCF iterations and the Hellmann–Feynman forces in ionic step iterations are set to 1.0×10^{-6} eV and 1.0×10^{-2} eV/Å, respectively. The Brillouin zone is represented by Monkhorst–Pack (Pack and Monkhorst, 1977) special k-point mesh with a grid spacing of 0.08 $Å^{-1}$. The phonon spectra are obtained based on finite-difference method as implemented in the Atomic Simulation Environment (ASE) (Bahn and Jacobsen, 2002; Larsen et al., 2017) software, where the forces are calculated by the python interface of DeePMD-kit. To calculate the phonon density of states, the q-point mesh is set to $20 \times 20 \times 20$. The local structure relaxation is carried out by the LAMMPS package (Plimpton, 1995), and the DP model used here has been reported and extensively tested in Zhang et al. (2019).

All structure data and convex hulls are analyzed by pymatgen software (Ong et al., 2013).

RESULTS AND DISCUSSIONS

To demonstrate the validity of CALYPSO+DP scheme, we perform some preliminary tests for several different stoichiometric proportions. Here, we take $Mg_{12}Al_8$ as an example to show the evolution of the energies of all the structures (**Figure 1A**), as well as the lowest energy (**Figure 1C**), during the CALYPSO structure search process. According to the energy histogram in **Figure 1B**, it can be found that there are about 458 structures, out of a total number of 3,200 structures, within an energy range <20 meV/atom (compared with the ground state structure). The evolution of the lowest energy for all generations shows that the global optimization converged quickly. It is remarkable to find that one potentially stable structures can be readily obtained in the first few generations (labeled by red star in **Figure 1A**). When refined by DFT, it shows that this structure is the ground state structure of the corresponding combination. In the tests of some other stoichiometric proportions, that is, Mg_3Al_3, Mg_2Al_2, Mg_1Al_1, and Mg_3Al, the corresponding known structures in the materials project database, that is, mp-1038779, mp-1094987, mp-1038934, and the $L1_2$ phase (Mendelev et al., 2009), are found by the CALYPSO+DP scheme, which further confirms the validity of our approach.

Next, we use the DP+CALYPSO scheme to construct the convex hull of the Al-Mg system systematically. We restrict the number of atom in the supercell to <32 atoms. In other words, we consider the systems $Mg_xAl_y\{x + y \leq 32, x \geq 1, y \geq 1, x, y \in \mathbb{Z}\}$, which represent 496 combinations, or 323 concentrations, in total. According to these parameter settings, the total number of local relaxations is up to 2×10^7. In the following, we consider three prerequisites that should be satisfied for a stable crystal structure: (i) thermodynamic stability, which is estimated by the formation energy and convex hull; (ii) dynamic stability, which

[6]See https://github.com/deepmodeling

can be assessed by phonon dispersions; and (iii) mechanical stability, which is evaluated via elastic constants (Xu et al., 2019).

According to the preliminary tests, we notice that although the DP model can generate energies and forces that are very close to the DFT reference model, small intrinsic error still exists. Therefore, if our goal is to calculate properties with the accuracy of the DFT-based PES landscape, an additional refinement step has to be adopted based on structures selected from a DP+CALYPSO process. Two concepts we pay particular attention to are the formation energy (E_{fa}) and the energy above convex hull (E_{abh}). The formation energy (Haastrup et al., 2018) of an alloy system is the energy required to produce the system from the most stable crystal structures of the individual components, which is defined as

$$E_{fa} = \frac{E(\mathrm{Mg_xAl_y}) - xE(\mathrm{Mg}) - yE(\mathrm{Al})}{x + y} \tag{4}$$

where $E(\mathrm{Mg_xAl_y})$ is the total energy of the material $\mathrm{Mg_xAl_y}$, and E(Mg) and E(Al) are the average energies of the elements Mg and Al in their stable crystal at 0 K. E_{abh} measures the energy for a material to decompose into the set of most stable materials with the same chemical composition. A positive E_{abh} indicates that this material is unstable with respect to such decomposition. A zero E_{abh} indicates that this is the most stable material at its composition. To accurately determine these properties, we use two criteria for an additional DFT refinement: E_{fa} being less than 20 meV/atom, and E_{abh} being <20 meV/atom. We use these two criteria at the same time based on the following reasons. Our goal is to find potential stable or meta-stable structures. Due to the error of the DP model, some structures with positive DP-predicted E_{fa} may turn negative if we refine it by DFT and vice versa. At the same time, the thermodynamic stability is also controlled by E_{abh}. If E_{abh} is too high, this structure will decompose into other phases even if this structure has a negative formation energy. In the tested example, we will show that since in general DP exhibits a ∼2 meV/atom average error. Compared with DFT, the criteria used here are fairly robust. In contrast, for a previously established empirical model, due to its large intrinsic error, the procedure introduced above is no longer practical, since the number of DFT refinements is so large that little efficiency can be gained.

As shown in **Figure 2A**, we first use the formation energy E_{fa} to screen the candidate structures, from which the number of structures is significantly reduced from 2×10^7 to 5,169. Based

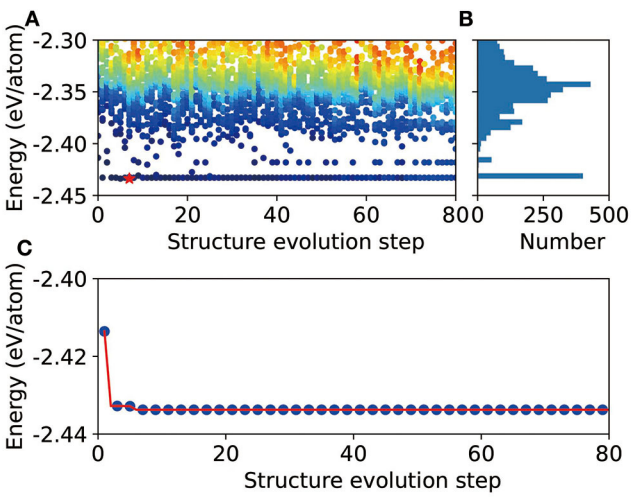

FIGURE 1 | Schematic illustration of the DP+CALYPSO process for the $Mg_{12}Al_8$ system. **(A)** Evolution of the energies of all structures during the structure prediction process. The red star indicates the global minimal structure. **(B)** Energy histogram. **(C)** Evolution of the lowest energy during the structure prediction process.

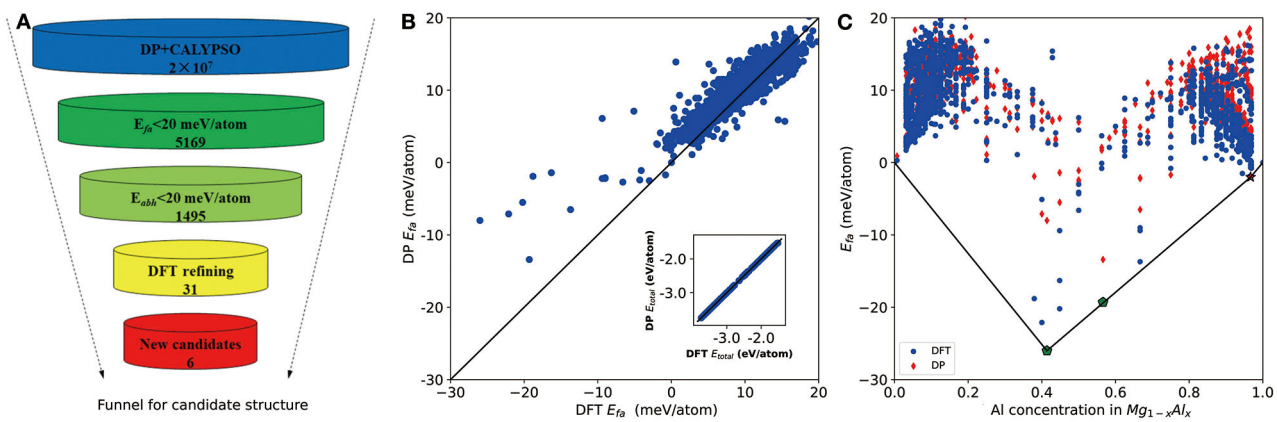

FIGURE 2 | (A) The funnel used for screening candidate structures of aluminum–magnesium (Al–Mg), the intermetallic compound via Deep Potential (DP) model. E_{fa} stands for formation energy and E_{abh} labels energy above the convex hull. **(B)** Scatter plot of the formation energy calculated by density functional theory (DFT) and DP for potentially candidate structures. The inset shows relationship between average energy calculated by DFT and DP. **(C)** The formation energy as a function of molar fraction of Al atom for different Mg–Al phases where solid line denotes the convex hull constructed by DFT results. The formation energies calculated by DFT are marked by blue circle and DP ones are marked by red diamond. All of known structures from materials project are re-optimized and directly used to construct the convex hull. Pentagon indicates the stable experimental phases and star labels new stable phase.

on these 5,169 DP-optimized structures, a crude convex hull is constructed. Then the criterion for E_{abh} is introduced to remove the thermodynamically unstable structures, and, finally, 1,495 structures are selected for further DFT refinements. As shown in the inset of **Figure 2B**, for all DFT refined structures, the energies per atom calculated by DP and DFT are in excellent agreement. The largest and root mean square error (RMSE) of the total energy per atom are about 17 and 2 meV/atom, respectively. As for the formation energy in **Figure 2B**, the largest error and RMSE are about 15 and 6 meV/atom, respectively. Among them, the errors of the formation energies of experimentally stable phases (Zhuang et al., 2017) $Mg_{17}Al_{12}$ and $Mg_{23}Al_{30}$ (labeled by pentagon in **Figure 2C**) are 15 and 6 meV/atom, which confirms the validity of our 20 meV/atom selection criteria.

The convex hull based on DFT results is then constructed and presented in **Figure 2C**, including two experimentally stable structures $Mg_{17}Al_{12}$ and $Mg_{23}Al_{30}$ labeled by green pentagon. One new stable structure with a formula of $MgAl_{29}$ is discovered and denoted by red star. Based on the DFT-refined convex hull, we look for new structures that are potentially synthesizable by experiments. We use the following criteria: $E_{abh} < 20$ meV/atom and $E_{fa} < 1$ meV/atom, where E_{abh} and E_{fa} are DFT-calculated values, and obtain 31 potentially candidates, including 1 stable structure and 30 meta-stable structures. However, we may note that most of those newly proposed meta-stable structures locate at the boundary region of the concentration range. That is to say, most of these structures have very low concentration of Al or Mg. Compared with these phases, the phases in the middle region are of more interesting, from which we propose six new meta-stable structures, namely $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27}. The corresponding side view of the crystal structures are shown in **Figures 3A–F** and the geometric structure parameters are listed in **Table 1**.

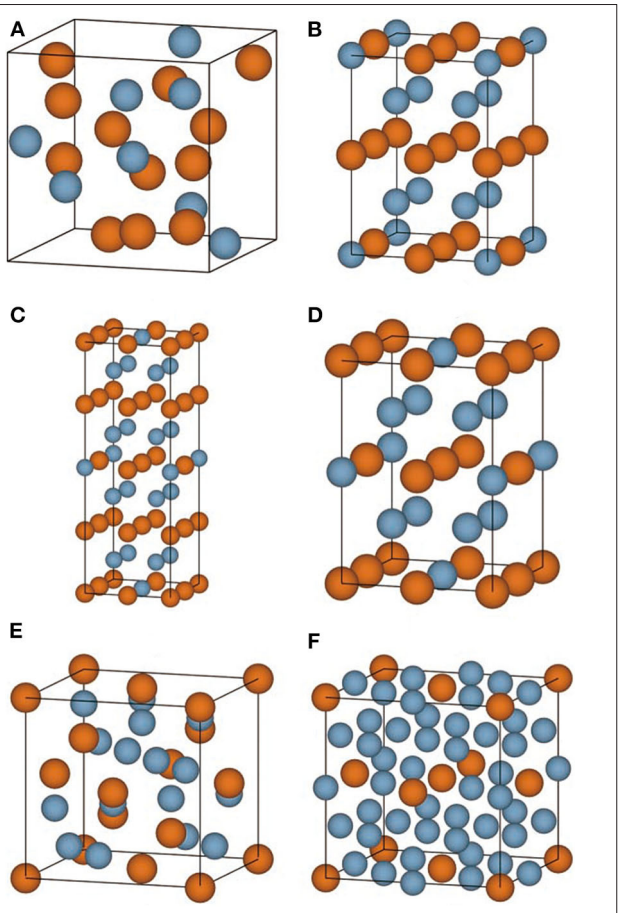

FIGURE 3 | (A–F) The side view of conventional crystal structures of $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27} is shown, respectively. The Mg and Al atoms are indicated by yellow and blue ball.

As listed in **Table 1**, the meta-stable structures can be divided into two groups according to their lattice types. Mg_7Al_9, $Mg_{14}Al_{18}$, and Mg_6Al_{10} have a tetragonal lattice, whereas $Mg_{12}Al_8$, Mg_8Al_{16}, and Mg_5Al_{27} have a cubic lattice. Moreover, all structures have nearly zero or negative formation and small energy above convex hull, which indicates that these structures may be synthesizable by experiments in future.

Given the encouraging stability metrics above, we proceed to study the dynamic and mechanical stability of these newly proposed intermetallic compounds via DP model. As shown in **Figure 4**, the phonon structures show no imaginary frequency, which indicates that all of those intermetallic compounds are dynamically stable. As for the mechanical aspect, the elastic stability conditions (Mouhat and Coudert, 2014) for cubic crystals are given as:

$$C_{11} - C_{12} > 0,\ C_{11} + 2C_{12} > 0,\ C_{44} > 0, \tag{5}$$

and those for tetragonal crystals are given as:

$$C_{11} > |C_{12}|,\ 2C_{13}^2 < C_{33}(C_{11} + 2C_{12}),\ C_{44} > 0,\ C_{66} > 0. \tag{6}$$

According to **Table 2**, both groups of structures meet the elastic stability conditions, which indicates that these 6 intermetallic compounds are mechanically stable. Further, we use the Pugh's ratio B_v/G_v to assess the expected average ductility (Pugh, 1954). According to Pugh, a larger B_v/G_v value implies a better ductility property. As shown in **Table 2**, the Pugh's ratio of both Mg_8Al_{16} and $Mg_{12}Al_8$ are larger than that of hcp Mg (2.08). In particular, the Pugh's ratio of $Mg_{12}Al_8$ (2.42) is comparable to that of Al (2.47), so it may have excellent ductility. Considering its

TABLE 1 | Lattice parameters a(Å), b(Å), c(Å), density ρ (g/cm^3), space group symbol S_m, lattice type, formation energy E_{fa} (meV/atom), and energy above the convex hull E_{abh}(meV/atom) of $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27} calculated by DFT.

Formula	N_{atom}	a	b	c	ρ	S_m	Lattice type	E_{fa}	E_{ah}
$Mg_{12}Al_8$	20	7.38	7.38	7.38	2.10	$P4_{332}$	Cubic	−22.05	3.09
Mg_7Al_9	16	5.98	5.98	8.44	2.27	P4/mmm	Tetragonal	−0.82	18.74
$Mg_{14}Al_{18}$	32	5.98	5.98	16.92	2.27	I4/mmm	Tetragonal	−0.10	19.45
Mg_6Al_{10}	16	5.93	5.93	8.43	2.33	I4/mmm	Tetragonal	0.87	17.71
Mg_8Al_{16}	24	7.67	7.67	7.67	2.30	Fd-3m	Cubic	−13.73	1.30
Mg_5Al_{27}	32	8.21	8.21	8.21	2.55	Pm-3m	Cubic	0.60	7.95

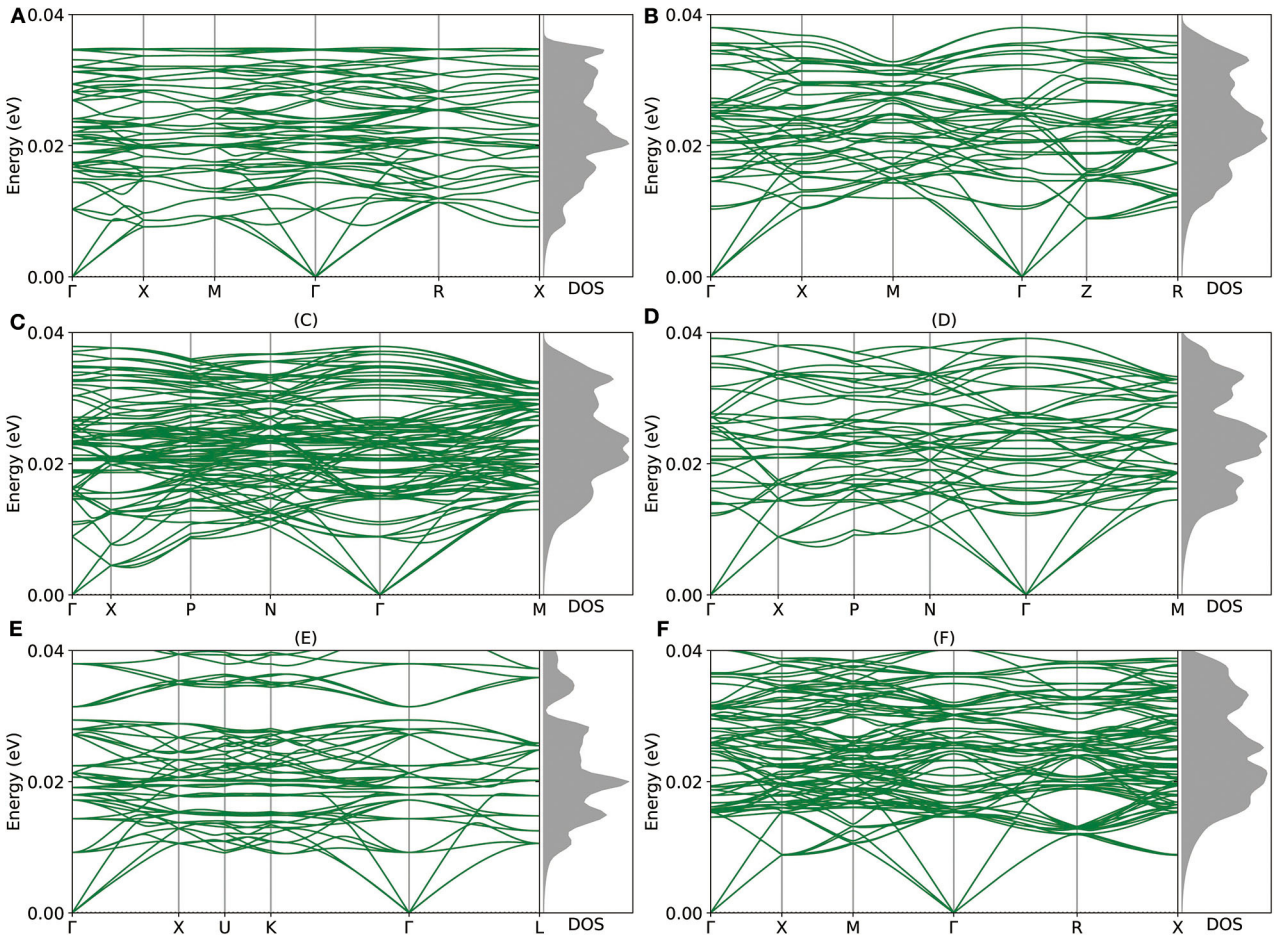

FIGURE 4 | (A–F) The phonon structures of $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27} calculated by Deep Potential (DP) model are shown.

TABLE 2 | Elastic constants, bulk modulus B_v (GPa), shear modulus G_v (GPa), Young's modulus E_v (GPa), Pugh's ratio (B_v/G_v), and Poisson's ratio ν of $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16} and Mg_5Al_{27} calculated by the DP model.

Formula	C_{11}	C_{12}	C_{13}	C_{33}	C_{44}	C_{66}	B_v	G_v	B_v/G_v	E_V	ν
Mg_9[a,1]	59.57	26.54	21.05	72.94	15.08	16.12	36.53	17.59	2.08	45.48	0.29
$Mg_{12}Al_8$	71.26	35.75	36.50	73.15	21.19	21.38	48.14	19.89	2.42	52.45	0.32
Mg_7Al_9	94.75	30.55	39.20	84.63	32.91	22.00	53.81	28.08	1.92	71.77	0.28
$Mg_{14}Al_{18}$	89.19	31.60	40.27	85.57	32.16	23.48	54.57	27.49	1.99	70.60	0.28
Mg_6Al_{10}	101.77	36.06	42.59	92.14	31.37	23.23	59.65	29.05	2.05	74.97	0.29
Mg_8Al_{16}	97.47	49.80	48.72	95.07	30.68	31.73	65.06	28.35	2.29	74.27	0.31
Mg_5Al_{27}	95.40	39.17	39.87	104.71	37.48	40.21	58.91	35.09	1.68	87.84	0.25
Al_4[a,2]	117.64	63.34	58.20	108.46	32.55	40.44	78.15	31.65	2.47	83.66	0.32

The subscript v denotes the Voigt expressions. The same properties of Mg and Al element are also calculated for comparison (here, the minor inconsistency of elastic constants comes from inherent error of DP model and computation error).

[a]*All values are calculated by authors based on DP model.*

[1]*Zhuang et al. (2017) give DFT values of C_{11} = 66 GPa, C_{12} =25 GPa, C_{13} = 19 GPa, C_{33} = 70 GPa, C_{44} = 20 GPa, B_v = 37 GPa, G_v = 21 GPa, B_v/G_v = 1.76 GPa, and E_V = 54 GPa.*

[2]*Zhang et al. (2019) give DFT values of C_{11} = 111.2 GPa, C_{12} = 61.4 GPa, C_{44} = 36.8 GPa, B_v = 78.0 GPa, and G_v = 32.1 GPa.*

lower density (2.10 g/cm^3) compared with Al (2.72 g/cm^3), this intermetallic compound may have potential applications in automotive, aerospace, electronic, and device industries if it can be synthesized. In addition, the Mg_5Al_{27} has a higher Young's modulus among all these structures, which indicates that this material may be applied to manufacture high strength devices.

Finally, we test the accuracy of a recent version of the MEAM potential (Jelinek et al., 2012) and see whether it can be used for a similar task or not. For a direct comparison, we test it on all DFT-refined structures and report the results in **Figure 5**. As shown by the red diamonds and green pentagons in **Figure 5A**, MEAM exhibits much larger errors compared with DP for most of the structures, and there are MEAM predictions outside the range of the plot (±100 meV/atom) due to large errors. MEAM results show a large spread on the convex hull plot constructed by DFT results (**Figure 5B**). For the tested structures, the largest error of E_{fa} is 204 meV/atom, and the RMSE is 44 meV/atom. Moreover, the largest error of per-atom total energy is 465 meV/atom, and the mean error is 236 meV/atom. As such, if we use MEAM+CALYPSO to do a screening of the structures, the selection criteria for further

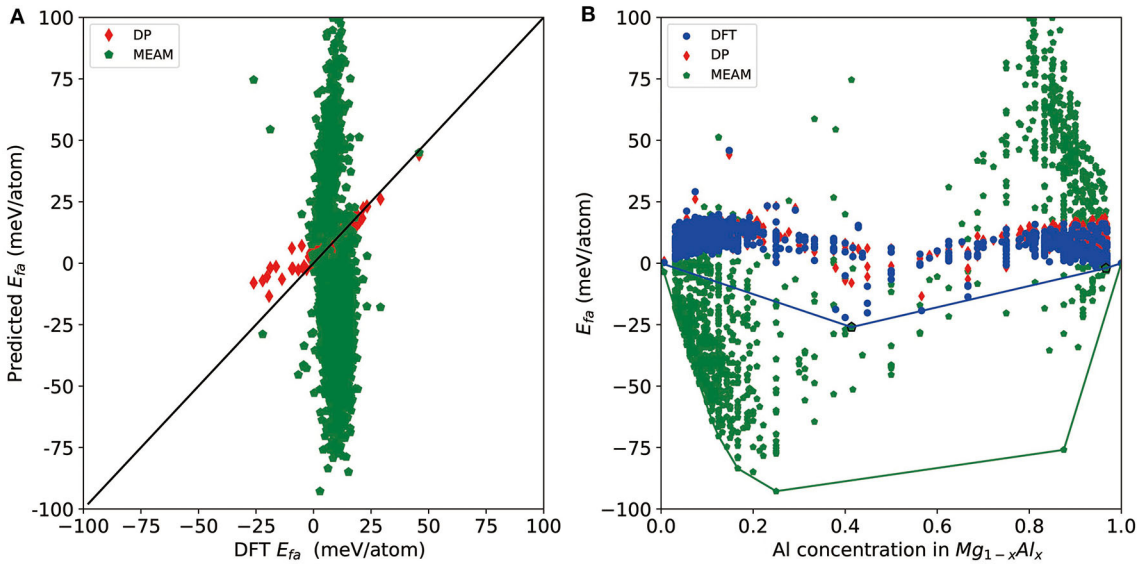

FIGURE 5 | (A) Scatter plot of the formation energy calculated by density functional theory (DFT) and Deep Potential (DP) for potentially candidate structures. **(B)** The formation energy as a function of molar fraction of Al atom for different Mg–Al phases where solid line denotes the convex hull constructed by DFT results. The formation energies calculated by DFT, DP, and MEAM are marked by blue circle, red diamond, and green pentagon, respectively.

DFT refinement would be E_{fa} and E_{abh} larger than at least 200 meV/atom. As a rough estimation, according to these criteria, $\sim 1 \times 10^6$ structures will need to be refined by DFT, which is not computationally feasible at all. Above all, we conclude that although MEAM is very efficient, it cannot be used to improve the efficiency of constructing a convex hull at the level of DFT accuracy.

To demonstrate the efficiency of our DP-based procedure, we use two groups of structures to compare the time performance. One group is composed of several $Mg_{31}Al$ structures and the other group consists of $MgAl_{31}$ structures. The test results shows that, compared with DFT relaxation, DP has an average speed-up ratio about 3,700 and 650 for $Mg_{31}Al$ and $MgAl_{31}$, respectively, which indicates DP has better time scaling and can be used for larger systems. All tests are performed on Intel(R) Xeon(R) Gold 6248 CPU @ 2.50 GHz.

CONCLUSIONS

In this paper, we demonstrate that the DP+CALYPSO scheme is reliable for crystal structure prediction for binary alloy system in a wide concentration range. As a concrete example, we use this scheme to predict potentially stable intermetallic compounds of the Al–Mg binary system. Six new meta-stable Al–Mg intermetallic compounds are successfully predicted, including $Mg_{12}Al_8$, Mg_7Al_9, $Mg_{14}Al_{18}$, Mg_6Al_{10}, Mg_8Al_{16}, and Mg_5Al_{27}. All the meta-stable structures are predicted to have thermodynamic stability, dynamic stability, and mechanical stability. In particular, $Mg_{12}Al_8$ shows excellent ductility and Mg_5Al_{27} has high Young's modulus. We remark that the exploration strategy proposed in this work can be combined with the DP-GEN protocol (Zhang et al., 2019, 2020) to generate more training data and improve the DP potential. Moreover, to serve a larger community, DeePMD-kit can be interfaced with other popular general-purpose crystal structure prediction software such as CALYPSO, USPEX, and Pychemia. However, some disadvantages also exits for current scheme, such as expensive cost for training a model, low interface efficiency with CALYPSO, and so on, which limits its application to search complex multicomponent systems with larger number of atoms. We will leave these problems in our future work.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

HaiW, HanW, and LZ conceived the idea. HaiW performed the calculations and wrote the script. All authors revised the manuscript.

FUNDING

The work of HaiW was supported by the Fundamental Research Funds for the Central Universities. The work of LZ was supported in part by the Center Chemistry in Solution and at Interfaces (CSI) funded by the DOE Award DE-SC0019394. The work of HanW was supported by the National Science Foundation of China under Grant No. 11871110, the National Key Research and Development

Program of China under Grants Nos. 2016YFB0201200 and 2016YFB0201203, and Beijing Academy of Artificial Intelligence (BAAI).

ACKNOWLEDGMENTS

We are grateful for computing resources provided by the High-performance Computing Platform of Peking University, and the Beijing Institute of Big Data Research, University of Science and Technology of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.589795/full#supplementary-material

REFERENCES

Artrith, N., and Urban, A. (2016). An implementation of artificial neural-network potentials for atomistic materials simulations: performance for TiO_2. *Comput. Mater. Sci.* 114, 135–150. doi: 10.1016/j.commatsci.2015.11.047

Avenda no-Franco, G., and Romero, A. H. (2016). Firefly algorithm for structural search. *J. Chem. Theory Comput*. 12, 3416–3428. doi: 10.1021/acs.jctc.5b01157

Bahn, S. R., and Jacobsen, K. W. (2002). An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* 4, 56–66. doi: 10.1109/5992.998641

Bartók, A. P., Payne, M. C., Kondor, R., and Csányi, G. (2010). Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett*. 104:136403. doi: 10.1103/PhysRevLett.104.136403

Behler, J., and Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett*. 98:146401. doi: 10.1103/PhysRevLett.98.146401

Dickel, D. E., Baskes, M. I., Aslam, I., and Barrett, C. D. (2018). New interatomic potential for Mg-Al-Zn alloys with specific application to dilute Mg-based alloys. *Model. Simul. Mater. Sci. Eng.* 26:045010. doi: 10.1088/1361-651X/aabaad

Glass, C. W., Oganov, A. R., and Hansen, N. (2006). USPEX-Evolutionary crystal structure prediction. *Comput. Phys. Commun*. 175, 713–720. doi: 10.1016/j.cpc.2006.07.020

Gupta, M., and Ling, S. N. M. (2011). *Magnesium, Magnesium Alloys, and Magnesium Composites*. Singapore: John Wiley & Sons. doi: 10.1002/9780470905098

Haastrup, S., Strange, M., Pandey, M., Deilmann, T., Schmidt, P. S., Hinsche, N. F., et al. (2018). The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater*. 5:042002. doi: 10.1088/2053-1583/aacfc1

Han, J., Zhang, L., Car, R., and E, W. (2018). Deep potential: a general representation of a many-body potential energy surface. *Commun. Comput. Phys*. 23, 629–639. doi: 10.4208/cicp.OA-2017-0213

He, J., Xia, Y., Naghavi, S. S., Ozolins, V., and Wolverton, C. (2019). Designing chemical analogs to PbTe with intrinsic high band degeneracy and low lattice thermal conductivity. *Nat. Commun*. 10, 1–8. doi: 10.1038/s41467-019-08542-1

Huang, S. D., Shang, C., Kang, P. L., and Liu, Z. P. (2018). Atomic structure of boron resolved using machine learning and global sampling. *Chem. Sci*. 9, 8644–8655. doi: 10.1039/C8SC03427C

Jelinek, B., Groh, S., Horstemeyer, M. F., Houze, J., Kim, S. G., Wagner, G. J., et al. (2012). Modified embedded atom method potential for Al, Si, Mg, Cu, and Fe alloys. *Phys. Rev. B* 85:245102. doi: 10.1103/PhysRevB.85.245102

Khorshidi, A., and Peterson, A. A. (2016). AMP: a modular approach to machine learning in atomistic simulations. *Comput. Phys. Commun*. 207, 310–324. doi: 10.1016/j.cpc.2016.05.010

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev*. 140:A1133. doi: 10.1103/PhysRev.140.A1133

Kresse, G., and Furthmüller, J. (1996). Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Phys. Rev. B* 6, 15–50. doi: 10.1016/0927-0256(96)00008-0

Larsen, A. H., Mortensen, J. J., Blomqvist, J., Castelli, I. E., Christensen, R., Dułak, M., et al. (2017). The atomic simulation environment-a python library for working with atoms. *J. Phys*. 29:273002. doi: 10.1088/1361-648X/aa680e

Mendelev, M. I., Asta, M., Rahman, M. J., and Hoyt, J. J. (2009). Development of interatomic potentials appropriate for simulation of solid-liquid interface properties in Al-Mg alloys. *Philos. Mag*. 89, 3269–3285. doi: 10.1080/14786430903260727

Mouhat, F., and Coudert, F. X. (2014). Necessary and sufficient elastic stability conditions in various crystal systems. *Phys. Rev. B* 90:224104. doi: 10.1103/PhysRevB.90.224104

Oganov, A. R. (2018). Crystal structure prediction: reflections on present status and challenges. *Faraday Discuss*. 211, 643–660. doi: 10.1039/C8FD90033G

Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., et al. (2013). Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci*. 68, 314–319. doi: 10.1016/j.commatsci.2012.10.028

Pack, J. D., and Monkhorst, H. J. (1977). Special points for Brillouin-zone integrations. *Phys. Rev. B* 16, 1748–1749. doi: 10.1103/PhysRevB.16.1748

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett*. 77, 3865–3868. doi: 10.1103/PhysRevLett.77.3865

Plimpton, S. (1995). Fast parallel algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys*. 117, 1–19. doi: 10.1006/jcph.1995.1039

Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V., and Oganov, A. R. (2019). Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys. Rev. B* 99:064114. doi: 10.1103/PhysRevB.99.064114

Pugh, S. (1954). XCII. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals. *Lond. Edinburgh Dubl. Philos. Mag. J. Sci*. 45, 823–843. doi: 10.1080/14786440808520496

Shapeev, A. V. (2016). Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multisc. Model. Simul*. 14, 1153–1173. doi: 10.1137/15M1054183

Tong, Q., Xue, L., Lv, J., Wang, Y., and Ma, Y. (2018). Accelerating CALYPSO structure prediction by data-driven learning of a potential energy surface. *Faraday Discuss*. 211, 31–43. doi: 10.1039/C8FD00055G

Wang, H., Zhang, L., Han, J., and E, W. (2018). DeePMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun*. 228, 178–184. doi: 10.1016/j.cpc.2018.03.016

Wang, Y., Lv, J., Zhu, L., and Ma, Y. (2010). Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B* 82:094116. doi: 10.1103/PhysRevB.82.094116

Wang, Y., Lv, J., Zhu, L., and Ma, Y. (2012). CALYPSO: A method for crystal structure prediction. *Comput. Phys. Commun*. 183, 2063–2070. doi: 10.1016/j.cpc.2012.05.008

Xie, H., Zhang, W., Duan, D., Huang, X., Huang, Y., Song, H., et al. (2020). Superconducting zirconium polyhydrides at moderate pressures. *J. Phys. Chem. Lett*. 11, 646–651. doi: 10.1021/acs.jpclett.9b03632

Xu, W. W., Yu, Q., Xia, F., Yin, G., Huang, Y., and Chen, L. (2019). Exploration of crystal structure and the origin of unexpected intrinsic ductility of χCo_7Ta_2. *J. Alloys Compounds* 797, 1198–1204. doi: 10.1016/j.jallcom.2019.05.223

Zhang, L., Han, J., Wang, H., Car, R., and E, W. (2018a). Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett*. 120:143001. doi: 10.1103/PhysRevLett.120.143001

Zhang, L., Han, J., Wang, H., Saidi, W., Car, R., and E, W. (2018b). End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. *Adv. Neural Inform. Process. Syst*. 31, 4441–4451.

Zhang, L., Lin, D.-Y., Wang, H., Car, R., and E, W. (2019). Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater*. 3:023804. doi: 10.1103/PhysRevMaterials.3.023804

Zhang, Y., Wang, H., Chen, W., Zeng, J., Zhang, L., Wang, H., et al. (2020). DP-GEN: a concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun*. 253:107206. doi: 10.1016/j.cpc.2020.107206

Zhao, Z., Zhang, S., Yu, T., Xu, H., Bergara, A., and Yang, G. (2019). Predicted pressure-induced superconducting transition in electride Li6 P. *Phys. Rev. Lett*. 122:97002. doi: 10.1103/PhysRevLett.122.097002

Zhuang, H. L., Chen, M., and Carter, E. A. (2017). Prediction and characterization of an Mg-Al intermetallic compound with potentially improved ductility via orbital-free and Kohn-Sham density functional theory. *Model. Simul. Mater. Sci. Eng*. 25:075002. doi: 10.1088/1361-651X/aa7e0c

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 04 December 2020
doi: 10.3389/fchem.2020.590047

Tests on the Accuracy and Scalability of the Full-Potential DFT Method Based on Multiple Scattering Theory

Peiyu Cao[1,2†], *Jun Fang*[3†], *Xingyu Gao*[3*], *Fuyang Tian*[4*] and *Haifeng Song*[3*]

[1] State Key Laboratory for Advanced Metals and Materials, University of Science and Technology Beijing, Beijing, China,
[2] State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing, China,
[3] Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing, China,
[4] Institute for Applied Physics, University of Science and Technology Beijing, Beijing, China

We investigate a reduced scaling full-potential DFT method based on the multiple scattering theory (MST) code MuST, which is released online (https://github.com/mstsuite/MuST) very recently. First, we test the accuracy by calculating structural properties of typical body-centered cubic (BCC) metals (V, Nb, and Mo). It is shown that the calculated lattice parameters, bulk moduli, and elastic constants agree with those obtained from the VASP, WIEN2k, EMTO, and Elk codes. Second, we test the locally self-consistent multiple scattering (LSMS) mode, which achieves reduced scaling by neglecting the multiple scattering processes beyond a cut-off radius. In the case of Nb, the accuracy of 0.5 mRy/atom can be achieved with a cut-off radius of 20 Bohr, even when small deformations are imposed on the lattice. Despite that the calculation of valence states based on MST exhibits linear scaling, the whole computational procedure has an overall scaling of about $\mathcal{O}(N^{1.6})$, due to the fact that the updating of Coulomb potential scales almost as $\mathcal{O}(N^2)$. Nevertheless, it can be still expected that MuST would provide a reliable and accessible way to large-scale first-principles simulations of metals and alloys.

Keywords: first principles, Korringa–Kohn–Rostoker (KKR), multiple scattering theory (MST), full potential, elastic constants

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Jun Kang,
Beijing Computational Science Research Center, China
Shen Zhang,
National University of Defense Technology, China

***Correspondence:**
Xingyu Gao
gao_xingyu@iapcm.ac.cn
Fuyang Tian
fuyang@ustb.edu.cn
Haifeng Song
song_haifeng@iapcm.ac.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *31 July 2020*
Accepted: *10 September 2020*
Published: *04 December 2020*

Citation:
Cao P, Fang J, Gao X, Tian F and Song H (2020) Tests on the Accuracy and Scalability of the Full-Potential DFT Method Based on Multiple Scattering Theory. Front. Chem. 8:590047. doi: 10.3389/fchem.2020.590047

1. INTRODUCTION

Kohn–Sham density functional theory (KS-DFT) (Kohn and Sham, 1965) transforms the many-body problem to a non-interacting system and has been widely used in modern first-principles calculations. Although many computational schemes are developed to solve the Kohn–Sham equation (Kohn and Sham, 1965) for the ground-state properties, the Korringa–Kohn–Rostoker Green's function (KKR-GF) method (Korringa, 1947; Kohn and Rostoker, 1954), also known as multiple scattering theory (MST), provides equivalent information by the single-particle GF (Economou, 2006). In the MST approach, the system is divided into non-overlapping atomic regions as a set of scatterers. To solve the single-site scattering problem, one numerically determines the angular momentum and energy-dependent solutions of the radial Schrödinger equation for a given potential. The coherent matching of the single-site solutions can be achieved if and only if the incoming wave of an atomic site is identical to the superposition of the outgoing waves from all other scatterers. This viewpoint not only gives access to the Kohn–Sham eigenstates but also to the single-electron GF of the system, which leads to the modern KKR-GF method.

The survey (Aarons et al., 2016) suggests that KKR-GF or MST method remains important for large-scale metallic systems. The favorable scaling in MST is attributed to the fact that the electron density, which is the fundamental quantity in DFT, can be obtained from the site-diagonal blocks of the scattering path matrix. And the site-diagonal block of the scattering path matrix for a particular atom can be solved with sufficient accuracy by considering only the electronic multiple scattering processes in a finite-sized region centered at this atom. This region is referred to as the local interaction zone (LIZ), which is originally introduced in the locally self-consistent multiple-scattering method (LSMS) (Wang et al., 1995). Base on the central idea of the LSMS method, the locally self-consistent GF (LSGF) approach (Abrikosov et al., 1996, 1997) can choose judiciously the effective medium to decrease the LIZ size. In particular, the linear scaling has been achieved in LSMS with muffin-tin approximation and in LSGF with tight-binding linear muffin-tin orbital (TB-LMTO) basis. It should be mentioned that besides the MST-based methods, other approaches to reduced scaling DFT methods for metallic systems have also been developed in recent years Pratapa et al. (2016), Suryanarayana et al. (2018), Aarons and Skylaris (2018), Mohr et al. (2018).

There is a trend toward the full-potential (FP) MST in which no shape approximation is assumed for the potential. Many questions in materials science, for example, on complex defects, interfaces, dislocations, as well as nanostructures, come to a great demand for the reduced scaling FP method. KKRnano, a massively parallel DFT package based on MST, has been developed and optimized for thousands of atoms without a compromise on the FP accuracy (Thiess et al., 2012). And this package has been applied to study the role of the vacancy clusters in metal-insulator transitions (Zhang et al., 2012).

However, most MST simulation packages are in-house, which impedes the application of MST as a powerful tool for large scale or disordered systems. Recently, the MuST package, an *ab initio* calculation software package based on FP MST (Rusanu et al., 2011), is open to public and is free to download online (https://github.com/mstsuite) under a BSD 3-clause license. We focus on the MST part in the MuST package, which not only provides features for calculating physical properties of materials but also serves as a platform for implementing and testing the numerical algorithms. At present, the MuST package is capable of performing the following calculations: (1) muffin-tin approximation, (2) FP method, (3) coherent potential approximation (CPA), and (4) LSMS method. And the fully relativistic MST by solving the Dirac equation has been implemented in MuST Liu et al. (2016, 2018). For such a newly released package, it is prerequisite to perform systematic tests both on the accuracy and efficiency.

A reliable FP method can be used to exactly capture the small energy difference for the lattice distortion or deformation. According to the elastic theory, we deform the crystalline cell to the distorted lattice structures and then calculate their energies. The small energy change with the lattice deformation can be used to calculate the elastic constants (Vitos, 2007). Asato et al. (1999) investigated total energy calculations for metals and semiconductors based on the FP MST method. But few work pay attention to validate the elastic properties based on MST, which is fundamental for applying MST to study the structural properties of materials. Considering the anomalies behavior of deformations in body-centered cubic (BCC) V and Nb (Nagasako et al., 2010; Dezerald et al., 2016), we employ the different *ab initio* methods including the FP MST method in MuST package to calculate the elastic constants of V, Nb, and Mo. By comparing with results of available experiments and other popular first-principles packages, we investigate the accuracy of the FP MST method in MuST package.

To estimate the parallel scalability, we carried out strong and weak scaling tests of the FP LSMS method in MuST package. It is seen that the LSMS method exhibits a good strong scalability. This is due to the two-level parallelism over atoms and energy points implemented in MuST package. However, in the weak scaling test, the overall computational procedure is not linear scaling, which seems to be inconsistent with the $\mathcal{O}(N)$ scaling of the muffin-tin LSMS proved in previous work (Wang et al., 1995). By analyzing the implementation scheme, we attribute it to the difference in updating the Coulomb potential between the muffin-tin approximation and the FP method. While the solution of eigenvalue problems is avoided in the MST method, the calculation of Coulomb potential could become the performance bottleneck in large-scale first-principles simulations. For example, PRinceton Orbital-Free Electronic Structure Software (PROFESS) (Ho et al., 2008; Hung et al., 2010; Chen et al., 2015) suggested that about 70% of computation time was spent on fast Fourier transforms (FFTs) to calculate the kinetic and electron–electron Coulomb interaction terms. PROFESS features plane-wave basis set and has been optimized for peta-scale computing (Chen et al., 2016). The calculation of MST is based on angular momentum expansion and new algorithms should be developed to optimize the overall scaling of the FP MST method.

The rest of this paper is arranged as follows. In section 2, we introduce the MST method and its LSMS variant. In section 3, we investigate the accuracy by calculating the equation of state (EOS) and elastic constants of typical BCC metals. In section 4, we test and analyze the scalability of the FP LSMS method. In section 5, some concluding remarks are drawn.

2. METHODOLOGY

2.1. MST Method

The term MST method in this manuscript refers to the modern version of the KKR method, that is, the KKR-GF method. The central quantity to be computed in the MST method changes from the Kohn–Sham orbitals in band theory methods to the one-electron GF, which can be defined as solutions of the following differential equation (Economou, 2006) (non-spin polarized cases assumed and atomic units $\hbar = 1$ and $m_e = 1/2$ used):

$$\{z + \nabla^2 - V_{\text{eff}}[\rho](\boldsymbol{r})\}G(\boldsymbol{r}, \boldsymbol{r}'; z) = \delta(\boldsymbol{r} - \boldsymbol{r}'), \tag{1}$$

where V_{eff} is the Kohn–Sham effective potential under exchange-correlation approximations like the local density approximation (LDA) or the generalized gradient approximation (GGA), $\rho(\boldsymbol{r})$ is the electron density, and $z \equiv \epsilon + \imath\eta$ is a complex variable. If z is

real and ϵ belongs to the continuous spectrum of $-\nabla^2 + V_{\text{eff}}[\rho]$, $G(\boldsymbol{r}, \boldsymbol{r}'; \epsilon)$ is not well-defined and one may define the retarded GF

$$G^+(\boldsymbol{r}, \boldsymbol{r}'; \epsilon) \equiv \lim_{\eta \to 0^+} G(\boldsymbol{r}, \boldsymbol{r}'; \epsilon + \imath\eta). \tag{2}$$

In the following, the superscript + will be omitted. Once the GF is known, the valence electron density can be obtained by (Gonis and Butler, 2000; Economou, 2006; Faulkner et al., 2018)

$$\rho(\boldsymbol{r}) = -\frac{2}{\pi} \text{Im} \int_{\epsilon_B}^{\epsilon_F} G(\boldsymbol{r}, \boldsymbol{r}; \epsilon) \text{d}\epsilon, \tag{3}$$

where ϵ_F is the Fermi energy, the bottom energy ϵ_B is chosen between the highest core-state energy and the valence band, and the factor 2 accounts for the number of electron spins. The energy integration in Equation (3) can be carried out along a contour in the complex energy plane so that only few tens of energy points are needed. Other physical quantities like the density of states (DOS) can also be obtained from the GF (Gonis and Butler, 2000; Economou, 2006; Faulkner et al., 2018).

The MST method provides a convenient access to the GF. In the MST method, atoms in the system are considered as scattering centers of which the scattering properties are described by the so-called single-site scattering t-matrix (Gonis and Butler, 2000; Faulkner et al., 2018). The space is divided into non-overlapping cells Ω_n centered at atomic positions $\boldsymbol{R}_n$, where n is the index of atoms in the system. In the vicinity of atomic site n, it is proved that the GF can be expressed as (Faulkner and Stocks, 1980; Gonis and Butler, 2000; Sébilleau, 2000; Zabloudil et al., 2006)

$$\begin{aligned} G(\boldsymbol{r}_n, \boldsymbol{r}_n; \epsilon) \; = \; & \sum_{LL'} Z_L^n(\boldsymbol{r}_n; \epsilon) \tau_{LL'}^{nn}(\epsilon) Z_{L'}^{n\times}(\boldsymbol{r}_n; \epsilon) \\ & - \sum_{L} Z_L^n(\boldsymbol{r}_n; \epsilon) J_L^{n\times}(\boldsymbol{r}_n; \epsilon), \end{aligned} \tag{4}$$

where L is the combined index of angular momentum quantum number l and magnetic quantum number m, $\boldsymbol{r}_n \equiv \boldsymbol{r} - \boldsymbol{R}_n$ the relative coordinate, $Z_L^n(\boldsymbol{r}_n; \epsilon)$ and $J_L^n(\boldsymbol{r}_n; \epsilon)$ regular and irregular solutions of the single-site problem in cell n for momentum L and energy ϵ, and $\tau_{LL'}^{nn}(\epsilon)$ site-diagonal matrix elements of the scattering path operator $\tau^{nm}(\epsilon)$ in the angular momentum representation. The $\times$ symbol in Equation (4) means that we take the complex conjugate of the spherical harmonics and keep remaining parts of the function unchanged.

The scattering path operator $\tau^{nm}(\epsilon)$ describes all possible scattering events of electron states with energy ϵ between atomic sites n and m. In the angular momentum representation, the corresponding scattering path matrix is given by (Gonis and Butler, 2000; Zabloudil et al., 2006)

$$\begin{aligned} \underline{\tau}^{nm} &= \underline{t}^n \delta_{nm} + \underline{t}^n \underline{G}_0^{nm} (1 - \delta_{nm}) \underline{t}^m + \underline{t}^n \sum_{k \neq n} \underline{G}_0^{nk} \underline{t}^k \underline{G}_0^{km} (1 - \delta_{km}) \underline{t}^m \\ &\quad + \underline{t}^n \sum_{k \neq n} \underline{G}_0^{nk} \underline{t}^k \sum_{j \neq k} \underline{G}_0^{kj} \underline{t}^j \underline{G}_0^{jm} (1 - \delta_{jm}) \underline{t}^m + \ldots \\ &= \underline{t}^n \delta_{nm} + \underline{t}^n \sum_{k \neq n} \underline{G}_0^{nk} \underline{\tau}^{km}, \end{aligned} \tag{5}$$

where the underline symbol indicates matrices with respect to the angular momentum index L, $\underline{t}^n(\epsilon)$ is the single site scattering t-matrix associated with site n, and $\underline{G}_0^{nm}(\epsilon)$ is the free-electron propagator matrix in the angular momentum representation, also known as KKR structure constant matrix, that describes the propagation of a free electron with energy ϵ from site n to site m. Note that we have omitted the dependence on energy ϵ in Equation (5) for a compact expression.

In the case of a finite system with N atoms, it is seen from the second equation in Equation (5) that the scattering path matrix can be computed directly by a matrix inversion:

$$\underline{\tau}^{nm}(\epsilon) = \begin{pmatrix} [\underline{t}^1(\epsilon)]^{-1} & -\underline{G}_0^{12}(\epsilon) & -\underline{G}_0^{13}(\epsilon) & \cdots & -\underline{G}_0^{1N}(\epsilon) \\ -\underline{G}_0^{21}(\epsilon) & [\underline{t}^2(\epsilon)]^{-1} & -\underline{G}_0^{23}(\epsilon) & \cdots & -\underline{G}_0^{2N}(\epsilon) \\ -\underline{G}_0^{31}(\epsilon) & -\underline{G}_0^{32}(\epsilon) & [\underline{t}^3(\epsilon)]^{-1} & \cdots & -\underline{G}_0^{3N}(\epsilon) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\underline{G}_0^{N1}(\epsilon) & -\underline{G}_0^{N2}(\epsilon) & -\underline{G}_0^{N3}(\epsilon) & \cdots & [\underline{t}^N(\epsilon)]^{-1} \end{pmatrix}^{-1}_{nm}, \tag{6}$$

where the subscript nm on the right hand side indicates the block at the nth row and mth column of the big matrix after the inversion has been taken. In the case of periodic systems, the equation in Equation (5) for the scattering path matrix can be solved by the lattice Fourier transform, leading to (we assume that there is only one atom in the unit cell):

$$\underline{\tau}^{nm}(\epsilon) = \frac{1}{\Omega_{\text{BZ}}} \int_{\Omega_{\text{BZ}}} \left[\underline{t}(\epsilon)^{-1} - \underline{G}_0(\boldsymbol{k}, \epsilon) \right]^{-1} \text{e}^{\imath \boldsymbol{k} \cdot (\boldsymbol{R}_n - \boldsymbol{R}_m)} \text{d}\boldsymbol{k}, \tag{7}$$

where Ω_{BZ} is the first Brillouin zone and $\underline{G}_0(\boldsymbol{k}, \epsilon)$ is the lattice Fourier transform of $\underline{\underline{G}}_0(\epsilon)$ (the double underline indicates matrices with respect to the angular momentum index and the atomic site index) (Gonis and Butler, 2000; Zabloudil et al., 2006).

2.2. LSMS Method

As described above, the MST method makes unnecessary the calculation of the Kohn–Sham orbitals, and consequently the time-consuming procedure for diagonalization and orthogonalization in the conventional KS-DFT calculations can be entirely avoided. The only global operation required by computing the GF is to obtain the scattering path matrix by an inversion of the matrix in Equation (6). Since the size of the matrix is proportional to the number of atoms in the unit cell, the MST method still suffers from cubic scaling limitation.

To reduce the scaling of the procedure, we can calculate the nth site-diagonal block of the scattering path matrix $\underline{\tau}^{nn}$ by neglecting the multiple scattering processes that involve atoms beyond some cut-off radius R_{LIZ} from atomic site n. This is based on the observation that the scattering processes involving far away atoms have little influence on the electronic scattering behavior in the vicinity of atomic site n, which is the essence of the LSMS method. The region within distance R_{LIZ} from the central atom is called the LIZ. If there are M atoms in the LIZ, the solution of the multiple scattering problem scales as $\mathcal{O}(NM^3)$, where N is the total number of atoms. Consequently, the LSMS method is expected to exhibit the linear scaling in N with a prefactor determined by M and the number of basis functions.

2.3. Coherent Potential Approximation

Due to the convenient access to the GF, the MST method plays a prominent role in first-principles alloy theory, in which a novel candidate is the CPA (Soven, 1967; Taylor, 1967; Gyorffy, 1972; Ruban and Abrikosov, 2008). The CPA is designed to obtain an ordered effective medium to describe properties of the multi-component random alloy. The scattering path operator of the CPA effective medium, denoted by τ_{CPA}, is determined by the following self-consistency condition (two-component alloy as the example):

$$\tau_{\mathrm{CPA}} = C_A\tau_A + C_B\tau_B, \tag{8}$$

where $\tau_{A(B)}$ is the scattering path operator of the auxiliary system constructed by replacing the central site in the ordered effective medium system by the alloy component A(B). Within the single-site approximation, it can be proved that the GF of the CPA effective medium system is equal to the targeted ensemble averaged GF (Faulkner, 1982; Ebert et al., 2011). The CPA condition in Equation (8) needs to be reformulated into a proper expression to be suited for numerical applications (Faulkner, 1982; Turek et al., 1997).

3. TEST ON ACCURACY

In this section, we investigate the accuracy of the FP MST method implemented in the MuST package by comparing equilibrium bulk properties and elastic constants with those calculated by the WIEN2k, EMTO, and VASP codes.

3.1. Calculation Details

In order for a meaningful comparison, we used the Perdew–Burke–Ernzerhof (PBE) exchange-correlation functional (Perdew et al., 1996) in all our calculations, and carried out convergence tests to determine the numerical parameters for each code. The relativistic effect of the core electrons was treated via the default scheme in each package. In the following, we enumerate the detailed settings of numerical parameters.

3.1.1. MuST

The uniform grid for the computation of the Coulomb potential was chosen as $64 \times 64 \times 64$. The Monkhorst–Pack $\boldsymbol{k}$-point mesh was set to be $21 \times 21 \times 21$ in all the KKR tests. The break condition for the electronic SCF (self-consistent field) iterations was that differences in the total energy and the potential are smaller than 5×10^{-8} Ry and 10^{-7} Ry, respectively. The maximum angular momentum used in the expansion of the wave functions and the GFs was set to $l_{\max} = 4$. The number of radial grid points from the atomic center to the muffin-tin radius was chosen to be 2001, which is sufficiently accurate for solving the single-site scattering problem.

3.1.2. WIEN2k

The WIEN2k package (Blaha et al., 2020) implements an FP linearized augmented plane-wave (LAPW) method. No shape approximations have been made on the potential and charge density inside the muffin-tin spheres and in the interstitial region. In our calculations, the muffin-tin sphere radius was fixed as 2.50 Bohr, the cutoff parameter $R_{\mathrm{MT}} \cdot K_{\max}$ was chosen to be 8.00, and the plane-wave expansion cutoff $G_{\max}$ was set as 14.00 Ry. And a $15 \times 15 \times 15$ Monkhorst–Pack $\boldsymbol{k}$-point mesh was used for the Brillouin zone sampling. The chosen $R_{\mathrm{MT}} \cdot K_{\max}$ and $\boldsymbol{k}$-mesh ensure that errors in the total energies of the deformed structures are converged to 10^{-4} Ry in elastic constant calculations.

3.1.3. EMTO

The EMTO package implements the so-called exact muffin-tin orbitals method, in which different from former muffin-tin methods, the single-electron states are calculated exactly for the optimized overlapping muffin-tin (OOMT) potential. We refer the readers to Vitos et al. (2001), Vitos (2001), and Vitos (2007) for the detailed theory and applications of the EMTO method. In our calculations, the EMTO basis set including *s*, *p*, *d*, and *f* orbitals was used in combination with soft-core approximation. For the integration over energy in the complex plane, we used 24 points along a semicircular contour. The Brillouin zone was sampled by a $21 \times 21 \times 21$ Monkhorst–Pack $\boldsymbol{k}$-point mesh to make the total energies of the deformed structures converge up to 3×10^{-5} Ry.

3.1.4. VASP

The Vienna *ab initio* simulation package (VASP) (Kresse and Furthmüller, 1996a,b; Kresse and Joubert, 1999) describes the electron-ion interactions by the projector-augmented wave (PAW) method. In our calculations, the kinetic energy cutoff for the plane-wave basis set was 400 eV. A $15 \times 15 \times 15$ Monkhorst–Pack $\boldsymbol{k}$-point mesh was used for the Brillouin zone sampling. And the SCF convergence criterion was set to be 10^{-7} Ry.

3.2. Equation of State

The lattice parameter a, bulk modulus B, and pressure derivative of the bulk modulus B' have been commonly used for accuracy assessments of DFT codes and (pseudo)potential libraries (Kucukbenli et al., 2014; Lejaeghere et al., 2014, 2016). These structural properties can be extracted from the EOS for a solid. For instance, in a Morse type of EOS, the total energy is fitted by an exponential function with four parameters (D_0, γ, a_0, and E_0)

$$E(a) = D_0 e^{-\gamma\left(\frac{a}{a_0}-1\right)} - 2D_0 e^{-\frac{\gamma}{2}\left(\frac{a}{a_0}-1\right)} + E_0. \tag{9}$$

Then a_0, B_0, and B' can be derived from the Morse function and used to assess the accuracy of DFT codes under investigation.

To investigate the accuracy of the FP MST method in the MuST package, we calculated a_0, B_0, and B' of three bulk systems V, Nb and Mo, and compared the results with all-electron packages including WIEN2k, EMTO, and VASP. In these tests, we employed the same exchange-correlation functional and equivalent numerical settings, as introduced in section 3.1. **Figure 1** shows the calculated $E(a)$ curves from these packages, which have been shifted so that all $E(a)$ points with the lowest energy are adjusted to zero. The fitted results of a_0, B_0, and B' are given in **Table 1**. In addition, results from the Elk package, those obtained by the FP-LMTO method, and experimental values from the literature are also listed in **Table 1** as a reference. The

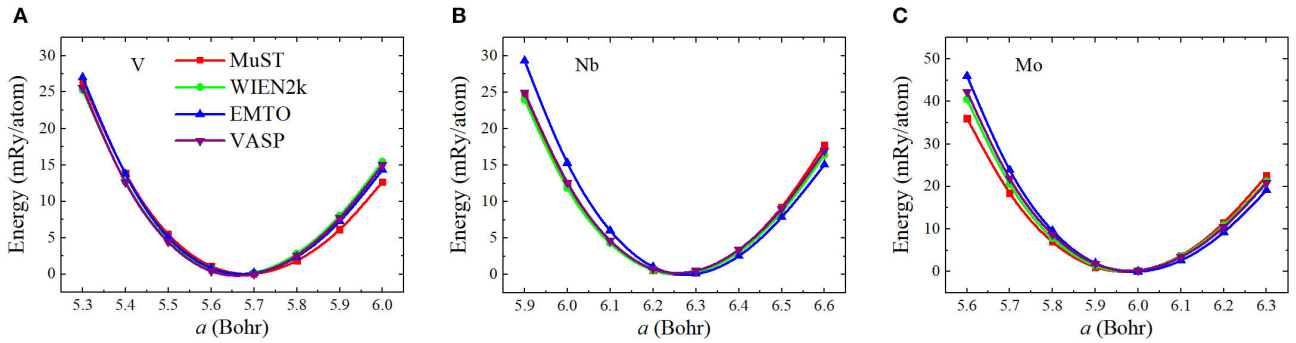

FIGURE 1 | (Color online) Equation of states (the total energy per atom vs. lattice parameter) for body-centered cubic (BCC) V **(A)**, Nb **(B)**, and Mo **(C)**. The total energies have been shifted so that all $E(a)$ points with the lowest energy are adjusted to zero.

TABLE 1 | Equilibrium bulk properties [lattice parameter a (Bohr), bulk modulus B (GPa), and pressure derivative of the bulk modulus B'], the elastic constants $c' = (c_{11} - c_{12})/2, c_{11}, c_{12}$, and c_{44}(GPa) for body-centered cubic (BCC) V, Nb, and Mo metals.

Method	a	B	B'	c'	c_{11}	c_{12}	c_{44}
				V			
MuST	5.685	170.32	3.72	60.42	250.88	130.04	52.46
WIEN2k	5.667	182.10	3.72	59.69	268.91	149.53	19.90
EMTO	5.673	178.73	3.06	79.72	285.03	125.58	51.61
VASP	5.666	183.18	3.22	60.71	264.13	142.71	20.50
Elk (Lejaeghere et al., 2016)	5.663	182.89	3.89	–	–	–	–
FP-LMTO (Landa et al., 2006a)	5.673	182.70	–	67.30	272.43	137.83	37.40
Exp.(Frederikse, 1972; Maschke and Levy, 1983; Young, 1991; Haas et al., 2009)	5.713 (5.715)	155.0	–	54.85 (57)	228.7	119.0	43.20 (46)
				Nb			
MuST	6.258	178.62	3.22	51.94	247.88	144.00	35.52
WIEN2k	6.258	168.77	3.01	47.87	232.60	136.85	14.88
EMTO	6.278	177.44	2.86	74.07	276.20	128.06	51.00
VASP	6.254	171.87	3.21	50.02	238.56	138.52	16.62
Elk (Lejaeghere et al., 2016)	6.256	170.92	3.84	–	–	–	–
FP-LMTO (Landa et al., 2006a)	6.270	170.7	–	63.9	225.9	128.1	25.5
Exp.(Frederikse, 1972; Ashkenazi et al., 1978; Young, 1991; Haas et al., 2009)	6.237 (6.225)	169	–	52.89 (60)	246.5	134.5	28.73 (31)
				Mo			
MuST	5.968	253.65	3.22	169.52	479.67	140.64	131.59
WIEN2k	5.973	263.47	3.99	147.45	460.07	165.17	103.15
EMTO	5.991	254.21	4.89	169.80	480.61	141.01	131.02
VASP	5.978	263.08	3.15	148.57	461.18	164.03	102.56
Elk (Lejaeghere et al., 2016)	5.973	259.07	4.22	–	–	–	–
FP-LMTO (Söderlind et al., 2000)	5.970	255	–	139	440	162	139
Exp.(Dickinson and Armstrong, 1967; Frederikse, 1972; Young, 1991; Haas et al., 2009)	5.947 (5.936)	261	–	152.95	463.7 (473.0)	157.8 (156.2)	109.2 (110.9)

Exp. stands for experimental results. The experimental values in parenthesis are the experimental lattice constants corrected for the zero-point anharmonic expansion (ZPAE) and experimental elastic constants corrected at 0 K. The MuST, WIEN2k, EMTO, VASP, Elk, and FP-LMTO are different ab initio methods.

differences with respect to the results from Elk are illustrated in **Figure 2**.

We see from **Table 1** and **Figure 2** that except for the bulk modulus of V, differences in the calculated a and B results are less than 0.5 % and 5 %, respectively, which could be considered as small discrepancies between different codes (Holzwarth et al., 1997; Kresse and Joubert, 1999; Lejaeghere et al., 2016). The lattice parameter a of BCC V metal obtained by MuST is slightly larger than that of other *ab initio* methods, but the difference of a is only 0.39%, with respect to the Elk's result. For BCC Nb and Mo, both MuST lattice parameters are very close to those results from Elk. The relative error is 0.03% for Nb and 0.08%

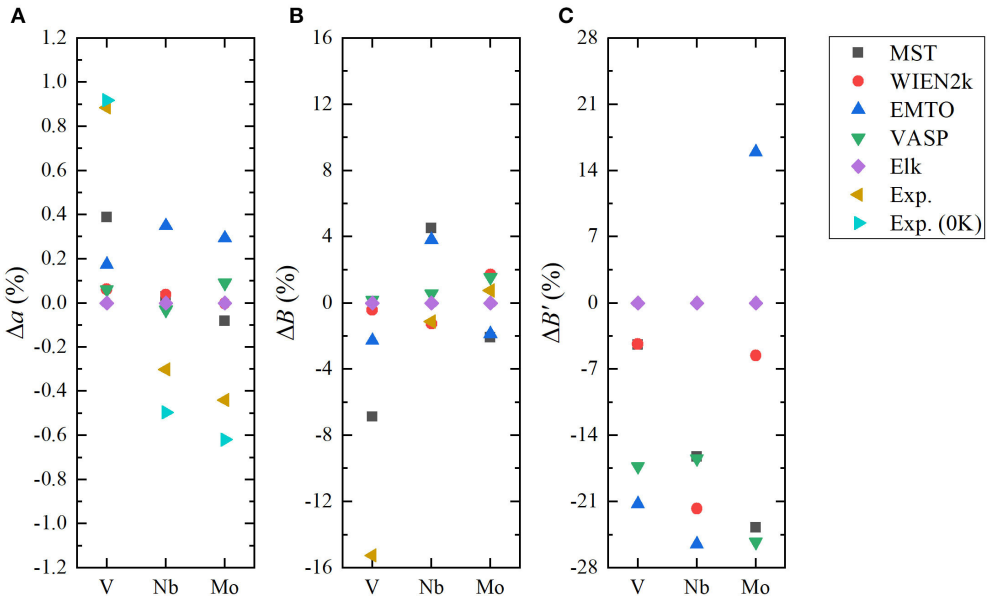

FIGURE 2 | (Color online) Relative errors in lattice parameter **(A)**, bulk modulus **(B)**, and its derivation **(C)** (Δa, B, $\Delta B'$) for body-centered cubic (BCC) V, Nb, and Mo, where the Elk results are taken as reference values.

for Nb, respectively. Generally speaking, the PBE predicted lattice parameter is overestimated, that is, theoretical lattice parameter is usually larger than experimental values, whereas for V, all present *ab initio* lattice parameters listed in **Table 1** are smaller than the experimental value at 0 K. But for Nb and Mo, the *ab initio* lattice parameters are slightly larger, with respect to their experimental values. The bulk modulus B represents the stress v.s. the volume expansion or compression. And its derivative B' can be used to describe the anharmonic effect in the vibrating lattice. Comparing the calculated bulk moduli and their derivatives, we find that for BCC V metal the MuST B is slightly smaller, within 6.9%, than the Elk bulk modulus, while EMTO, WIEN2k, and VASP results agree well with each other. This is consistent with the fact that the MuST lattice constant is slightly larger, within 0.4%, than the Elk result, whereas the relative discrepancy is within 0.2% among the results of other codes.

Finally, it is necessary to mention that the energy-lattice curve of a solid is sensitive to the treatment of semi-core states. For example, Nb has core (1*s*, 2*s*, 2*p*, 3*s*, 3*p*, 3*d*), semi-core (4*s*, 4*p*), and valence (4*d*, 5*s*) states. Due to the limitation in the current implementation of MuST, only the 4*d*5*s* electrons of Nb are considered as valence electrons, and the semi-core states are treated as core states. The same treatment is imposed for the semi-core states of V (3*s*, 3*p*) and those of Mo (4*s*, 4*p*). In MST as well as in other all-electron methods including LAPW and EMTO, both the core and the valence states participate in the self-consistent iteration. The difference is that the core states are calculated using the spherical part of the crystal potential in the atomic sphere Singh and Nordström (2006). The wave function for each core state is confined and normalized within the sphere radius. In the case that semi-core states are treated as core states, since their charges are no longer well confined inside the atomic sphere, the so-called confinement error appears and a proper setting of the bounding sphere radius becomes important Asato et al. (1999). Different from MuST, in the WIEN2k calculations, a recommended separation energy of -6.0 Ry automatically defines the core- and band-states. Accordingly, both the semi-core and valence states of V, Nb, and Mo metals are treated as band states and solved using the full potential of the crystal. In the PAW method of the VASP package, the frozen-core approximation is used, so the core electrons will not participate in the self-consistent calculations. And the PAW atomic datasets including semi-core states for V, Nb, and Mo are provided by the VASP POTCAR library to be utilized for accurate calculations. The main point is that: the differences in the treatment of the semi-core states may cause noteworthy discrepancies in the calculated results, and we suggest that the semi-core states are allowed to be treated as band states in the future version of MuST.

3.3. Elastic Constants

In a cubic lattice there are three independent elastic constants, c_{11}, c_{12}, and c_{44}, of which c_{11} and c_{12} are connected to the bulk modulus $B = (c_{11} + 2c_{12})/3$ and the tetragonal shear modulus $c' = (c_{11} - c_{12})/2$. The two shear elastic parameters c' and c_{44} were computed according to the standard methodology (Vitos, 2007). For example, we used the following volume conserving orthorhombic and monoclinic deformations:

$$\begin{pmatrix} 1+\delta_o & 0 & 0 \\ 0 & 1-\delta_o & 0 \\ 0 & 0 & \frac{1}{1-\delta_o^2} \end{pmatrix} \text{and} \begin{pmatrix} 1 & \delta_m & 0 \\ \delta_m & 1 & 0 \\ 0 & 0 & \frac{1}{1-\delta_m^2} \end{pmatrix},$$

which lead to the energy change $\triangle E(\delta_o) = 2V_c'\delta_o^2 + O(\delta_o^4)$ and $\triangle E(\delta_m) = 2V_{c44}'\delta_m^2 + O(\delta_m^4)$. Both energies were computed

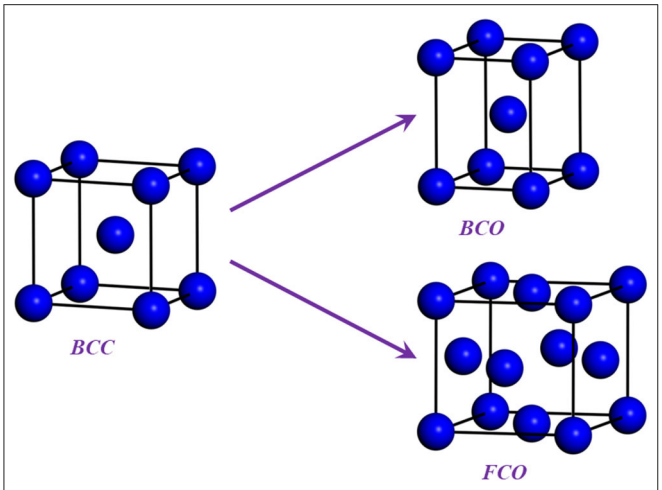

FIGURE 3 | (Color online) The deformation configurations [body center orthorhombic (BCO) for the calculation of c' and face center orthogonal (FCO) for the calculations of c_{44}] for body center cubic (BCC) crystal.

for six distortions, $\delta = 0.00, 0.01, ..., 0.05$. The body center orthorhombic (BCO) for c' and faced center orthorhombic (FCO) for c_{44} are shown in **Figure 3**.

Results of elastic constants from different *ab initio* methods and experiments are listed in **Table 1**. Their differences with respect to experiments are shown in **Figure 4**. Due to the calculations of c_{11} and c_{12} via the combination of c' and bulk modulus, the accuracy of c' plays a key role in the quality of c_{11} and c_{12} results. From **Figure 4**, we can see that for the c' of V, the MuST result agrees well with the results from WIEN2k and VASP. Due to the small bulk modulus, our MuST calculated c_{11} and c_{12} are slightly different from those of WIEN2k and VASP. For c_{11} and c_{12} of Nb, results from MuST, WIEN2k, and VASP are all close to experiments at room temperature, whereas the difference of c' between calculations and experiments at 0 K is up to 13.4% for MuST, 20.2% for WIEN2k, and 16.6% for VASP. For Mo, the discrepancy of c' with the experimental value is up to 11.0% for MuST and EMTO, but it is only 2.9%/3.6% for VASP/WIEN2k. This results in the large difference for c_{11} and c_{12} between MuST/EMTO and WIEN2k/VASP calculations. Although EMTO and FP-LMTO can be regarded as similar muffin-tin type methods, their calculated elastic constants are very different. The main reason is that the available FP-LMTO results were calculated based on the LDA (Söderlind et al., 2000).

From **Table 1**, we can find for V and Nb that c_{44} results of MuST and EMTO are close to experimental values, while those from WIEN2k and VASP much smaller. We note that the early work on elastic constants c_{44} is 17.1 GPa for V and 10.3 GPa for Nb (Koči et al., 2008; Nagasako et al., 2010). There is an anomalous dispersion of transverse acoustic phonons propagating along the <100> direction in V and Nb. Softening of acoustic phonons induces small values of the shear constant. The soft acoustic phonons and small shear constants are related to the nesting properties of the Fermi surface, which produce a van Hove singularity in the electronic DOS near the Fermi level (Landa et al., 2006b; Nagasako et al., 2010). Due to the presence of van Hove singularity, an extremely fine mesh for Brillouin zone integration suggested in Nagasako et al. (2010) was expected to determine the exact c_{44}. However, in practice, the convergence of c_{44} with respect to the $\boldsymbol{k}$-point density may be very slow (Nagasako et al., 2010). Instead of using an extremely dense $\boldsymbol{k}$-mesh, the smearing methods can be used to handle the singularity in DOS. It is reported in Nagasako et al. (2010) that the smearing method has a clear impact on the c_{44} results. We note that smearing is performed in WIEN2k and VASP calculations, but in the MuST and EMTO codes no smearing methods are used. This might be the reason on the discrepancy between theoretical results. For Mo, the MuST calculated c_{44} is 18.7% larger than the experimental value at 0 K, but the c_{44} from WIEN2k and VASP are in good agreement with experiments. The ultimate reason of differences in the elastic constant between *ab initio* calculations and experiments is still far from resolved. We noted that there exist variations between experiment results, for example, for Mo the experimental value of c_{44} is about 110.9 GPa at 0 K reported in Dickinson and Armstrong (1967), while another experiment is about 125 GPa at 0 K (Featherston and Neighbours, 1963). So it may be necessary to estimate the accuracy of experiments at low temperatures and the improved extrapolation method may also be desirable.

4. TEST ON SCALABILITY

In this section, we investigate the strong and weak scalabilities of the FP LSMS method implemented in MuST package.

4.1. Convergence on LIZ Size

In practice, the first question on the LSMS method may be the proper choice of the LIZ size for an atomic site. We can calculate the total energy of the bulk system using the LSMS method with increasing LIZ sizes and compare the results with those obtained by the standard MST method. The convergence tests on the LIZ radius have been performed for face-centered cubic (FCC) Cu and BCC Mo in Faulkner et al. (2018). For FCC Cu, the LSMS energy agrees with the reference MST result to better than 0.5 mRy when 6 neighboring shells are included in the LIZ. This corresponds to a cluster of 87 atomic sites with a LIZ radius of 11.7 Bohr. For BCC Mo, on the other hand, a larger LIZ is required in order to achieve better than 0.5 mRy accuracy. We test the convergence of the LIZ radius for BCC Nb and its deformed structures. As shown in **Figure 5**, the LSMS total energies converge to the MST results when the LIZ radius is larger than 20 Bohr. Indeed, we have achieved the accuracy of 0.5 mRy by including 14 neighboring shells into the LIZ, which corresponds to about 330 atoms. This is due to the fact that the Fermi energy falls in the *d* bands so that the DOS near the Fermi energy is significant.

We would like to mention that in the MuST code, the LIZ is embedded in the vacuum with free-electron GF. The LIZ size may be effectively decreased by choosing the effective medium instead (Zeller et al., 1995; Abrikosov et al., 1997), which could provide clues for improving the performance of the LSMS method in a future release.

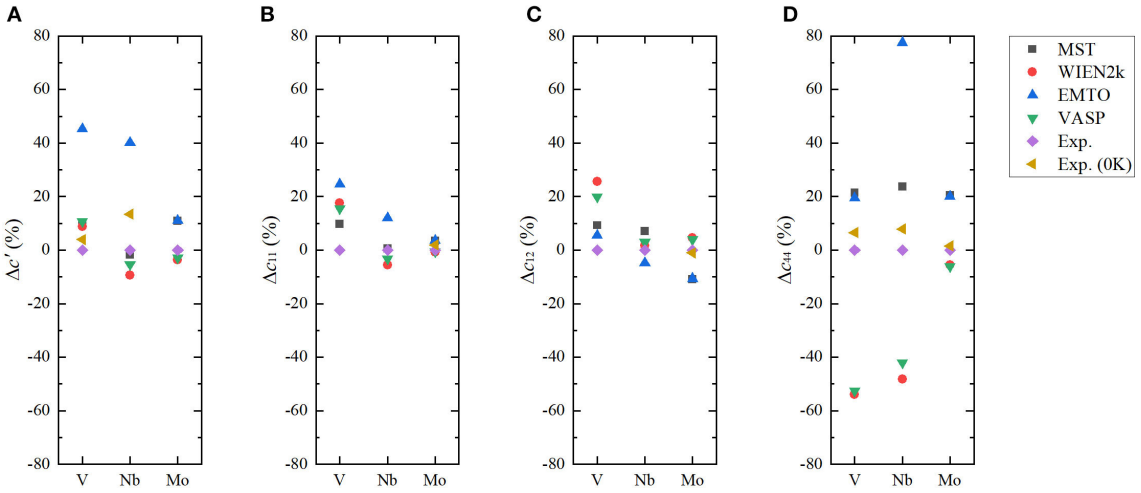

FIGURE 4 | (Color online) Differences in the calculated c' **(A)**, c_{11} **(B)**, c_{12} **(C)**, and c_{44} **(D)** from *ab initio* methods for body-centered cubic (BCC) V, Nb, and Mo with respect to experimental values at room temperature.

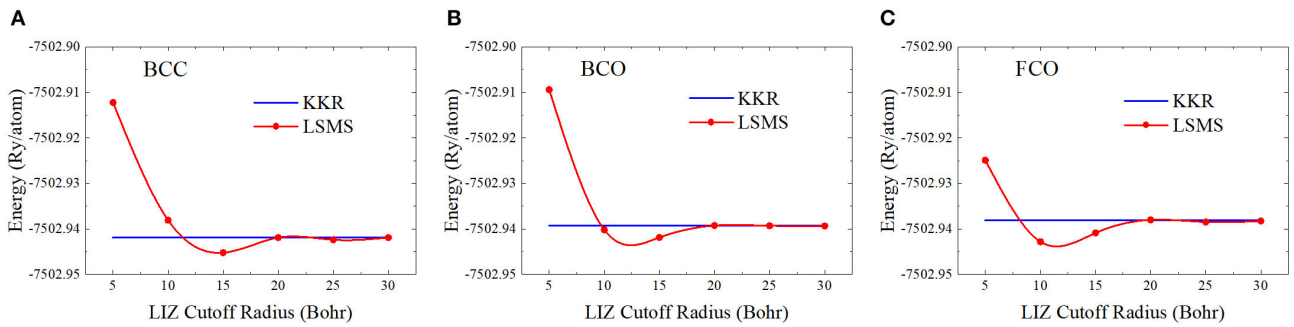

FIGURE 5 | (Color online) Energy as a function of the local interaction zone (LIZ) radius for **(A)** body-centered cubic (BCC), **(B)** body center orthorhombic (BCO), and **(C)** faced center orthorhombic (FCO) structures, where the "KKR" stands for the results from the standard multiple scattering theory (MST) method.

4.2. Strong Scalability

The complexity of the FP MST method can be estimated by the weak scaling test. A good strong scalability is a prerequisite to an effective weak scaling test. Under the premise of good strong scalability, the computational overhead can be revealed by the execution time since the communication overhead contributes a small percentage. We constructed a BCC supercell consisting of 1024 niobium (Nb) atoms. The LIZ of each atomic site contains 89 atoms. As illustrated in **Table 2**, the LSMS method exhibits a good strong scalability. This is due to the two-level parallelism over atoms and energy points implemented in MuST package. The 1024 atoms are distributed over from 128 to 1024 MPI (message passing interface) processes. When the number of MPI processes exceeds the number of atoms, a second level of parallelization over energy points is performed.

The intrinsic parallelism comes from the fact that the computation of the GF for each atom and each energy point along the complex contour is essentially independent. Each MPI process exchanges t-matrix with the others treating the neighboring atoms in the LIZ region. There are no global operations involved in the process of calculating the GF other than few global sum operations such as the summation of the net charge in each atomic cell for the determination of the electron chemical potential. Consequently, the MST method can be highly parallelized, as shown in **Figure 6**.

TABLE 2 | Strong scalability test of the full-potential multiple scattering theory (MST).

# MPI	128	256	512	1024	2048
Execution time (s) per SCF iteration	17581	8804	4855	2456	1255
Parallel efficiency (%)	100.0	99.8	90.5	89.5	87.6

4.3. Weak Scalability

In the weak scaling test, the system size and the number of MPI processes are increased concurrently when one atom per MPI process is kept unchanged. In the pure atom parallelization, we observe a significant growth in the execution time with increasing atoms, as shown in **Table 3**. Consequently, the overall complexity

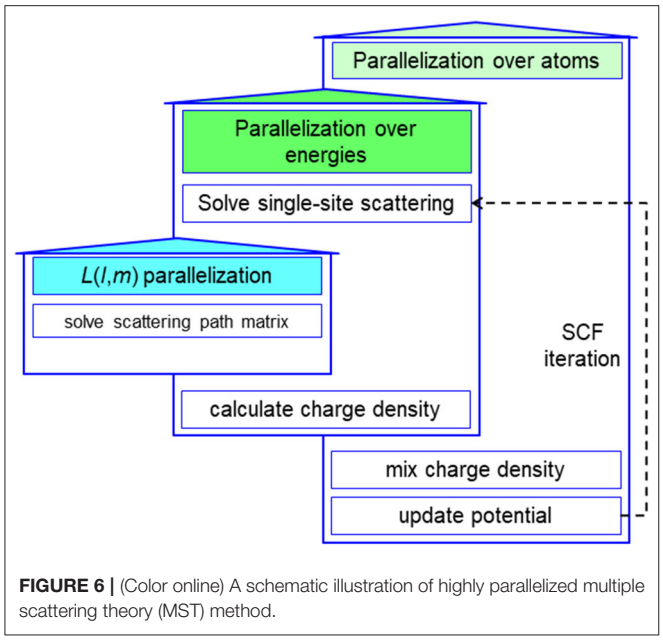

FIGURE 6 | (Color online) A schematic illustration of highly parallelized multiple scattering theory (MST) method.

TABLE 3 | Weak scalability test of the full-potential locally self-consistent multiple scattering (LSMS) method where t_{scf} stands for the execution time on one SCF iteration, t_{val} the execution time on calculating the valence states based on multiple scattering theory (MST), t_{pot} the execution time on updating the effective potential, t_{xc} the execution time on updating the exchange-correlation potential, and t^* to be defined in Section 4.4 is the execution time for an interpolation step in updating the Coulomb potential.

# atoms	t_{scf} (s)	t_{val} (s)	t_{pot} (s)	t^* (s)	t_{xc} (s)
64	391.45	249.28	141.86	120.84	19.94
128	517.29	252.65	263.03	241.74	19.79
256	818.45	253.18	562.11	539.58	20.46
512	1374.58	255.09	1115.19	1092.71	19.97
1024	2479.77	255.05	2216.79	2182.25	19.98

of the FP MST method is not $\mathcal{O}(N)$. The execution time of one SCF iteration can be divided into two parts. One part is for the solution of the valence state by KKR-GF method. The other is used to update the effective potential. They are denoted by t_{val} and t_{pot} in **Table 3**. It can be observed that t_{val} remains almost the same while t_{pot} grows with increasing system size. So the linear scaling is achieved in solving the GF function, which is consistent with the tests in Thiess et al. (2012). As the system size becomes large, the computational overhead on updating the effective potential becomes gradually dominant, which deserves further analysis.

4.4. Scaling Analysis for Updating Potential

As shown in **Table 3**, the execution time on updating the exchange-correlation potential, denoted by t_{xc}, remains almost the same as system size increases. Therefore, we concentrate on the Coulomb potential. In the FP method, the total charge density is divided into the following two parts:

$$\rho(\boldsymbol{r}) = \tilde{\rho}(\boldsymbol{r}) + \hat{\rho}(\boldsymbol{r}), \tag{10}$$

where $\tilde{\rho}$ is chosen as a smoothly varying density and $\hat{\rho}$ is the sphere-bounding non-overlapping charge density. The associated Coulomb potential with $\hat{\rho}$ can be formulated as like

$$\hat{V}_{\text{Coul}}(\boldsymbol{r}) = 2\int_{\mathbb{R}^3} \frac{\hat{\rho}(\boldsymbol{r}') + \rho_0}{|\boldsymbol{r}' - \boldsymbol{r}|} d\boldsymbol{r}' - 2\sum_j \frac{Z_j}{|\boldsymbol{r} - \boldsymbol{R}_j|}, \tag{11}$$

which can be calculated by the multi-pole expansion technique together with the periodic boundary condition and the constraint

$$\int_{\mathbb{R}^3} \hat{\rho}(\boldsymbol{r}) d\boldsymbol{r} + \rho_0 \int_{\mathbb{R}^3} d\boldsymbol{r} = \sum_j Z_j. \tag{12}$$

The procedure is somewhat analogous to the calculation of the Coulomb potential in muffin-tin approximation. The difference is that the non-spherical potential in FP method has multi-pole expansion while the spherical one in muffin-tin approximation has only zero-order moment. Actually, both the two schemes have linear scaling.

The charge density $\tilde{\rho}$ can be regarded as a pseudo electron density varying smoothly. The associated Coulomb potential can be determined by solving the Poisson equation:

$$-\nabla^2 \tilde{V}(\boldsymbol{r}) = 4\pi \tilde{\rho}(\boldsymbol{r}). \tag{13}$$

And fast Fourier transform (FFT) is used for solving Equation (13). In the MST method, both the electron density and one-electron potential are discretized on the spherical mesh around each atom. Therefore, an interpolation from the uniform FFT mesh to the spherical mesh is required. More specifically, the radial part $\tilde{V}_L$ is calculated from $\tilde{\rho}$ on the uniform FFT grids. The computational scheme can be formulated as the integral form:

$$\tilde{V}_L(r_j) = \frac{2}{\pi}(-\imath)^l \int_{\mathbb{R}^3} \tilde{\rho}(\boldsymbol{r}') F_L(r_j, r_j') d\boldsymbol{r}', \tag{14}$$

where $F_L(r_j, r_j')$ is defined as follows:

$$F_L(r_j, r_j') \equiv \int_{\mathbb{R}^{-3}\setminus\{\boldsymbol{0}\}} \frac{1}{|\boldsymbol{G}|^2} j_l(|\boldsymbol{G}|r_j) Y_L^*(\hat{\boldsymbol{G}}) e^{\imath \boldsymbol{G}\cdot\boldsymbol{r}_j'} d\boldsymbol{G}, \tag{15}$$

where j_l is the spherical Bessel function and Y_L is the spherical harmonic. In Equations (14) and (15), r_j stands for the radial grid point, $\boldsymbol{r}'$ and $\boldsymbol{G}$ represent the real-space FFT grid and the corresponding reciprocal grid, respectively, and $r_j' = |\boldsymbol{r}' - \boldsymbol{R}_j|$.

Since $F_L(r_j, r_j')$ is independent of the electron density, it can be setup once before the SCF iteration. However, the summation in Equation (14) scales as $N_{\text{rad}} \cdot N \cdot N_{\text{FFT}}$, where N_{rad} is the number of radial grids and set to be 2001 in the test. And N_{FFT} is proportional to the number of atoms N, which yields $\mathcal{O}(N^2)$ scaling in performing the interpolation like Equation (14). We locate the code segment to perform (14) and denote its execution time by t^* in **Table 3**. As shown in **Figure 7**, the calculations

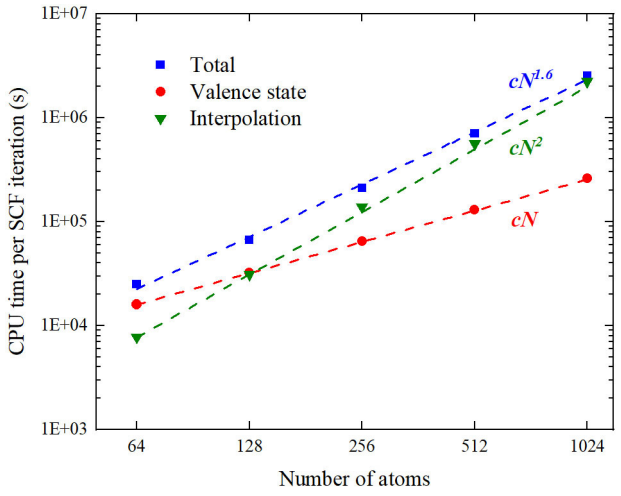

FIGURE 7 | (Color online) The log–log diagram of CPU time vs. system size where the CPU time is the product of the execution time by the number of MPI processes.

of valence states scale as $\mathcal{O}(N)$, while the interpolation from FFT uniform grid to the atom-centered radial grid exhibits an $\mathcal{O}(N^2)$ scaling, which results in an overall scaling of $\mathcal{O}(N^{1.6})$. Therefore, a more efficient algorithm to update the Coulomb potential in angular momentum expansion is critical to achieve a linear scaling FP MST method.

5. CONCLUSION

We have investigated the accuracy and scalability of the FP MST method implemented in MuST package. The MST predicted lattice parameter for V, Nb, and Mo are consistent with the other calculations and the available experiments. The MST predicted bulk moduli, pressure derivative of the bulk modulus, and the c' elastic constant are acceptable, expect for a relatively larger difference in the bulk modulus of V. While for c_{44}, there exists large difference between theoretical and experimental results, the possible reasons have been discussed in details. It is suggested that a proper treatment of the semi-core states should be considered in the future version of the MuST package.

A significant advantage of the MST method is the reduced scaling in the calculations of metallic systems. Although the linear scaling has been reported previously under the muffin-tin approximation, tests in this work imply that the overall scaling of the FP method is not $\mathcal{O}(N)$. It is suggested that the updating of the Coulomb potential in angular momentum expansion should be further improved. Nevertheless, a favorable scaling as $\mathcal{O}(N^{1.6})$ can be achieved in the full-potential MST method, compared to the $\mathcal{O}(N^2)$ to $\mathcal{O}(N^3)$ scaling of frequently-used methods. Another advantages in MuST is the treatment of chemical and magnetic disorders based on the CPA.

In summary, the FP MST method shows the potential to simulate more complicated materials on massively parallel supercomputers. And MuST provides a reliable and accessible way to large-scale first-principle simulations of metals and alloys.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank Prof. Yang Wang at Pittsburgh Supercomputing Center, USA for profound discussion of MST method, Dr. Yu Liu for helpful discussion of WIEN2k tool, and Dr. De-Ye Lin for helpful discussion of elastics. This work was partially supported by the Science Challenge Project under Grant No. TZ2018002, the National Key Research and Development Program of China under Grant No. 2016YFB0201200, and the National Natural Science Foundation of China under Grant Nos. 51771015 and 11701037.

REFERENCES

Aarons, J., Sarwar, M., Thompsett, D., and Skylaris, C.-K. (2016). Perspective: methods for large-scale density functional calculations on metallic systems. *J. Chem. Phys.* 145:220901. doi: 10.1063/1.4972007

Aarons, J., and Skylaris, C.-K. (2018). Electronic annealing Fermi operator expansion for DFT calculations on metallic systems. *J. Chem. Phys.* 148:074107. doi: 10.1063/1.5001340

Abrikosov, I. A., Niklasson, A. M. N., Simak, S. I., Johansson, B., Ruban, A. V., and Skriver, H. L. (1996). Order-*N* Green's function technique for local environment effects in alloys. *Phys. Rev. Lett.* 76, 4203–4206. doi: 10.1103/PhysRevLett.76.4203

Abrikosov, I. A., Simak, S. I., Johansson, B., Ruban, A. V., and Skriver, H. L. (1997). Locally self-consistent Green's function approach to the electronic structure problem. *Phys. Rev. B* 56, 9319–9334. doi: 10.1103/PhysRevB.56.9319

Asato, M., Settels, A., Hoshino, T., Asada, T., Blügel, S., Zeller, R., et al. (1999). Full-potential KKR calculations for metals and semiconductors. *Phys. Rev. B* 60, 5202–5210. doi: 10.1103/PhysRevB.60.5202

Ashkenazi, J., Dacorogna, M., Peter, M., Talmor, Y., Walker, E., and Steinemann, S. (1978). Elastic constants in NB-ZR alloys from zero temperature to the melting point: experiment and theory. *Phys. Rev. B* 18:4120. doi: 10.1103/PhysRevB.18.4120

Blaha, P., Schwarz, K., Tran, F., Laskowski, R., Madsen, G. K., and Marks, L. D. (2020). WIEN2K: an APW+lo program for calculating the properties of solids. *J. Chem. Phys.* 152:074101. doi: 10.1063/1.5143061

Chen, M., Jiang, X.-W., Zhuang, H., Wang, L.-W., and Carter, E. A. (2016). Petascale orbital-free density functional theory enabled by small-box algorithms. *J. Chem. Theory Comput.* 12, 2950–2963. doi: 10.1021/acs.jctc.6b00326

Chen, M., Xia, J., Huang, C., Dieterich, J. M., Hung, L., Shin, I., et al. (2015). Introducing profess 3.0: an advanced program for orbital-free density

functional theory molecular dynamics simulations. *Comput. Phys. Commun*. 190, 228–230. doi: 10.1016/j.cpc.2014.12.021

Dezerald, L., Rodney, D., Clouet, E., Ventelon, L., and Willaime, F. (2016). Plastic anisotropy and dislocation trajectory in BCC metals. *Nat. Commun*. 7:11695. doi: 10.1038/ncomms11695

Dickinson, J., and Armstrong, P. (1967). Temperature dependence of the elastic constants of molybdenum. *J. Appl. Phys*. 38, 602–606. doi: 10.1063/1.1709381

Ebert, H., Koedderitzsch, D., and Minar, J. (2011). Calculating condensed matter properties using the KKR-Green's function method–recent developments and applications. *Rep. Prog. Phys*. 74:096501. doi: 10.1088/0034-4885/74/9/096501

Economou, E. N. (2006). *Green's Functions in Quantum Physics*, Vol. 7. Berlin: Springer Science & Business Media. doi: 10.1007/3-540-28841-4

Faulkner, J., and Stocks, G. (1980). Calculating properties with the coherent-potential approximation. *Phys. Rev. B* 21:3222. doi: 10.1103/PhysRevB.21.3222

Faulkner, J., Stocks, G. M., and Wang, Y. (2018). *Multiple Scattering Theory: Electronic Structure of Solids*. Bristol: IOP Publishing. doi: 10.1088/2053-2563/aae7d8ch3

Faulkner, J. S. (1982). The modern theory of alloys. *Prog. Mater. Sci*. 27, 1–187. doi: 10.1016/0079-6425(82)90005-6

Featherston, F. H., and Neighbours, J. (1963). Elastic constants of tantalum, tungsten, and molybdenum. *Phys. Rev*. 130:1324. doi: 10.1103/PhysRev.130.1324

Frederikse, H. (1972). *American Institute of Physics Handbook*. New York, NY: McGrraw-Hill.

Gonis, A., and Butler, W. H. (2000). *Multiple Scattering in Solids*. New York, NY: Springer. doi: 10.1007/978-1-4612-1290-4

Gyorffy, B. (1972). Coherent-potential approximation for a nonoverlapping-muffin-tin-potential model of random substitutional alloys. *Phys. Rev. B* 5:2382. doi: 10.1103/PhysRevB.5.2382

Haas, P., Tran, F., and Blaha, P. (2009). Calculation of the lattice constant of solids with semilocal functionals. *Phys. Rev. B* 79:085104. doi: 10.1103/PhysRevB.79.085104

Ho, G. S., Lignéres, V. L., and Carter, E. A. (2008). Introducing profess: a new program for orbital-free density functional theory calculations. *Comput. Phys. Commun*. 179, 839–854. doi: 10.1016/j.cpc.2008.07.002

Holzwarth, N. A. W., Matthews, G. E., Dunning, R. B., Tackett, A. R., and Zeng, Y. (1997). Comparison of the projector augmented-wave, pseudopotential, and linearized augmented-plane-wave formalisms for density-functional calculations of solids. *Phys. Rev. B* 55:2005. doi: 10.1103/PhysRevB.55.2005

Hung, L., Huang, C., Shin, I., Ho, G. S., Lignéres, V. L., and Carter, E. A. (2010). Introducing profess 2.0: a parallelized, fully linear scaling program for orbital-free density functional theory calculations. *Comput. Phys. Commun*. 181, 2208–2209. doi: 10.1016/j.cpc.2010.09.001

Koči, L., Ma, Y., Oganov, A., Souvatzis, P., and Ahuja, R. (2008). Elasticity of the superconducting metals V, Nb, Ta, Mo, and W at high pressure. *Phys. Rev. B* 77:214101. doi: 10.1103/PhysRevB.77.214101

Kohn, W., and Rostoker, N. (1954). Solution of the Schrödinger equation in periodic lattices with an application to metallic lithium. *Phys. Rev*. 94, 1111–1120. doi: 10.1103/PhysRev.94.1111

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev*. 140, A1133–A1138. doi: 10.1103/PhysRev.140.A1133

Korringa, J. (1947). On the calculation of the energy of a Bloch wave in a metal. *Physica* 13, 392–400. doi: 10.1016/0031-8914(47)90013-X

Kresse, G., and Furthmüller, J. (1996a). Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mat. Sci*. 6, 15–50. doi: 10.1016/0927-0256(96)00008-0

Kresse, G., and Furthmüller, J. (1996b). Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* 54, 11169–11186. doi: 10.1103/PhysRevB.54.11169

Kresse, G., and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* 59, 1758–1775. doi: 10.1103/PhysRevB.59.1758

Kucukbenli, E., Monni, M., Adetunji, B., Ge, X., Adebayo, G., Marzari, N., et al. (2014). Projector augmented-wave and all-electron calculations across the periodic table: a comparison of structural and energetic properties. *arXiv preprint arXiv:1404.3015*.

Landa, A., Klepeis, J., Söderlind, P., Naumov, I., Velikokhatnyi, O., Vitos, L., et al. (2006a). *Ab initio* calculations of elastic constants of the bcc V-Nb system at high pressures. *J. Phys. Chem. Solids* 67, 2056–2064. doi: 10.1016/j.jpcs.2006.05.027

Landa, A., Klepeis, J., Söderlind, P., Naumov, I., Velikokhatnyi, O., Vitos, L., et al. (2006b). Fermi surface nesting and pre-martensitic softening in V and Nb at high pressures. *J. Phys. Cond. Matter* 18:5079. doi: 10.1088/0953-8984/18/22/008

Lejaeghere, K., Bihlmayer, G., Björkman, T., Blaha, P., Blügel, S., Blum, V., et al. (2016). Reproducibility in density functional theory calculations of solids. *Science* 351:6280. doi: 10.1126/science.aad3000

Lejaeghere, K., Van Speybroeck, V., Van Oost, G., and Cottenier, S. (2014). Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals. *Crit. Rev. Solid State Mater. Sci*. 39, 1–24. doi: 10.1080/10408436.2013.772503

Liu, X., Wang, Y., Eisenbach, M., and Stocks, G. M. (2016). A full-potential approach to the relativistic single-site green's function. *J. Phys. Cond. Matter* 28:355501. doi: 10.1088/0953-8984/28/35/355501

Liu, X., Wang, Y., Eisenbach, M., and Stocks, G. M. (2018). Fully-relativistic full-potential multiple scattering theory: a pathology-free scheme. *Comput. Phys. Commun*. 224, 265–272. doi: 10.1016/j.cpc.2017.10.011

Maschke, K., and Lévy, F. (1983). *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, eds K.-H. Hellwege and O. Madelung, New Series, Group III: Crystal and Solid State Physics, (Berlin: Springer).

Mohr, S., Eixarch, M., Amsler, M., Mantsinen, M. J., and Genovese, L. (2018). Linear scaling DFT calculations for large Tungsten systems using an optimized local basis. *Nuclear Mater. Energy* 15, 64–70. doi: 10.1016/j.nme.2018.01.002

Nagasako, N., Jahnátek, M., Asahi, R., and Hafner, J. (2010). Anomalies in the response of V, Nb, and Ta to tensile and shear loading: *ab initio* density functional theory calculations. *Phys. Rev. B* 81:094108. doi: 10.1103/PhysRevB.81.094108

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett*. 77:3865. doi: 10.1103/PhysRevLett.77.3865

Pratapa, P. P., Suryanarayana, P., and Pask, J. E. (2016). Spectral Quadrature method for accurate $\mathcal{O}(N)$ electronic structure calculations of metals and insulators. *Comp. Phys. Commun*. 200, 96–107. doi: 10.1016/j.cpc.2015.11.005

Ruban, A. V., and Abrikosov, I. A. (2008). Configurational thermodynamics of alloys from first principles: effective cluster interactions. *Rep. Prog. Phys*. 71:046501. doi: 10.1088/0034-4885/71/4/046501

Rusanu, A., Stocks, G. M., Wang, Y., and Faulkner, J. S. (2011). Green's functions in full-potential multiple-scattering theory. *Phys. Rev. B* 84:035102. doi: 10.1103/PhysRevB.84.035102

Sébilleau, D. (2000). Basis-independent multiple-scattering theory for electron spectroscopies: general formalism. *Phys. Rev. B* 61:14167. doi: 10.1103/PhysRevB.61.14167

Singh, D. J., and Nordström, L. (2006). *Planewaves, Pseudopotentials, and the LAPW Method*, 2nd Edn, New York, NY: Springer Science & Business Media.

Söderlind, P., Yang, L., Moriarty, J. A., and Wills, J. (2000). First-principles formation energies of monovacancies in bcc transition metals. *Phys. Rev. B* 61:2579. doi: 10.1103/PhysRevB.61.2579

Soven, P. (1967). Coherent-potential model of substitutional disordered alloys. *Phys. Rev*. 156:809. doi: 10.1103/PhysRev.156.809

Suryanarayana, P., Pratapa, P., P., Sharma, A., and Pask, J. E. (2018). SQDFT: Spectral quadrature method for large-scale parallel$\mathcal{O}(N)$ Kohn-Sham calculations at high temperature. *Comp. Phys. Commun*. 224, 288–298. doi: 10.1016/j.cpc.2017.12.003

Taylor, D. (1967). Vibrational properties of imperfect crystals with large defect concentrations. *Phys. Rev*. 156:1017. doi: 10.1103/PhysRev.156.1017

Thiess, A., Zeller, R., Bolten, M., Dederichs, P. H., and Blügel, S. (2012). Massively parallel density functional calculations for thousands of atoms: KKRnano. *Phys. Rev. B* 85:235103. doi: 10.1103/PhysRevB.85.235103

Turek, I., Drchal, V., Kudrnovsky, J., Sob, M., and Weinberger, P. (1997). *Electronic Structure of Disordered Alloys, Surfaces and Interfaces*. New York, NY: Springer Science & Business Media. doi: 10.1007/978-1-4615-6255-9

Vitos, L. (2001). Total-energy method based on the exact muffin-tin orbitals theory. *Phys. Rev. B* 64:014107. doi: 10.1103/PhysRevB.64.014107

Vitos, L. (2007). *Computational Quantum Mechanics for Materials Engineers*. London: Springer.

Vitos, L., Abrikosov, I. A., and Johansson, B. (2001). Anisotropic lattice distortions in random alloys from first-principles theory. *Phys. Rev. Lett*. 87:156401. doi: 10.1103/PhysRevLett.87.156401

Wang, Y., Stocks, G. M., Shelton, W. A., Nicholson, D. M. C., Szotek, Z., and Temmerman, W. M. (1995). Order-*n* multiple scattering approach to electronic structure calculations. *Phys. Rev. Lett*. 75, 2867–2870. doi: 10.1103/PhysRevLett.75.2867

Young, D. A. (1991). *Phase Diagrams of the Elements*. Berkeley: Univ of California Press.

Zabloudil, J., Hammerling, R., Szunyogh, L., and Weinberger, P. (2006). *Electron Scattering in Solid Matter: A Theoretical and Computational Treatise*, Vol. 147. Berlin: Springer Science & Business Media. doi: 10.1007/b138290

Zeller, R., Dederichs, P. H., Újfalussy, B., Szunyogh, L., and Weinberger, P. (1995). Theory and convergence properties of the screened Korringa-Kohn-Rostoker method. *Phys. Rev. B* 52, 8807–8812. doi: 10.1103/PhysRevB.52.8807

Zhang, W., Thiess, A., Zalden, P., Zeller, R., Dederichs, P. H., Raty, J. Y., et al. (2012). Role of vacancies in metal-insulator transitions of crystalline phase-change materials. *Nat. Mater*. 11, 952–956. doi: 10.1038/nmat3456

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 08 December 2020
doi: 10.3389/fchem.2020.588808

Density Functional Prediction of Quasiparticle, Excitation, and Resonance Energies of Molecules With a Global Scaling Correction Approach

Xiaolong Yang[1], *Xiao Zheng*[1]* *and Weitao Yang*[2,3]*

[1] *Hefei National Laboratory for Physical Sciences at the Microscale and Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, China,* [2] *Department of Chemistry, Duke University, Durham, NC, United States,* [3] *Key Laboratory of Theoretical Chemistry of Environment, School of Chemistry and Environment, South China Normal University, Guangzhou, China*

Molecular quasiparticle and excitation energies determine essentially the spectral characteristics measured in various spectroscopic experiments. Accurate prediction of these energies has been rather challenging for ground-state density functional methods, because the commonly adopted density function approximations suffer from delocalization error. In this work, by presuming a quantitative correspondence between the quasiparticle energies and the generalized Kohn–Sham orbital energies, and employing a previously developed global scaling correction approach, we achieve substantially improved prediction of molecular quasiparticle and excitation energies. In addition, we also extend our previous study on temporary anions in resonant states, which are associated with negative molecular electron affinities. The proposed approach does not require any explicit self-consistent field calculation on the excited-state species, and is thus highly efficient and convenient for practical purposes.

Keywords: density functional theory, delocalization error, scaling correction approach, quasiparticle energies, electronic excitation energies, electron affinity

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Bingbing Suo,
Northwest University, China
Jun Cheng,
Xiamen University, China

***Correspondence:**
Xiao Zheng
xz58@ustc.edu.cn
Weitao Yang
weitao.yang@duke.edu

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *29 July 2020*
Accepted: *23 September 2020*
Published: *08 December 2020*

Citation:
Yang X, Zheng X and Yang W (2020) Density Functional Prediction of Quasiparticle, Excitation, and Resonance Energies of Molecules With a Global Scaling Correction Approach. Front. Chem. 8:588808. doi: 10.3389/fchem.2020.588808

1. INTRODUCTION

Density function theory (DFT) (Hohenberg and Kohn, 1964) has made great success in practical calculations for ground-state electronic properties because of its outstanding balance between accuracy and computational cost. In the Kohn–Sham (KS) scheme of DFT (Hohenberg and Kohn, 1964; Kohn and Sham, 1965), the effective single-particle equations can be written as (by omitting the spin indices and adopting the atomic units)

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r})\right]\phi_m(\mathbf{r}) = \varepsilon_m\,\phi_m(\mathbf{r}). \tag{1}$$

Here, $v_{\mathrm{ext}}(\mathbf{r})$ is the external potential, $v_{\mathrm{H}}(\mathbf{r})$ is the Hartree potential, $v_{\mathrm{xc}}(\mathbf{r})$ is the local exchange-correlation (XC) potential, and $\{\phi_m(\mathbf{r})\}$ and $\{\varepsilon_m\}$ are the KS/generalized KS (GKS) orbitals and their eigenvalues, respectively. In the GKS scheme, $v_{\mathrm{xc}}(\mathbf{r})$ is replaced by a non-local potential, $v_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$. The KS equations can be solved self-consistently to produce the ground-state

energy and charge density. However, it is challenging to apply conventional ground-state density functional methods to calculate excited-state-related properties, such as the quasiparticle (QP) energies and the electronic excitation energies, which will be introduced as follows.

When an electronic system is perturbed by incoming photons or electrons, in order to preserve a single-particle picture, the concept of QP is often adopted. In a direct photoemission experiment, an electron on a molecule absorbs the energy of a photon and gets excited away from the molecule. Such a process leaves a quasihole in the molecule whose energy level is renormalized by the presence of the other electrons. Similarly, in an inverse photoemission experiment, an electron attaches to a molecule by emitting a photon, which leads to the formation of a quasielectron whose energy level is influenced by the existing electrons in the molecule (Onida et al., 2002).

The actual QP energies and wavefunctions can be obtained by solving the QP equations as follows (Hedin, 1965; Aulbur et al., 2000),

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{ext}}(\mathbf{r})\right]\psi_m(\mathbf{r}) + \int \Sigma(\mathbf{r},\mathbf{r}';\omega_m)\,\psi_m(\mathbf{r}')d\mathbf{r}' = \omega_m\,\psi_m(\mathbf{r}). \quad (2)$$

Here, $\{\psi_m(\mathbf{r})\}$ and $\{\omega_m\}$ are the QP wavefunctions and energies, respectively, and Σ is a non-local and energy-dependent self-energy operator, with the imaginary part of its eigenvalues giving the lifetime of the QPs. To enable practical calculations, an approximate scheme for Σ is to be employed. The most widely used scheme is the *GW* approximation (Hedin, 1965; Hybertsen and Louie, 1986; Aulbur et al., 2000; Dvorak et al., 2014). Therefore, regarding the calculation of QP energies, many-body perturbation theory within the *GW* approximation has become a popular method at present (Hedin, 1965; Hybertsen and Louie, 1986; Louie and Hybertsen, 1987; Aulbur et al., 2000; Onida et al., 2002; Dvorak et al., 2014). However, the somewhat large computational cost makes it difficult to apply the *GW* method to complex systems. Thus, a highly efficient and accurate method for the prediction of QP energies is sought for.

It is tempting to relate the KS/GKS orbital energies to QP energies, because the KS and GKS schemes are in conformity with an effective single-electron description. However, with conventional density functional approximations (DFAs), such as the local density approximation (LDA) (Slater, 1951; Vosko et al., 1980), generalized gradient approximations (GGAs), and hybrid functionals, the calculated KS/GKS orbital energies usually deviate severely from the QP energies. Such deviations have also led to significant underestimation of band gaps, which is largely due to the delocalization error associated with the DFAs (Cohen et al., 2008a). In the exact DFT, the ground-state energy of a system with a fractional number of electrons, $E_0(N+n)$ (subscript 0 denotes the ground state corresponding to the fixed v_{ext}), should satisfy the Perdew–Parr–Levy–Balduz (PPLB) condition (Perdew et al., 1982, 2007; Yang et al., 2000): $E_0(N+n) = (1-n)E_0(N) + nE_0(N+1)$, where $0 < n < 1$ is a fractional number.

The PPLB condition infers that $(\frac{\partial E_0}{\partial N})_- = -I$ and $(\frac{\partial E_0}{\partial N})_+ = -A$, where $I \equiv E_0(N-1) - E_0(N)$ and $A \equiv E_0(N) - E_0(N+1)$ are the vertical ionization potential (IP) and electron affinity (EA) of the N-electron system, respectively. It has been proved (Cohen et al., 2008a; Yang et al., 2012) that if the XC energy is an explicit and differentiable functional of the electron density or the KS reduced density matrix, we have $(\frac{\partial E_0}{\partial N})_- = \varepsilon_{\mathrm{HOMO}}$ and $(\frac{\partial E_0}{\partial N})_+ = \varepsilon_{\mathrm{LUMO}}$, where $\varepsilon_{\mathrm{HOMO}}$ and $\varepsilon_{\mathrm{LUMO}}$ are the energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), respectively. Therefore, if the PPLB condition can be satisfied, we should have $I = -\varepsilon_{\mathrm{HOMO}}$ and $A = -\varepsilon_{\mathrm{LUMO}}$.

Within the framework of ground-state DFT, a rigorous mapping between the other remaining KS/GKS orbital energies apart from the HOMO and LUMO and the QP energies has not been established. However, in practice the Koopmans-like relations have been proposed and adopted by many authors (Hill et al., 2000; Coropceanu et al., 2002; Vargas et al., 2005; Bartlett, 2009; Gritsenko and Baerends, 2009; Tsuneda et al., 2010; Dauth et al., 2011; Körzdörfer et al., 2012; Baerends et al., 2013; Bartlett and Ranasinghe, 2017; Puschnig et al., 2017; Ranasinghe et al., 2017; Thierbach et al., 2017). These relations have the form of $\varepsilon_i \approx -I_i^{\mathrm{v}} = -[E_i(N-1) - E_0(N)]$ and $\varepsilon_a \approx -A_a^{\mathrm{v}} = -[E_0(N) - E_a(N+1)]$. Here, the index i (a) denotes the occupied (virtual) KS/GKS orbital of the N-electron system from (to) which an electron is deprived (added), with I_i^{v} (A_a^{v}) being the corresponding vertical IP (EA). It is easily recognized that these vertical IPs and EAs coincide with the energies of quasiholes and quasielectrons, i.e., $\omega_i = -I_i^{\mathrm{v}}$ and $\omega_a = -A_a^{\mathrm{v}}$, respectively. Computationally, approximating QP energies by KS/GKS orbital energies has the advantage of requiring only a single self-consistent field (SCF) calculation for the ground state of the interested molecule.

The excited-state properties of molecular systems can be probed by photon absorption experiments (Onida et al., 2002). However, theoretical characterization of the excited states is rather challenging because the excited electron and the resulting hole cannot be treated separately. Numerous methods have been developed for the calculation of excitation energies. Coupled cluster (CC) (Schreiber et al., 2008; Silva-Junior et al., 2008; Winter et al., 2013; Wang et al., 2014; Dreuw and Wormit, 2015; Jacquemin et al., 2015) and multi-reference methods (Andersson et al., 1990; Potts et al., 2001; Slavicek and Martinez, 2010; Hoyer et al., 2016) are able to describe electronic excited states with a high accuracy. However, the expensive computational cost makes the application of these methods to large systems rather difficult. As a straightforward extension of the *GW* approach (Hedin, 1965; Hybertsen and Louie, 1986; Onida et al., 2002), the Bethe–Salpeter equation (BSE) (Rohlfing and Louie, 2000; Onida et al., 2002; Jacquemin et al., 2017) provides another method for the calculation of excited states, which is however also quite expensive. The time-dependent DFT (TDDFT) (Runge and Gross, 1984; Casida, 1995) is in principle an exact extension of the ground-state DFT, and it has been widely employed to study neutral excitations. Despite its success, TDDFT faces several challenges, such as double excitation character, multi-reference problems, and high-spin excited states (Ipatov et al.,

2009; Laurent and Jacquemin, 2013; Santoro and Jacquemin, 2016).

Recently, a simple method has been proposed, which attempts to acquire excitation energies by using only the KS/GKS orbital energies of the molecular cations calculated by ground-state DFT (Haiduke and Bartlett, 2018; Mei et al., 2019). Such a method is referred to as the QE-DFT (QP energies from DFT), which has been employed to describe excited-state potential energy surfaces and conical intersections (Mei and Yang, 2019). Following the idea of QE-DFT, molecular excitation energies have also been expressed by KS/GKS orbital energies obtained with long-range corrected functionals (Hirao et al., 2020). Details about QE-DFT are to be presented in section 2.1.

In addition to neutral molecules, in this work we also consider the resonance states of temporary anions. A temporary anion has an energy higher than that of the neutral species, and thus its EA has a negative value. This means the anion is unstable and lasts only a short time. Although temporary anions cannot be studied by traditional spectroscopic techniques, they can be observed via resonances (sharp variations) in the cross-sections of various electron scattering processes (Jordan and Burrow, 1987). In the gas phase, the resonances can be identified by the electron transmission spectroscopy (Sanche and Schulz, 1972; Schulz, 1973; Jordan and Burrow, 1987). Since these resonances belong to the continuous part of the spectrum, they cannot be addressed by conventional electronic structure methods for bound states. A number of theoretical methods have been proposed to tackle the problem of temporary anions. For instance, it has been proposed that the attractive components of electron-molecule interaction are combined with a long-range repulsive potential to produce a barrier, behind which the excess electron can be temporarily trapped (Jordan et al., 2014). Moreover, the negative EAs have been studied by the *GW* method (Hedin, 1965; Hybertsen and Louie, 1986; Govoni and Galli, 2018), the electron-propagator methods (Longo et al., 1995; Ortiz, 2013; Dolgounitcheva et al., 2016), and the equation-of-motion coupled cluster (EOM-CC) approach (Stanton and Bartlett, 1993; Nooijen and Bartlett, 1995; Dutta et al., 2014; Jagau et al., 2017; Skomorowski et al., 2018; Ma et al., 2020), again with considerable computational cost.

In order to describe the unbound resonance states within the DFT approach, Tozer and De Proft (2005) have proposed an approximate approach to evaluate the EA based on the KS frontier orbital energies (Kohn and Sham, 1965) and the accurate IP. Zhang et al. (2018) have used directly the negative of GKS eigenvalue of the neutral ground-state molecule as an approximation of EA corresponding to the resonance state of the anion. The good accuracy was made possible because of the use of the global scaling correction (GSC) (Zheng et al., 2011), which will be introduced later. At the same time, another method has been developed to evaluate the negative EA from the GKS eigenvalue of the neutral ground states (Carmona-Espíndola et al., 2020). Different from GSC, this method is designed to impose the derivative discontinuity of the exact XC potential. Our work (Zhang et al., 2018) proceeded the work of Mei et al. (2019) and that of Haiduke and Bartlett (Haiduke and Bartlett, 2018) in the direct use of GKS eigenvalues of the N-electron ground state to approximate the excited state energy of the corresponding $(N+1)$-electron system, with the excited state of the $(N+1)$-electron system being a unbound resonance state.

For achieving an accurate prediction of QP energies with ground-state density functional methods, it is crucial to reduce the delocalization error associated with the adopted DFA. Enormous efforts have been made, which have led to the development of the GSC (Zheng et al., 2011) and local scaling correction (LSC) (Li et al., 2015) approaches, which alleviate substantially the delocalization error of various DFAs for systems involving global and local fractional electron distributions, respectively. This is done by imposing explicitly the PPLB condition on the form of DFA. Recently, a localized orbital scaling correction (LOSC) (Li et al., 2017; Su et al., 2020) has been constructed to join the merits of GSC and LSC. The LOSC approach is capable of correcting the energy, energy derivative, and electron density of any finite system in a self-consistent and size-consistent manner. In particular, the LOSC approach has been applied in conjunction with the QE-DFT to predict QP and excitation energies of molecules (Mei et al., 2019).

In this work, we revisit the non-empirical GSC approach (Zheng et al., 2011, 2013, 2015; Zhang et al., 2015) and explore the possibility of using it to achieve an accurate prediction of QP and excitation energies. With a perturbative treatment of the orbital relaxation induced by the addition (removal) of an infinitesimal amount of electron to (from) a molecule, the GSC approach has been demonstrated to improve systematically the prediction of KS frontier orbital energies and band gaps of molecules (Zhang et al., 2015). Based on the idea of QE-DFT, we will extend the scope of GSC from the frontier orbitals to the other KS/GKS orbitals.

The remainder of this paper is organized as follows. In section 2, we present the QE-DFT method to calculate the QP and excitation energies within the framework of ground-state DFT, as well as the GSC approach to achieve the accurate KS/GKS orbital energies. In section 3, numerical results of the QP energies, electronic excitation energies, and resonance energies are presented and discussed. Finally, we summarize this work in section 4.

2. METHODOLOGY

2.1. QE-DFT Method for the Calculation of QP, Excitation, and Resonance Energies

In the QE-DFT method, the following Koopmans-like relations are adopted, which use the energies of occupied and virtual KS/GKS orbitals to approximate the quasihole and quasielectron energies, respectively.

$$\begin{aligned} \varepsilon_a(N) \approx \omega_a(N) &= E_a(N+1) - E_0(N), \\ \varepsilon_i(N) \approx \omega_i(N) &= E_0(N) - E_i(N-1). \end{aligned} \tag{3}$$

Here, $\{\varepsilon_i(N)\}$ and $\{\varepsilon_a(N)\}$ are the occupied and virtual orbital energies of the N-electron system, respectively. $E_a(N+1)$ denotes the energy of the $(N+1)$-electron system formed by adding an excess electron to the ath virtual orbital of the N-electron system at its ground state. Note that the subscript a refers to the N-electron system, and the value of such an orbital index may

vary in the $(N+1)$-electron system because of the relaxation and re-ordering of the orbitals upon the perturbation induced by electron addition. A similar argument applies to the ith occupied orbital of the N-electron system.

From Equation (3), it is obvious that the excitation energy of an electron from the HOMO to a virtual orbital of the N-electron system, which corresponds to the ath orbital of the $(N-1)$-electron system, can be calculated as follows (Haiduke and Bartlett, 2018; Mei et al., 2019):

$$\begin{aligned}\Delta E_a(N) &\equiv E_a(N) - E_0(N)\\ &= [E_a(N) - E_0(N-1)] - [E_0(N) - E_0(N-1)]\\ &= \omega_a(N-1) - \omega_{\mathrm{LUMO}}(N-1)\\ &\approx \varepsilon_a(N-1) - \varepsilon_{\mathrm{LUMO}}(N-1).\end{aligned} \tag{4}$$

Here, in the second equality of Equation (4), we have chosen to use the $(N-1)$-electron system as a reference system. This means the electronic excitation from the HOMO to a virtual orbital can be regarded as consisting of two steps: first an electron is removed from the HOMO of the N-electron system, giving rise to an $(N-1)$-electron system in its ground state, and then an excess electron is put to the ath virtual orbital of the $(N-1)$-electron system, which is energetically higher than the frontier orbitals, resulting in an excited N-electron system. Accordingly, $E_0(N-1)$ is the ground-state energy of the $(N-1)$-electron system, and $E_a(N)$ denotes the energy of the N-electron system that is finally obtained. Thus, Equation (4) can describe the excitation of an electron from the HOMO to any virtual KS/GKS orbital (LUMO and above), as long as the orbital finds its counterpart in the $(N-1)$-electron system.

Specifically, if we presume the $(N-1)$-electron reference system contains one more spin-α electrons than spin-β electrons, the first triplet-state excitation energy of the N-electron system is calculated by

$$\Delta E_{\mathrm{T1}}(N) \equiv E_{\mathrm{T1}}(N) - E_0(N) \approx \varepsilon_{\alpha,\mathrm{LUMO}}(N-1) - \varepsilon_{\beta,\mathrm{LUMO}}(N-1). \tag{5}$$

Higher triplet-state excitation energies can be calculated similarly.

We now consider the first singlet excited state formed by adding a spin-β electron to the ath virtual orbital of the ground-state $(N-1)$-electron system. It is well-known that the accurate calculation of open-shell singlet states is quite challenging for the density functional methods within the KS/GKS scheme. This is because the electronic wavefunction naturally involves more than one Slater determinant, and such a multireference character is hardly captured by the presently used DFAs due to their intrinsic static correlation error (Cohen et al., 2008b). Moreover, an open-shell singlet wavefunction in the form of a single Slater determinant of KS/GKS orbitals is not an eigenstate of the total spin operator. In practice, people have attempted to circumvent the problem of static correlation error by explicitly using more than one Slater determinant. For instance, the singlet-state energy of an N-electron system has been written as (Ess et al., 2011)

$$E_{\mathrm{S}}(N) = E_{\mathrm{M}}(N) + \chi[E_{\mathrm{M}}(N) - E_{\mathrm{T}}(N)]. \tag{6}$$

Here, $E_{\mathrm{M}}(N)$ represents the energy of a single-Slater-determinant wavefunction with the excited spin-β electron occupying the virtual orbital. The second term on the right-hand side is a correction to the singlet-state energy, which accounts for the spin contamination of the single-Slater-determinant wavefunction, with χ being a parameter. The singlet-state excitation energy of the N-electron is thus obtained as

$$\begin{aligned}\Delta E_{\mathrm{S}}(N) &\equiv E_{\mathrm{S}}(N) - E_0(N)\\ &= [E_{\mathrm{M}}(N) - E_0(N)] + \chi[E_{\mathrm{M}}(N) - E_{\mathrm{T}}(N)]\\ &\approx [\varepsilon_{\beta,\mathrm{LUMO}+a}(N-1) - \varepsilon_{\beta,\mathrm{LUMO}}(N-1)]\\ &\quad + \chi[\varepsilon_{\beta,\mathrm{LUMO}+a}(N-1) - \varepsilon_{\alpha,\mathrm{LUMO}+a}(N-1)].\end{aligned} \tag{7}$$

To improve the accuracy of ΔE_{S}, a spin purification procedure has been proposed (Ziegler et al., 1977), which amounts to $\chi = 1$ in Equation (6). Specifically, the first singlet-state excitation energy is calculated by

$$\begin{aligned}\Delta E_{\mathrm{S1}}(N) \equiv E_{\mathrm{S1}}(N) - E_0(N) \approx 2\varepsilon_{\beta,\mathrm{LUMO}+1}(N-1)\\ -\varepsilon_{\alpha,\mathrm{LUMO}}(N-1) - \varepsilon_{\beta,\mathrm{LUMO}}(N-1).\end{aligned} \tag{8}$$

Obviously, with the QE-DFT method, the calculation of excitation energies requires the SCF calculations to be performed explicitly only for the ground-state $(N-1)$-electron system.

Regarding temporary anions, we only consider the scenario that the LUMO of the neutral molecule is already an unbound orbital, which corresponds to a negative EA. Consequently, addition of an excess electron to the LUMO gives rise to a resonant state. Traditionally, the molecular EA is obtained by performing SCF calculations separately for the neutral and anionic systems and taking the energy difference between them. This is referred to the ΔSCF method. However, in practice it is extremely difficult to carry out an SCF calculation for the anionic species if it is in a resonant state.

Since the LUMO is a frontier orbital, the PPLB condition holds exactly, and thus the negative EA can be obtained directly from the positive LUMO energy via the following equality:

$$A = -\varepsilon_{\mathrm{LUMO}}(N). \tag{9}$$

By using Equation (9), the SCF calculation on the temporary anion that is potentially problematic is no longer needed.

2.2. GSC Approach for the Accurate Prediction of KS/GKS Orbital Energies

From section 2.1, the prediction of QP and excitation energies is transformed to the accurate calculation of KS/GKS orbital energies. To this end, we employ a non-empirical GSC approach developed in our previous works Zheng et al. (2013), Zhang et al. (2015) to reduce the delocalization error of some frequently used DFAs. It has been demonstrated that the GSC approach greatly improves the accuracy of the frontier KS/GKS orbitals. In the following, we shall go beyond the frontier orbitals and extend the application of GSC to all the KS/GKS orbitals.

In the KS or GKS scheme, the total electronic energy in the ground state is $E_0(N) = T_{\mathrm{s}} + V_{\mathrm{ext}} + J + E_{\mathrm{xc}}$. With the KS/GKS

orbitals fixed as the electron number is varied, the KS kinetic energy T_s and external potential energy V_{ext} are linear in $\rho(\mathbf{r})$, while the electron Coulomb energy J is quadratic, and E_{xc} is usually non-linear in $\rho(\mathbf{r})$. The GSC approach establishes a linear energy function that satisfies the PPLB condition,

$$\tilde{E}_0(N+n) \equiv (1-n)E_0(N) + nE_0(N+1), \tag{10}$$

by linearizing both J and E_{xc} with respect to the fractional electron number n. The difference between $\tilde{E}_0$ and E_0 is just the GSC for the energy:

$$\Delta E_0^{\mathrm{GSC}} = \tilde{E}_0(N+n) - E_0(N+n). \tag{11}$$

Here, $\Delta E_0^{\mathrm{GSC}}$ can express explicitly by the electron density $\rho(\mathbf{r}) = \sum_m n_m[\phi_m(\mathbf{r})]^2$ and some other quantities, where $\phi_m(\mathbf{r})$ and n_m are the mth KS/GKS orbital and electron occupation number, respectively. For simplicity, the spin indices are omitted.

The addition of the n fractional electron to the LUMO presents a perturbation to the N-electron system, and the change in electron density in response to such a perturbation is

$$\delta\rho(\mathbf{r}) = \rho^{N+n}(\mathbf{r}) - \rho^N(\mathbf{r}) = nf(\mathbf{r}) + n^2\gamma(\mathbf{r}) + \cdots, \tag{12}$$

where $f(\mathbf{r}) \equiv \lim_{n\to 0} \frac{\partial\rho(\mathbf{r})}{\partial n}|_{v_{ext}}$ and $\gamma(\mathbf{r}) \equiv \lim_{n\to 0} \frac{1}{2}\frac{\partial^2\rho(\mathbf{r})}{\partial n^2}|_{v_{ext}}$ are the first- and second-order Fukui functions (Parr and Yang, 1984; Yang et al., 1984; Yang and Parr, 1985), respectively. Accordingly, the relaxation of KS/GKS orbitals upon the addition of n fractional electron can be expanded in a perturbative series as $\delta\phi_m(\mathbf{r}) = \phi_m^{N+n}(\mathbf{r}) - \phi_m^N(\mathbf{r}) = n\delta\phi_m^{(1)}(\mathbf{r}) + n^2\delta\phi_m^{(2)}(\mathbf{r}) + \cdots$, with $\delta\phi_m^{(k)}(\mathbf{r})$ being the kth-order orbital relaxation. Thus, the Fukui functions can be expressed explicitly in terms of orbital relaxation. For instance, the first-order Fukui function is

$$f(\mathbf{r}) = |\phi_f(\mathbf{r})|^2 + 2\sum_m n_m\delta\phi_m^{(1)}(\mathbf{r})\,\phi_m(\mathbf{r}). \tag{13}$$

Here, the subscript f denotes the frontier orbital, with f = LUMO (f = HOMO) in the case of electron addition (removal). The explicit forms of orbital relaxation up to the third order have been derived and provided in Zhang et al. (2015), with all the perturbation Hamiltonian matrices determined by a self-consistent process. Ultimately, all orders of orbital relaxation and Fukui quantities are expressed in terms of $\{\phi_m(\mathbf{r})\}$ and $\{\varepsilon_m\}$ of the N-electron system. The scaling correction to the frontier orbital energy is then evaluated by the Janak's theorem (Janak, 1978) in a post-SCF manner,

$$\Delta\varepsilon_f^{\mathrm{GSC}} = \frac{\partial\Delta E_0^{\mathrm{GSC}}}{\partial n} = \Delta\varepsilon_f^{(1)} + \Delta\varepsilon_f^{(2)} + \cdots, \tag{14}$$

where $\Delta\varepsilon_f^{(k)}$ is the kth-order correction to the frontier orbital energy.

An accurate prediction of molecular IP and EA has been achieved by employing the GSC approach (Zheng et al., 2013, 2015; Zhang et al., 2015, 2018) via

$$I = -\varepsilon_{\mathrm{HOMO}}^{\mathrm{GSC-DFA}} = -(\varepsilon_{\mathrm{HOMO}}^{\mathrm{DFA}} + \Delta\varepsilon_{\mathrm{HOMO}}^{\mathrm{GSC}}), \tag{15}$$

$$A = -\varepsilon_{\mathrm{LUMO}}^{\mathrm{GSC-DFA}} = -(\varepsilon_{\mathrm{LUMO}}^{\mathrm{DFA}} + \Delta\varepsilon_{\mathrm{LUMO}}^{\mathrm{GSC}}). \tag{16}$$

In practical calculations, the perturbative series needs to be truncated at a certain order. It is worth pointing out that the accuracy of the GSC does not necessarily increase with further inclusion of higher order orbital relaxation. This is because the present form of GSC only treats the exchange energy E_x, while the correlation energy E_c is presumed to be much smaller and hence its correction is omitted. However, the correction to E_c may have a comparable magnitude to the high-order corrections to E_x. For instance, regarding the prediction of EA, while the inclusion of first-order orbital relaxation is found optimal for the LDA and GGA (such as BLYP, Becke, 1988; Lee et al., 1988), the inclusion of orbital relaxation up to second-order is most favorable for the hybrid functional B3LYP (Lee et al., 1988; Becke, 1993).

To extend the GSC approach beyond the frontier KS/GKS orbitals, we presume that the PPLB condition could be generalized to the following energy linearity relation:

$$\tilde{E}_a(N+n) \equiv (1-n)E_0(N) + nE_a(N+1). \tag{17}$$

The GSC to the energy of the $(N+n)$-electron system is

$$\Delta E_a^{\mathrm{GSC}} = \tilde{E}_a(N+n) - E_a(N+n), \tag{18}$$

where $E_a(N+n)$ is the energy of the $(N+n)$-electron system in an excited state, since the n fractional electron is now added to the ath virtual orbital of the N-electron system. Similarly, the changes of electron density and KS/GKS orbitals in response to the perturbation caused by the electron addition process, as well as their contributions to $\Delta E_a^{\mathrm{GSC}}$, are calculated by using the self-consistent perturbation theory presented in Zhang et al. (2015). This finally gives rise to the GSC to the KS/GKS orbital energies:

$$\Delta\varepsilon_a^{\mathrm{GSC}} = \frac{\partial\Delta E_a^{\mathrm{GSC}}}{\partial n} = \Delta\varepsilon_a^{(1)} + \Delta\varepsilon_a^{(2)} + \cdots. \tag{19}$$

Likewise, for the scenario that n fractional electron is deprived from the ith occupied orbital of the N-electron system, we have

$$\Delta\varepsilon_i^{\mathrm{GSC}} = \frac{\partial\Delta E_i^{\mathrm{GSC}}}{\partial n} = \Delta\varepsilon_i^{(1)} + \Delta\varepsilon_i^{(2)} + \cdots. \tag{20}$$

With the QE-DFT method, we can now use the scaling corrected KS/GKS orbital energies to approximate the QP energies and the related vertical IPs and EAs as follows:

$$\begin{aligned} \omega_i &= -I_i^{\mathrm{v}} \approx \varepsilon_i^{\mathrm{GSC-DFA}} = \varepsilon_i^{\mathrm{DFA}} + \Delta\varepsilon_i^{\mathrm{GSC}}, \\ \omega_a &= -A_a^{\mathrm{v}} \approx \varepsilon_a^{\mathrm{GSC-DFA}} = \varepsilon_a^{\mathrm{DFA}} + \Delta\varepsilon_a^{\mathrm{GSC}}. \end{aligned} \tag{21}$$

3. RESULTS AND DISCUSSIONS

3.1. QP Energies

3.1.1. Quasihole Energies of Molecules

Because of the lack of highly accurate experimental or theoretical data for the molecular quasielectron energies (except for those associated with the LUMOs), in this work we only compare the calculated quasihole energies that are associated with the

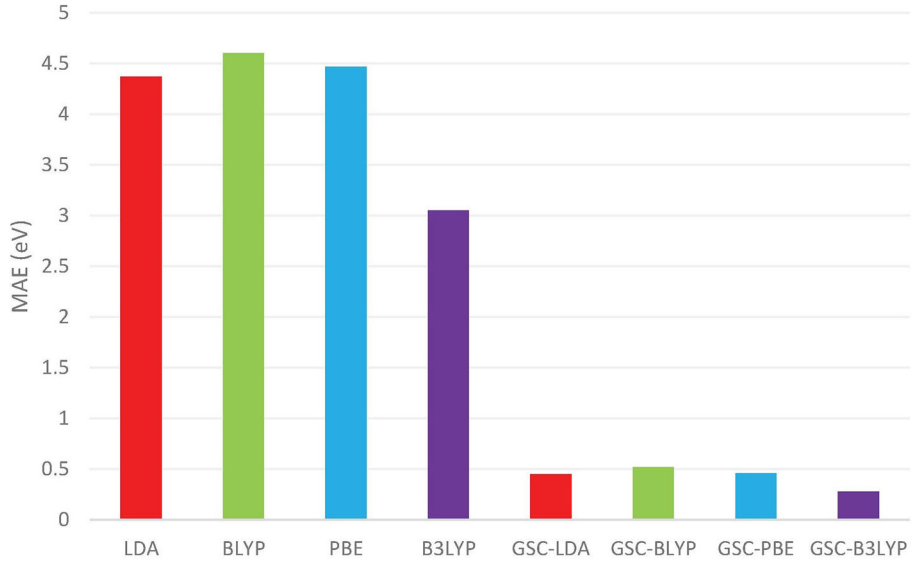

FIGURE 1 | The mean absolute errors (MAEs) (in units of eV) between the occupied Kohn–Sham (KS)/generalized KS (GKS) orbital energies $\{\varepsilon_i\}$ calculated by various density functional approximations (DFAs) and the experimentally measured quasihole energies $\{\omega_i\}$. The experimental data are extracted from Chong et al. (2002) and Schmidt (1977). The basis set adopted in the density functional calculations is aug-cc-pVTZ (Kendall et al., 1992; Woon and Dunning Jr, 1993).

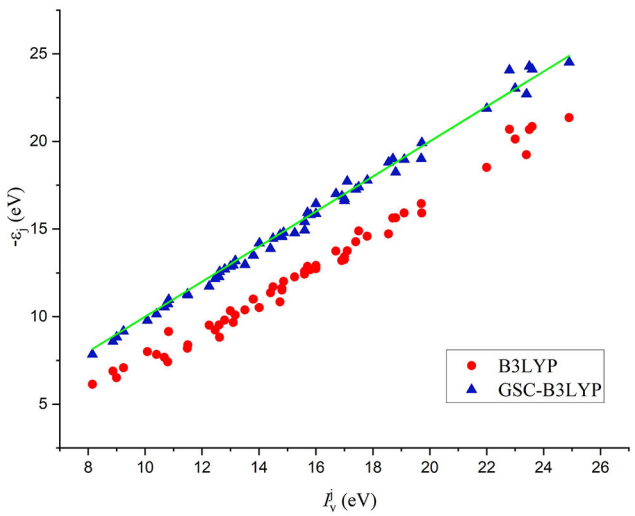

FIGURE 2 | A comparison between 56 Kohn–Sham (KS)/generalized KS (GKS) orbital energies $\{-\varepsilon_i\}$ calculated by B3LYP and GSC-B3LYP and the corresponding vertical ionization potentials (IPs) $\{I_i^V\}$ measured experimentally for 12 molecules (see the main text). The green solid line indicates exact equality.

occupied KS/GKS orbitals to the reference data available in the literature.

We first look into 56 quasihole energies of 12 molecules by calculating the scaling corrected orbital energies, and make comparison with experimentally measured vertical IPs. The examined molecules are cyanogen, CO, acetylene, water, ethylene, ammonia, acetonitrile, fluoromethane, benzene, naphthalene, furan, O2, and formic acid, which exhibit diversified geometric and electronic features. Among these molecules, the geometries of benzene and naphthalene are extracted from Mei et al. (2019), while the structures of the other molecules are optimized with the B3LYP/6-311+g** method by using the Gaussian09 package (Frisch et al., 2009).

The GSC approach presented in section 2.2 is employed to correct the orbital energies calculated by various mainstream DFAs, including the LDA, the GGAs (BLYP and PBE, Perdew et al., 1996), and the hybrid functional B3LYP. For these DFAs, the orbital relaxation up to the second order is considered for calculating the scaling corrections of the occupied orbital energies. The GSC approach is implemented in an in-house built quantum chemistry software package QM^{4}D (Hu et al., 2020).

Figure 1 compares the averaged deviations of the calculated $\{\varepsilon_i^{\mathrm{DFA}}\}$ and $\{\varepsilon_i^{\mathrm{GSC-DFA}}\}$ from the quasihole energies $\{\omega_i\}$ extracted from the experimentally measured vertical IPs. It is shown clearly that the mean absolute errors (MAEs) associated with the original DFAs are as large as several eVs, while by applying the GSC approach, the MAEs are substantially reduced to less than 0.5 eV. Take the B3LYP functional as an example. It yields an MAE of 3.05 eV, which is the smallest among all the uncorrected DFAs, and the MAE is greatly reduced to 0.28 eV by using the GSC-B3LYP. If instead the orbital relaxation is treated up to the first and third order, the MAE becomes 0.74 eV and 0.43 eV, respectively. The dependence on the order of orbital relaxation is consistent with the trend observed in our previous work (Zhang et al., 2015).

In a previous study by Chong et al. (2002), 10 out of 12 molecules examined in **Figure 1** (without benzene and naphthalene) have been investigated by calculating their KS/GKS

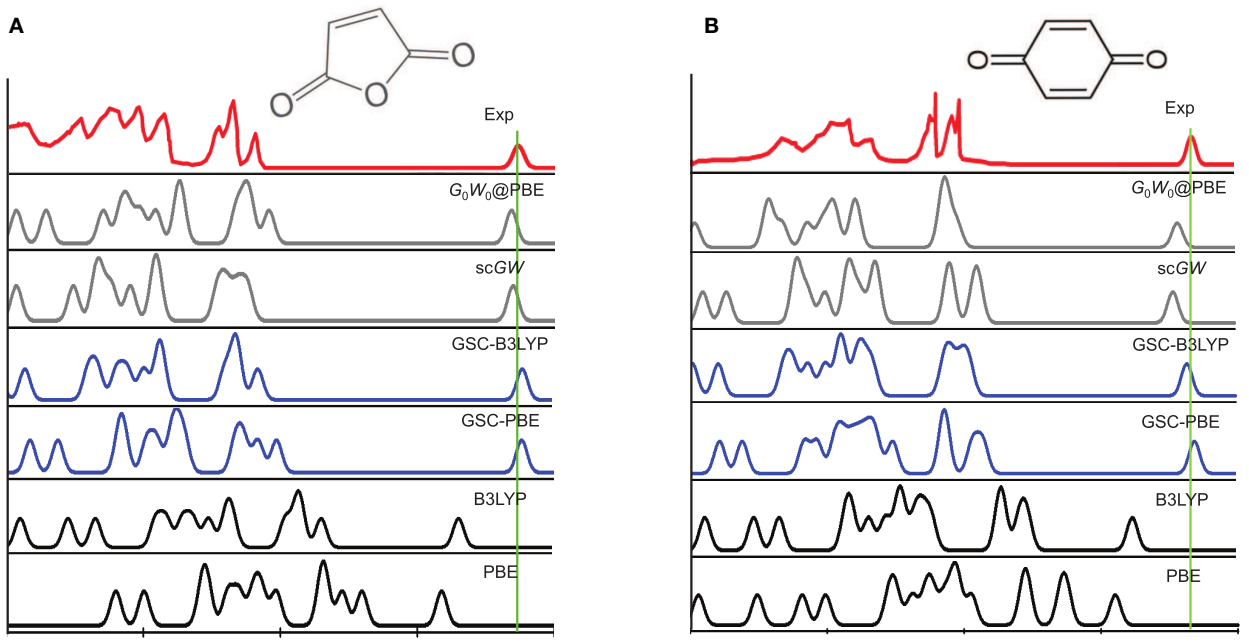

FIGURE 3 | A comparison between the experimental and simulated photoemission spectra (PES) of **(A)** a maleic anhydride and **(B)** a benzoquinone. The experimental PES are extracted from Brundle et al. (1972), Knight et al. (2016), with the rightmost quasielectron peak added manually. The quasielectron peak is centered at the experimental EA of the molecule, and is broadened artificially by a Gaussian function with the width of $\lambda = 0.2$ eV. The simulated PES by using the QE-DFT is explained in the main text, and the results of the self-consistent *GW* (sc*GW*) and non-self-consistent G_0W_0 methods are extracted from Knight et al. (2016).

orbital energies by using an approximate XC potential obtained with the statistical averaging of (model) orbital potentials (SAOP). For these 10 molecules, the MAE reported in Chong et al. (2002) is 0.38 eV, while the GSC-B3LYP yields a somewhat smaller MAE of 0.28 eV, albeit the different molecular geometries and basis sets adopted.

The comparison between the individual orbital energies $\{-\varepsilon_i\}$ calculated by B3LYP and GSC-B3LYP and the experimentally measured vertical IPs $\{I_i^{\mathrm{v}}\}$ is depicted in **Figure 2**. It is apparent that the uncorrected orbital energies deviate systematically and significantly from the experimental QP energies, while such deviations are largely alleviated by applying the GSC approach.

3.1.2. Photoemission Spectra

The QP energies can also be extracted from the peak positions of experimentally measured photoemission spectra (PES). We employ the QE-DFT to study the PES of 14 molecules. The same molecular geometries and basis set (cc-pVTZ, Dunning, 1989; Woon and Dunning Jr, 1993) as those adopted in Mei et al. (2019) are used here. The PES are simulated by setting the energy of each KS/GKS orbital as the center of a QP peak, and assuming all QP peaks have the same amplitude and are broadened by the same Gaussian function $e^{-(\varepsilon-\varepsilon_i)^2/2\lambda^2}$ with $\lambda = 0.2$ eV.

Figure 3 depicts the experimentally measured and theoretically simulated PES of a maleic anhydride and a benzoquinone, while those of the other 12 molecules are presented in **Supplementary Material**. Clearly, both the PBE and B3LYP yield considerable errors in the peak positions of the simulated PES. More specifically, they tend to predict much too high quasihole energies and too low quasielectron energies. This is because the uncorrected DFAs (PBE and B3LYP) suffer from delocalization error, as they violate the rigorous PPLB condition and the extended energy linearity relation.

The use of GSC improves significantly the simulated PES. For GSC-PBE, the orbital relaxation is considered up to the first and second order for the virtual and occupied KS/GKS orbitals, respectively; while for GSC-B3LYP, the orbital relaxation is included up to second order for all the KS/GKS orbitals. From the comparison shown in **Figure 3**, it is evident that the GSC-DFAs achieve at least the same level of accuracy as the results of *GW* method (Knight et al., 2016). Moreover, the computational cost of the QE-DFT method by using a GSC-DFA is supposedly much cheaper than that of the *GW* method, because the former requires only a single SCF calculation at the DFT level.

3.2. Energies of Low-Lying Excited States

We now turn to the energies of low-lying excited states of molecules. By employing the QE-DFT method, we carry out calculations on 48 low excitation energies of the 16 molecules investigated previously in Mei et al. (2019). The cationic species of all these molecules indeed contain one more spin-α electrons than spin-β electrons, and hence their triplet and singlet excitation energies are computed by using equations 5 and 7, respectively. Since the calculations involve only the virtual KS/GKS orbitals of the cations, the orbital relaxation is

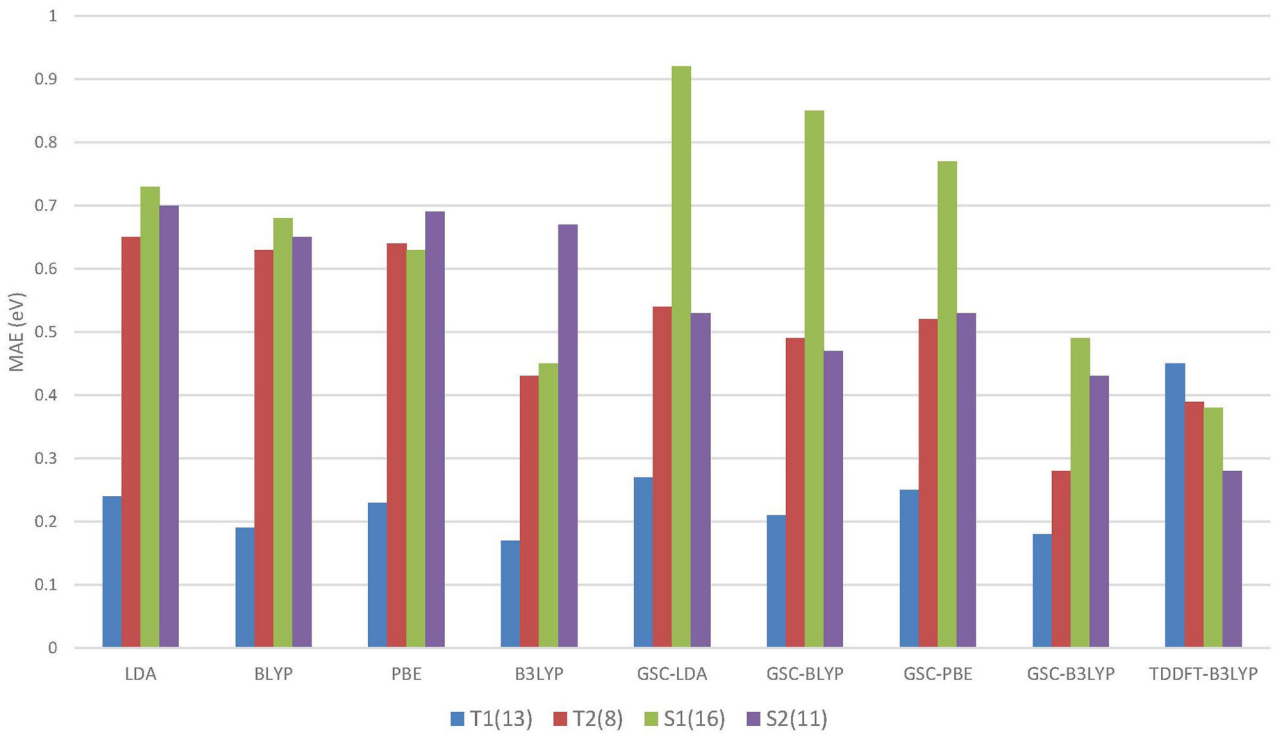

FIGURE 4 | The mean absolute errors (MAEs) of the energies of different types of excitations calculated by using the QE-DFT method with various density functional approximations (DFAs) and the basis set 6-311++G(3df, 3pd). For comparison purpose, the MAEs of the TDDFT-B3LYP results extracted from Mei et al. (2019) are also displayed. Excitation energies calculated by high-level wavefunction methods are used as the reference data (Schreiber et al., 2008). T1 and S1 (T2 and S2) refer to the triplet and singlet HOMO to LUMO (LUMO+1) excitations, respectively. The numbers in the parentheses record the numbers of energy data belonging to the different types of excitations.

considered up to the second order for GSC-B3LYP and up to the first order for other GSC-DFAs, respectively.

Figure 4 compares the MAEs of different types of excitation energies calculated by various DFAs, and the detailed results can be found in the **Supplementary Material**.

Intriguingly, for the lowest (HOMO-to-LUMO) triplet excitations, the uncorrected DFAs yield reasonably accurate excitation energies, and the application of the GSC approach does not lead to any improvement. In particular, B3LYP yields an MAE as small as 0.17 eV for the T1 excitations. Such an appealing accuracy is likely due to the cancellation of delocalization error. Equation (5) involves the difference between a pair of virtual orbital energies. Thus, when two virtual orbitals are close in energy, their associated delocalization errors are expected to cancel out (Mei et al., 2019). Consequently, the GSC approach does not help. Such an error cancellation mechanism becomes less favorable for higher excitations. For instance, as displayed in **Figure 4**, the uncorrected DFAs tend to yield a larger MAE for the T2 excitations, and applying the GSC indeed leads to improved accuracy. The latter is because the scaling correction to each individual QP energy starts to take effect.

For the S1 excitations, the GSC-DFAs yield MAEs that are somewhat larger than the original DFAs. This is because a second type of systematic error of the DFAs, the static correlation error, becomes prominent and significantly affects the calculated ΔE_{S1}. Taking the LDA functional as an example. If the first singlet excited state is described by a single Slater determinant, i.e., by setting $\chi = 0$ in Equation (6) and (7), the MAEs of ΔE_{S1} are 1.44 eV and 1.33 eV for the original LDA and GSC-LDA, respectively. In contrast, after adopting the spin purification procedure (by setting $\chi = 1$), the MAEs reduce to 0.73 eV and 0.92 eV for the original LDA and GSC-LDA, respectively. Therefore, the spin purification procedure indeed diminishes the MAE of ΔE_{S1} by circumventing the problem of static correlation error. However, the somewhat larger MAE of the GSC-LDA seems to indicate that the spin purification formula is not entirely compatible with the present GSC scheme. For higher singlet excitations, the static correlation error becomes much less significant, as signified by the much smaller second term on the right-hand side of Equation (6). Consequently, the MAE of ΔE_{S2} experiences a rather minor change by invoking the spin purification. For instance, the MAE increases slightly from 0.68 eV to 0.70 eV with the original LDA, while it reduces slightly from 0.57 eV to 0.53 eV with the GSC-LDA.

Among all the DFAs examined, the GSC-B3LYP functional achieves an optimal performance for all the low-lying excitations studied. The overall accuracy of GSC-B3LYP is comparable to the TDDFT-B3LYP. This affirms that it is entirely possible and practical to access excited-state properties of molecules within the framework of ground-state DFT.

TABLE 1 | Calculated and experimental ionization potentials (IPs) and T1 excitation energies of various transition metal atoms and compounds.

		Final state	Exp.	CCSD(T)	B3LYP	GSC-B3LYP
IP	Cu	d^{10}	7.73	7.69	5.35	7.90
	Ag	d^{10}	7.58	7.52	5.30	7.80
	Au	d^{10}	9.23	9.13	6.61	9.28
MAE[a]				0.07	2.43	0.14
T1	Cu	$d^{10}p^1$	3.81	3.87	4.92	3.94
	Ag	$d^{10}p^1$	3.84	3.74	4.80	3.87
	Au	$d^{10}p^1$	4.95	5.01	6.06	5.00
	CuF	$^3\Sigma^+$	1.81	1.81	2.73	2.05
	CuCl	$^3\Sigma^+$	2.35	2.43	3.24	2.35
	CuBr	$^3\Sigma^+$	2.54	2.56	3.19	2.35
	AgF	$1^3\Sigma^+$	3.09	3.27	3.26	2.54
	AgCl	$1^3\Sigma^+$	N/A	3.50	3.66	2.97
	AgBr	$1^3\Sigma^+$	N/A	3.37	3.43	2.83
	AuF	$1^3\Sigma^+$	N/A	2.08	2.17	1.80
	AuCl	$1^3\Sigma^+$	N/A	2.53	2.61	2.14
	AuBr	$1^3\Sigma^+$	N/A	2.60	2.52	2.09
MAE[a]				0.07	0.83	0.17
MAE[b]					0.50	0.31

The experimental excitation energies and the CCSD(T) calculation results are extracted from Guichemerre et al. (2002). The experimental geometries (Guichemerre et al., 2002) and the def2-TZVPD basis set (Schäfer et al., 1994; Rappoport and Furche, 2010) are used for all the calculations.

[a] *The MAE is between the calculated values and the experimental data.*

[b] *The MAE is between the DFT and the CCSD(T) results.*

We further extend our test to cover three transition metal atoms M (M = Cu, Ag, Au) and nine transition metal compounds MX (X = F, Cl, Br). The calculated results are presented in **Table 1** along with the CCSD(T) results and the available experimental data. The small MAEs between the calculated results and experimental data further affirm the applicability of the GSC approach.

3.3. Resonance Energies of Temporary Anions

For a temporary anion in a resonant state, the corresponding neutral molecule has a negative EA, for which the conventional ΔSCF method often yields problematic results. This is because the choice of an appropriate basis set is difficult for the SCF calculation of a temporary anion. On the one hand, the energy of a temporary anion is rather sensitive to the inclusion of diffuse basis functions (Guerra, 1990). On the other hand, the diffuse basis functions may artificially delocalize the excess electron (Cohen et al., 2008c, 2012), and thus result in incorrect electron density distribution.

Alternatively, using the scaling corrected LUMO energy to determine the energy of the temporary anion has made impressive progress. It has been demonstrated that the GSC-PBE functional predicts highly accurate negative EAs by using Equation (16) (Zhang et al., 2018). For a set of 38 molecules proposed in Tozer and De Proft (2005), the resulting MAE is as small as 0.18 eV with the aug-cc-pVTZ basis set. Recently, a similar accuracy has been reached by the explicit inclusion of derivative discontinuity in the GGA exchange potential (Carmona-Espíndola et al., 2020). In this section, we extend our calculation to 26 new molecules that are beyond the above mentioned works, and hence expand the test set to a total of 64 molecules. The molecular geometries are optimized at the B3LYP/6-311+G** level with the Gaussian09 suite of programs (Frisch et al., 2009). For the GSC approach, the relaxation of KS/GKS orbitals is considered up to second-order for GSC-B3LYP, and to first-order for other GSC-DFAs, respectively.

Table 2 lists the experimental and calculated EAs of the newly added 26 molecules. The experimental data are extracted from Jordan and Burrow (1978), Chiu et al. (1979), and Ng et al. (1983), while the theoretical data take either the values of $-\varepsilon_{\mathrm{LUMO}}$ (or $-\varepsilon_a$ if it is the ath virtual orbital that is related to the resonant state) or the energy difference between the neutral and anionic species (the ΔSCF method). More details are given in the **Supplementary Material**.

Figure 5 visualizes the MAEs of the calculated EAs of the extended set of 64 molecules. Obviously, the application of the GSC approach greatly improves the accuracy of the virtual orbital energies (particularly the $\varepsilon_{\mathrm{LUMO}}$). The MAE is reduced from several eVs with the original DFAs to less than 0.5 eV with the GSC-DFAs. Moreover, the MAE is further reduced by adopting a more diffuse basis set. This is because a more complete basis set is more favorable for a perturbative treatment of scaling correction and orbital relaxation. The lowest MAE reached for the whole extended set is 0.14 eV with the GSC-LDA.

As already been pointed out in Zhang et al. (2018), the use of a very diffuse basis set (such as aug-cc-pVTZ) may give rise to highly delocalized virtual KS/GKS orbitals with energies close to the molecular chemical potential. These orbitals are actually not relevant to the resonant state of the temporary anion of our interest, and should be left out of theoretical analysis. Therefore, we need to choose carefully the virtual orbital, which is genuinely pertinent to the formation of the temporary anion. For instance, in the case of a *cis*-butene molecule, the few lowest virtual orbitals calculated at the B3LYP/cc-pVTZ and B3LYP/aug-cc-pVTZ levels are depicted in **Figure 6**. Apparently, with the B3LYP/aug-cc-pVTZ method, the three lowest virtual orbitals (from LUMO to LUMO+2) are rather diffuse. Occupation on any of these orbitals by an excess electron will lead to an unbound state. Therefore, these orbitals are not relevant to the formation of the temporary anion. By scrutinizing the spatial distribution of the virtual KS/GKS orbitals, it is recognized that $\phi_{\mathrm{LUMO+3}}(\mathbf{r})$ would give rise to the resonant state of the temporary anion, as it exhibits a same shape as $\phi_{\mathrm{LUMO}}(\mathbf{r})$ obtained with the cc-pVTZ basis set. In such a case, instead of using Equation (16), the EA is predicted by $A = -\varepsilon^{\mathrm{GSC-B3LYP}}_{\mathrm{LUMO+3}} = -(\varepsilon^{\mathrm{B3LYP}}_{\mathrm{LUMO+3}} + \Delta\varepsilon^{\mathrm{GSC}}_{\mathrm{LUMO+3}})$. Similarly, with the B3LYP/aug-cc-pVTZ method there are some other molecules for which a virtual orbital other than the LUMO should be chosen. The virtual orbital pertinent to the temporary anion is $\phi_{\mathrm{LUMO+1}}(\mathbf{r})$ for 9 molecules (aniline, propene, CO_2, guanine, 1,4-cyclohexadiene, *cis*-1,2-difluoroethylene,

TABLE 2 | Experimental and calculated electron affinities (EAs) of 26 new molecules that are not included in Tozer and De Proft (2005) and Zhang et al. (2018).

Molecule	Exp.	LDA	BLYP	B3LYP	PBE	GSC- LDA	GSC- BLYP	GSC- B3LYP	GSC- PBE	ΔSCF- PBE	ΔSCF- B3LYP
Monofluoroethylene	−1.91	1.25	0.95	0.31	1.03	−1.97	−2.18	−2.06	−2.15	−0.47	−0.52
trans-1,2-difluoroethylene	−1.84	1.32	1.01	0.39	1.07	−1.95	−2.18	−2.09	−2.16	−0.50	−0.58
cis-1,2-difluoroethylene[a]	−2.18	1.14	0.82	0.19	0.89	−2.18	−2.42	−2.34	−2.40	−0.36	−0.40
1,1-Difluoroethylene[a]	−2.39	1.09	0.79	0.17	0.84	−2.15	−2.32	−2.09	−2.33	−0.42	−0.47
Trifluoroethylene[a]	−2.45	1.06	0.74	0.11	0.77	−2.30	−2.53	−2.38	−2.53	−0.41	−0.45
Tetrafluoroethylene[b]	−3.00	0.82	0.47	0.21	0.49	−2.71	−3.04	−3.07	−3.03	−0.89	−0.91
Nitrogen	−2.20	2.18	1.88	0.98	1.92	−2.20	−2.41	−2.40	−2.41	N/A	−1.83
Formaldehyde	−0.86	2.91	2.57	1.75	2.66	−0.92	−1.18	−1.14	−1.14	N/A	−0.46
Butadiene	−0.62	2.12	1.76	1.18	1.90	−0.61	−0.93	−0.85	−0.82	N/A	−0.73
Biphenyl	−0.30	2.04	1.63	1.13	1.80	0.01	−0.40	−0.39	−0.24	N/A	−0.37
Trichloromethane	−0.35	2.47	2.22	1.55	2.31	−0.24	−0.44	−0.46	−0.38	−0.15	−0.26
Dichlorofluoromethane	−0.96	1.99	1.74	1.08	1.81	−0.90	−1.06	−1.00	−1.04	−0.36	−0.44
Dichlorodifluoromethane	−0.98	2.37	2.11	1.43	2.17	−0.61	−0.80	−0.79	−0.77	−0.42	−0.48
Dichloromethane	−1.23	1.74	1.52	0.90	1.59	−1.02	−1.08	−0.89	−1.09	−0.31	−0.38
Benzene	−1.15	1.44	1.06	0.50	1.21	−1.18	−1.52	−1.48	−1.40	−0.36	−0.42
CO	−1.80	2.24	1.94	1.12	2.00	−1.89	−2.07	−1.96	−2.08	−1.05	−1.11
Cyanogen	−0.58	3.87	3.48	2.84	3.61	−0.48	0.12	0.23	0.23	0.21	0.29
Propyne[d]	−2.95	0.13	0.02	−0.40	0.01	−2.29	−1.83	−1.23	−2.13	−0.40	−0.47
Butadiyne	−1.00	2.09	1.73	1.16	1.87	−0.77	−1.07	−0.96	−0.97	−0.25	−0.36
Tetramethylethylene[e]	−2.27	0.42	0.20	−0.31	0.28	−1.81	−1.79	−1.48	−1.86	−0.34	−0.41
Acetylene[c]	−2.60	0.57	0.32	−0.28	0.39	−2.51	−2.58	−2.36	−2.61	−0.46	−0.53
Acrylonitrile	−0.21	3.00	2.62	2.01	2.76	−0.01	−0.35	−0.27	−0.24	0.02	−0.16
1,4-Cyclohexadiene[a]	−1.75	1.05	0.67	0.34	0.81	−1.45	−1.80	−1.89	−1.68	−0.34	−0.56
Toluene	−1.11	1.39	1.01	0.46	1.16	−1.11	−1.45	−1.43	−1.33	−0.34	−0.42
Ethylbenzene	−1.17	1.37	0.99	0.47	1.14	−1.07	−1.43	−0.90	−1.31	−0.28	−0.37
Isopropylbenzene	−1.08	1.39	0.99	0.44	1.16	−1.01	−1.38	−1.41	−1.25	−0.26	−0.34
MAE		3.17	2.85	2.23	2.95	0.17	0.25	0.30	0.21	1.22	1.01

The calculated EAs are obtained by using the ΔSCF method, or take the values of the uncorrected $-\varepsilon_{LUMO}^{DFA}$ or scaling corrected $-\varepsilon_{LUMO}^{GSC-DFA}$. All energies are in units of eV. The aug-cc-pVTZ basis set is adopted for all the calculated data listed in this table.

[a] *The $-\varepsilon_{LUMO+1}$ calculated with the GSC-B3LYP is taken as the EA of this molecule.*

[b] *The $-\varepsilon_{LUMO+1}$ calculated with the GSC-B3LYP, GSC-BLYP, and GSC-PBE are taken as the EA of this molecule.*

[c] *The $-\varepsilon_{LUMO+2}$ calculated with the GSC-B3LYP is taken as the EA of this molecule.*

[d] *The $-\varepsilon_{LUMO+2}$ calculated with the GSC-B3LYP and $-\varepsilon_{LUMO+1}$ calculated with other DFAs are taken as the EA of this molecule.*

[e] *The $-\varepsilon_{LUMO+3}$ calculated with the GSC-B3LYP and $-\varepsilon_{LUMO+1}$ calculated with other DFAs are taken as the EA of this molecule.*

1,1-difluoroethylene, trifluoroethylene and tetrafluoroethylene), $\phi_{\mathrm{LUMO+2}}(\mathbf{r})$ for 3 molecules (trimethylethylene, propyne and acetylene), $\phi_{\mathrm{LUMO+3}}(\mathbf{r})$ for 3 molecules (pyrrole, *trans*-butene, and tetramethylethylene), and $\phi_{\mathrm{LUMO+4}}(\mathbf{r})$ for one molecule (cyclohexene).

As shown in **Figure 5**, unlike the QE-DFT method, increasing the size of basis set does not improve the accuracy of the ΔSCF method. This is because through the SCF calculation of the molecular anion by using a diffuse basis set, the excess electron is more inclined to reside on the delocalized orbital, which has a lower energy. Consequently, it is difficult to have the excess electron correctly occupying the virtual orbital that is pertinent to the resonant state of temporary anion. In contrast, the QE-DFT method in conjunction with the GSC approach does not require an SCF calculation for the anionic species, and is clearly more favorable for the prediction of resonance energies of temporary anions.

4. CONCLUSION

To summarize, we have calculated the QP, excitation, and resonance energies of molecules by employing the QE-DFT method. A non-empirical GSC approach is used to reduce the delocalization error associated with the DFAs by imposing an energy linearity condition for systems with a fractional number of electrons. The accuracy of the results obtained in this work with the GSC-DFAs is overall similar to that achieved in Mei et al. (2019) by the LOSC method (Li et al., 2017). For instance, the GSC-B3LYP yields an MAE of 0.36 eV and a mean sign error (MSE) of −0.16 eV for the 48 excitation energies of 16 molecules (see section 3.2). These errors are slightly smaller than the MAE of 0.49 eV and the MSE of −0.19 eV resulted by the LOSC-B3LYP method (Mei et al., 2019). The marginal superiority in the performance of the GSC is because of the explicit treatment of the relaxation of KS/GKS orbitals upon electron addition or removal. Relaxation of KS/GKS orbitals and electron density

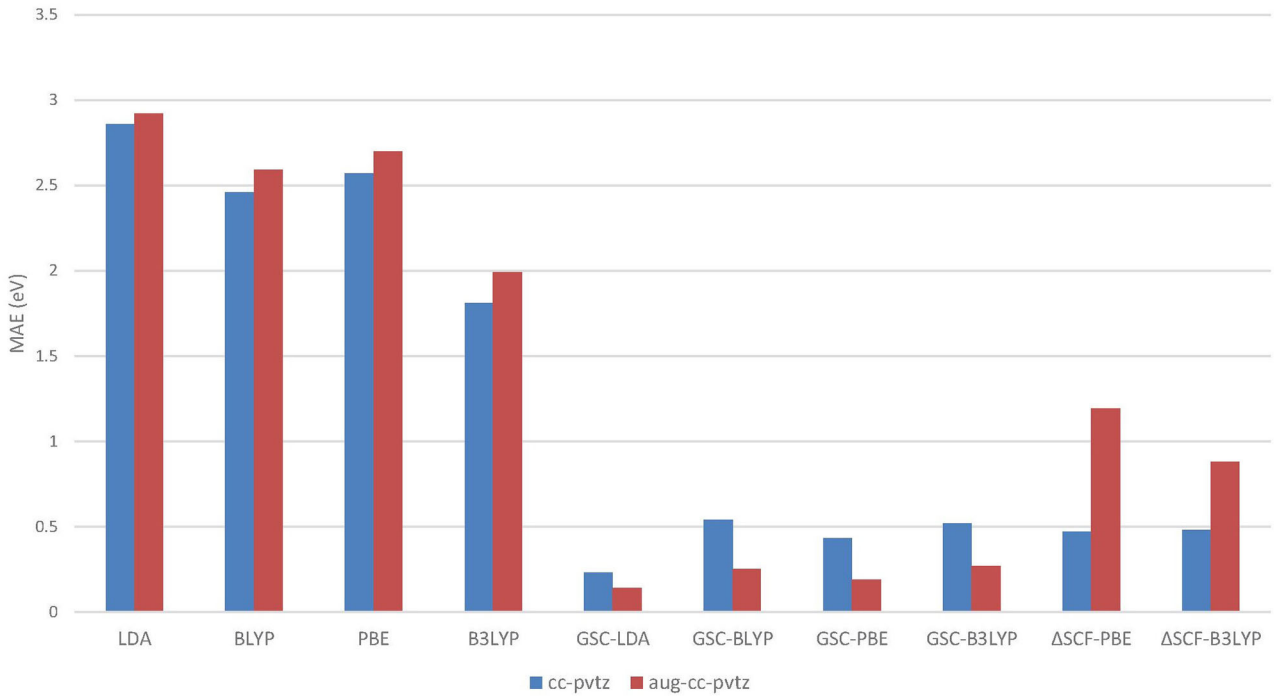

FIGURE 5 | The mean absolute errors (MAEs) of the EAs for an extended set of 64 molecules calculated by employing the QE-DFT method with various density functional approximations (DFAs) and by using the ΔSCF method. Note that when the more diffuse aug-cc-pVTZ basis set is adopted, the energy of a certain virtual Kohn–Sham (KS)/generalized KS (GKS) orbital should be taken as the predicted EA; see **Table 2** for details. If the orbital relaxation is considered up to the first order for the GSC-B3LYP, the MAEs become 0.21 and 0.28 eV with the cc-pVTZ and aug-cc-pVTZ basis sets, respectively.

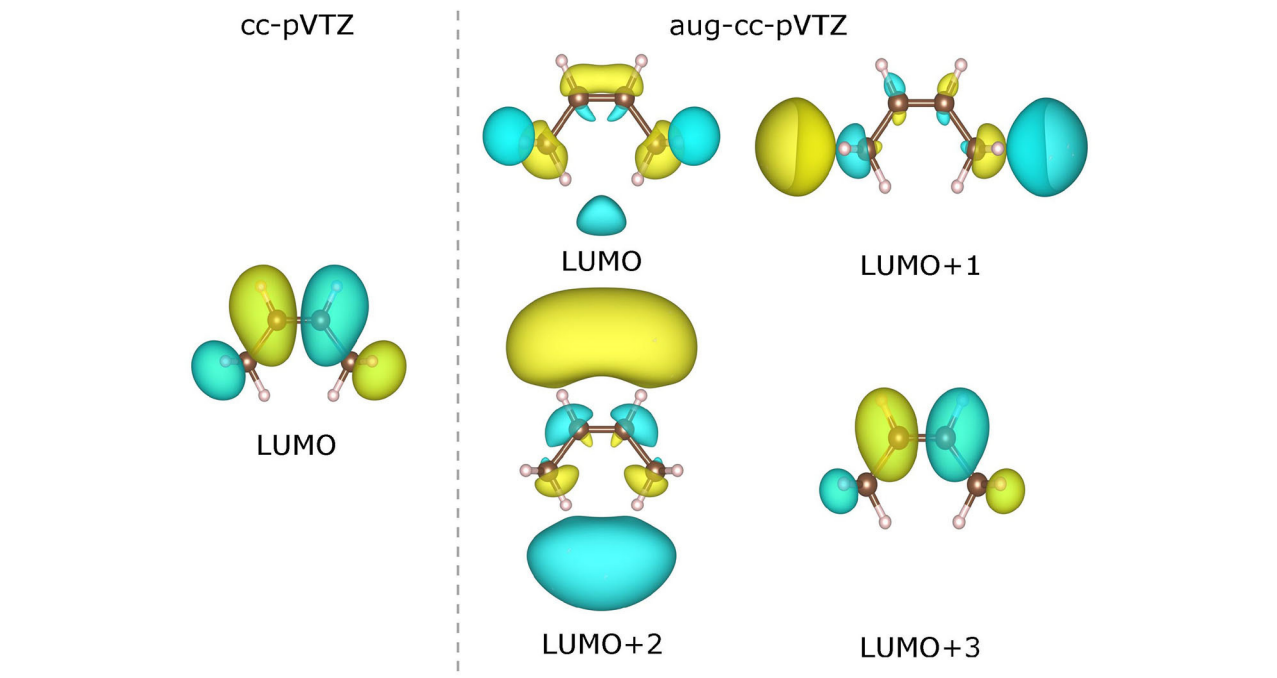

FIGURE 6 | Contour plots of the lowest virtual Kohn–Sham (KS)/generalized KS (GKS) orbitals of the neutral *cis*-butene molecule calculated at the B3LYP/cc-pVTZ and B3LYP/aug-cc-pVTZ levels. The isosurfaces of ±0.022 a.u. are shaded in yellow and green, respectively).

could also be included in the LOSC calculations (Su et al., 2020), which will further improve the accuracy of predicted QP energies. Moreover, as the size of the molecule increases to a certain extent, the energies at integer electron numbers may become less accurate. For such systems, the correction offered by the GSC approach may be inadequate, and the LOSC method with size-consistent corrections (Su et al., 2020; Yang et al., 2020) to DFA should be used.

For the various DFAs considered in this paper, the GSC-B3LYP yields the overall best performance. Our calculation results achieve at least the same level of accuracy as some more expensive methods, such as the *GW* method for QP energies, the TDDFT method for excitation energies, and the EOM-CC method for resonance energies. This thus affirms that it is entirely possible and practical to study excited-state properties within the framework of ground-state DFT. Despite the promising results, the prediction of singlet excitation energies still has plenty of room for improvement. This is because another source of error associated with the DFAs, the static correlation error, comes into play, which may be corrected by imposing a constancy condition on systems with fractional spins (Cohen et al., 2008b). Further work is needed along this direction.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

XZ and WY conceived the project. XY conducted the numerical calculations. XY, XZ, and WY analyzed the data and WY wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

XY and XZ acknowledge the support from the Ministry of Science and Technology of China (Grant Nos. 2016YFA0400900 and 2016YFA0200600), the National Natural Science Foundation of China (Grant No. 21973086), and the Ministry of Education of China (111 Project Grant No. B18051). WY has been supported by the National Science Foundation (CHE-1900338) and the National Institute of General Medical Sciences of National Institute of Health (R01-GM061870).

ACKNOWLEDGMENTS

The support from the Supercomputing Center of University of Science and Technology of China is appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.588808/full#supplementary-material

REFERENCES

Andersson, K., Malmqvist, P. A., Roos, B. O., Sadlej, A. J., and Wolinski, K. (1990). Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem*. 94, 5483–5488. doi: 10.1021/j100377a012

Aulbur, W. G., Jönsson, L., and Wilkins, J. W. (2000). Quasiparticle calculations in solids. *Solid State Phys*. 54, 1–218. doi: 10.1016/S0081-1947(08)60248-9

Baerends, E., Gritsenko, O., and Van Meer, R. (2013). The Kohn-Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn-Sham orbital energies. *Phys. Chem. Chem. Phys*. 15, 16408–16425. doi: 10.1039/c3cp52547c

Bartlett, R. J. (2009). Towards an exact correlated orbital theory for electrons. *Chem. Phys. Lett*. 484, 1–9. doi: 10.1016/j.cplett.2009.10.053

Bartlett, R. J., and Ranasinghe, D. S. (2017). The power of exact conditions in electronic structure theory. *Chem. Phys. Lett*. 669, 54–70. doi: 10.1016/j.cplett.2016.12.017

Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38:3098. doi: 10.1103/PhysRevA.38.3098

Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys*. 98, 5648–5652. doi: 10.1063/1.464913

Brundle, C., Robin, M., and Kuebler, N. (1972). Perfluoro effect in photoelectron spectroscopy. II. Aromatic molecules. *J. Am. Chem. Soc*. 94, 1466–1475. doi: 10.1021/ja00760a008

Carmona-Espíndola, J., Gázquez, J. L., Vela, A., and Trickey, S. B. (2020). Negative electron affinities and derivative discontinuity contribution from a generalized gradient approximation exchange functional. *J. Phys. Chem. A* 124, 1334–1342. doi: 10.1021/acs.jpca.9b10956

Casida, M. E. (1995). "Time-dependent density functional response theory for molecules," in *Recent Advances in Density Functional Methods: (Part I)*, ed D. P. Chong (Singapore: World Scientific), 155–192.

Chiu, N. S., Burrow, P. D., and Jordan, K. D. (1979). Temporary anions of the fluoroethylenes. *Chem. Phys. Lett*. 68, 121–126. doi: 10.1016/0009-2614(79)80082-2

Chong, D. P., Gritsenko, O. V., and Baerends, E. J. (2002). Interpretation of the Kohn-Sham orbital energies as approximate vertical ionization potentials. *J. Chem. Phys*. 116, 1760–1772. doi: 10.1063/1.1430255

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2008a). Fractional charge perspective on the band gap in density-functional theory. *Phys. Rev. B* 77:115123. doi: 10.1103/PhysRevB.77.115123

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2008b). Fractional spins and static correlation error in density functional theory. *J. Chem. Phys*. 129:121104. doi: 10.1063/1.2987202

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2008c). Insights into current limitations of density functional theory. *Science* 321, 792–794. doi: 10.1126/science.1158722

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2012). Challenges for density functional theory. *Chem. Rev*. 112, 289–320. doi: 10.1021/cr200107z

Coropceanu, V., Malagoli, M., da Silva Filho, D., Gruhn, N., Bill, T., and Brédas, J. (2002). Hole-and electron-vibrational couplings in oligoacene crystals: intramolecular contributions. *Phys. Rev. Lett*. 89:275503. doi: 10.1103/PhysRevLett.89.275503

Dauth, M., Körzdörfer, T., Kümmel, S., Ziroff, J., Wiessner, M., Schöll, A., et al. (2011). Orbital density reconstruction for molecules. *Phys. Rev. Lett*. 107:193002. doi: 10.1103/PhysRevLett.107.193002

Dolgounitcheva, O., Díaz-Tinoco, M., Zakrzewski, V. G., Richard, R. M., Marom, N., Sherrill, C. D., et al. (2016). Accurate ionization potentials and electron affinities of acceptor molecules IV: electron-propagator methods. *J. Chem. Theory Comput*. 12, 627–637. doi: 10.1021/acs.jctc.5b00872

Dreuw, A., and Wormit, M. (2015). The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscipl. Rev. Comput. Mol. Sci*. 5, 82–95. doi: 10.1002/wcms.1206

Dunning, T. H. Jr. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* 90, 1007–1023. doi: 10.1063/1.456153

Dutta, A. K., Gupta, J., Pathak, H., Vaval, N., and Pal, S. (2014). Partitioned EOMEA-MBPT (2): an efficient N^5 scaling method for calculation of electron affinities. *J. Chem. Theory Comput.* 10, 1923–1933. doi: 10.1021/ct4009409

Dvorak, M., Chen, X.-J., and Wu, Z. (2014). Quasiparticle energies and excitonic effects in dense solid hydrogen near metallization. *Phys. Rev. B* 90:035103. doi: 10.1103/PhysRevB.90.035103

Ess, D. H., Johnson, E. R., Hu, X., and Yang, W. (2011). Singlet-triplet energy gaps for diradicals from fractional-spin density-functional theory. *J. Phys. Chem. A* 115, 76–83. doi: 10.1021/jp109280y

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2009). *Gaussian 09, RevisionA.01*. Wallingford, CT: Gaussian, Inc.

Govoni, M., and Galli, G. (2018). GW100: Comparison of methods and accuracy of results obtained with the WEST code. *J. Chem. Theory Comput.* 14, 1895–1909. doi: 10.1021/acs.jctc.7b00952

Gritsenko, O., and Baerends, E. J. (2009). The analog of Koopmans' theorem for virtual Kohn-Sham orbital energies. *Can. J. Chem.* 87, 1383–1391. doi: 10.1139/V09-088

Guerra, M. (1990). On the use of diffuse functions for estimating negative electron affinities with LCAO methods. *Chem. Phys. Lett.* 167, 315–319. doi: 10.1016/0009-2614(90)87174-P

Guichemerre, M., Chambaud, G., and Stoll, H. (2002). Electronic structure and spectroscopy of monohalides of metals of group I-B. *Chem. Phys.* 280, 71–102. doi: 10.1016/S0301-0104(02)00510-4

Haiduke, R. L. A., and Bartlett, R. J. (2018). Communication: can excitation energies be obtained from orbital energies in a correlated orbital theory? *J. Chem. Phys.* 149:131101. doi: 10.1063/1.5052442

Hedin, L. (1965). New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* 139:A796. doi: 10.1103/PhysRev.139.A796

Hill, I., Kahn, A., Cornil, J., Dos Santos, D., and Brédas, J. (2000). Occupied and unoccupied electronic levels in organic π-conjugated molecules: comparison between experiment and theory. *Chem. Phys. Lett.* 317, 444–450. doi: 10.1016/S0009-2614(99)01384-6

Hirao, K., Chan, B., Song, J.-W., Bhattarai, K., and Tewary, S. (2020). Excitation energies expressed as orbital energies of Kohn-Sham density functional theory with long-range corrected functionals. *J. Comput. Chem.* 41, 1368–1383. doi: 10.1002/jcc.26181

Hohenberg, P., and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.* 136:B864. doi: 10.1103/PhysRev.136.B864

Hoyer, C. E., Ghosh, S., Truhlar, D. G., and Gagliardi, L. (2016). Multiconfiguration pair-density functional theory is as accurate as CASPT2 for electronic excitation. *J. Phys. Chem. Lett.* 7, 586–591. doi: 10.1021/acs.jpclett.5b02773

Hu, X., Hu, H., Zheng, X., Zeng, X., Wu, P., Peng, D., et al. (2020). *An In-house Program for QM/MM Simulations*. Available online at: http://www.qm4d.info

Hybertsen, M. S., and Louie, S. G. (1986). Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. *Phys. Rev. B* 34:5390. doi: 10.1103/PhysRevB.34.5390

Ipatov, A., Cordova, F., Doriol, L. J., and Casida, M. E. (2009). Excited-state spin-contamination in time-dependent density-functional theory for molecules with open-shell ground states. *J. Mol. Struct. Theochem.* 914, 60–73. doi: 10.1016/j.theochem.2009.07.036

Jacquemin, D., Duchemin, I., and Blase, X. (2015). 0-0 energies using hybrid schemes: Benchmarks of TD-DFT, CIS(D), ADC(2), CC2, and BSE/GW formalisms for 80 real-life compounds. *J. Chem. Theory Comput.* 11, 5340–5359. doi: 10.1021/acs.jctc.5b00619

Jacquemin, D., Duchemin, I., and Blase, X. (2017). Is the Bethe-Salpeter formalism accurate for excitation energies? Comparisons with TD-DFT, CASPT2, and EOM-CCSD. *J. Phys. Chem. Lett.* 8, 1524–1529. doi: 10.1021/acs.jpclett.7b00381

Jagau, T., Bravaya, K. B., and Krylov, A. I. (2017). Extending quantum chemistry of bound states to electronic resonances. *Annu. Rev. Phys. Chem.* 68, 525–553. doi: 10.1146/annurev-physchem-052516-050622

Janak, J. (1978). Proof that $\frac{\partial E}{\partial n_i} = \varepsilon_i$ in density-functional theory. *Phys. Rev. B* 18:7165. doi: 10.1103/PhysRevB.18.7165

Jordan, K. D., and Burrow, P. D. (1978). Studies of the temporary anion states of unsaturated hydrocarbons by electron transmission spectroscopy. *ChemInform* 10, 341–348. doi: 10.1021/ar50129a004

Jordan, K. D., and Burrow, P. D. (1987). Temporary anion states of polyatomic hydrocarbons. *Chem. Rev.* 87, 557–588. doi: 10.1021/cr00079a005

Jordan, K. D., Voora, V. K., and Simons, J. (2014). Negative electron affinities from conventional electronic structure methods. *Theor. Chem. Acc.* 133:1445. doi: 10.1007/s00214-014-1445-1

Kendall, R. A., Dunning, T. H. Jr., and Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* 96, 6796–6806. doi: 10.1063/1.462569

Knight, J. W., Wang, X., Gallandi, L., Dolgounitcheva, O., Ren, X., Ortiz, J. V., et al. (2016). Accurate ionization potentials and electron affinities of acceptor molecules III: a benchmark of GW methods. *J. Chem. Theory Comput.* 12, 615–626. doi: 10.1021/acs.jctc.5b00871

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140:A1133. doi: 10.1103/PhysRev.140.A1133

Körzdörfer, T., Parrish, R. M., Marom, N., Sears, J. S., Sherrill, C. D., and Brédas, J.-L. (2012). Assessment of the performance of tuned range-separated hybrid density functionals in predicting accurate quasiparticle spectra. *Phys. Rev. B* 86:205110. doi: 10.1103/PhysRevB.86.205110

Laurent, A. D., and Jacquemin, D. (2013). TD-DFT benchmarks: a review. *Int. J. Quant. Chem.* 113, 2019–2039. doi: 10.1002/qua.24438

Lee, C., Yang, W., and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37:785. doi: 10.1103/PhysRevB.37.785

Li, C., Zheng, X., Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2015). Local scaling correction for reducing delocalization error in density functional approximations. *Phys. Rev. Lett.* 114:053001. doi: 10.1103/PhysRevLett.114.053001

Li, C., Zheng, X., Su, N. Q., and Yang, W. (2017). Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations. *Nat. Sci. Rev.* 5, 203–215. doi: 10.1093/nsr/nwx111

Longo, R., Champagne, B., and Öhrn, Y. (1995). Electron propagator theory and application. *Theoret. Chim. Acta* 90, 397–419. doi: 10.1007/BF01113544

Louie, S. G., and Hybertsen, M. S. (1987). Theory of quasiparticle energies: band gaps and excitation spectra in solids. *Int. J. Quant. Chem.* 32, 31–44. doi: 10.1002/qua.560320706

Ma, F., Wang, Z., Guo, M., and Wang, F. (2020). Approximate equation-of-motion coupled-cluster methods for electron affinities of closed-shell molecules. *J. Chem. Phys.* 152:124111. doi: 10.1063/1.5142736

Mei, Y., Li, C., Su, N. Q., and Yang, W. (2019). Approximating quasiparticle and excitation energies from ground state generalized Kohn-Sham calculations. *J. Phys. Chem. A* 123, 666–673. doi: 10.1021/acs.jpca.8b10380

Mei, Y., and Yang, W. (2019). Excited-state potential energy surfaces, conical intersections, and analytical gradients from ground-state density functional theory. *J. Phys. Chem. Lett.* 10, 2538–2545. doi: 10.1021/acs.jpclett.9b00712

Ng, L., Balaji, V., and Jordan, K. D. (1983). Measurement of the vertical electron affinities of cyanogen and 2,4-hexadiyne. *Chem. Phys. Lett.* 101, 171–176. doi: 10.1016/0009-2614(83)87365-5

Nooijen, M., and Bartlett, R. J. (1995). Equation of motion coupled cluster method for electron attachment. *J. Chem. Phys.* 102, 3629–3647. doi: 10.1063/1.468592

Onida, G., Reining, L., and Rubio, A. (2002). Electronic excitations: density-functional versus many-body Green-function approaches. *Rev. Modern Phys.* 74:601. doi: 10.1103/RevModPhys.74.601

Ortiz, J. V. (2013). Electron propagator theory: an approach to prediction and interpretation in quantum chemistry. *Wiley Interdiscipl. Rev. Comput. Mol. Sci.* 3, 123–142. doi: 10.1002/wcms.1116

Parr, R. G., and Yang, W. (1984). Density functional approach to the frontier-electron theory of chemical reactivity. *J. Am. Chem. Soc.* 106, 4049–4050. doi: 10.1021/ja00326a036

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77:3865. doi: 10.1103/PhysRevLett.77.3865

Perdew, J. P., Parr, R. G., Levy, M., and Balduz, J. L. Jr. (1982). Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.* 49:1691. doi: 10.1103/PhysRevLett.49.1691

Perdew, J. P., Ruzsinszky, A., Csonka, G. I., Vydrov, O. A., Scuseria, G. E., Staroverov, V. N., et al. (2007). Exchange and correlation in open systems of fluctuating electron number. *Phys. Rev. A* 76:040501. doi: 10.1103/PhysRevA.76.040501

Potts, D. M., Taylor, C. M., Chaudhuri, R. K., and Freed, K. F. (2001). The improved virtual orbital-complete active space configuration interaction method, a "packageable" efficient *ab initio* many-body method for describing electronically excited states. *J. Chem. Phys*. 114, 2592–2600. doi: 10.1063/1.1337053

Puschnig, P., Boese, A., Willenbockel, M., Meyer, M., Lüftner, D., Reinisch, E., et al. (2017). Energy ordering of molecular orbitals. *J. Phys. Chem. Lett*. 8, 208–213. doi: 10.1021/acs.jpclett.6b02517

Ranasinghe, D. S., Margraf, J. T., Jin, Y., and Bartlett, R. J. (2017). Does the ionization potential condition employed in QTP functionals mitigate the self-interaction error? *J. Chem. Phys*. 146, 034102–034102. doi: 10.1063/1.4973727

Rappoport, D., and Furche, F. (2010). Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys*. 133:134105. doi: 10.1063/1.3484283

Rohlfing, M., and Louie, S. G. (2000). Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B* 62:4927. doi: 10.1103/PhysRevB.62.4927

Runge, E., and Gross, E. K. (1984). Density-functional theory for time-dependent systems. *Phys. Rev. Lett*. 52:997. doi: 10.1103/PhysRevLett.52.997

Sanche, L., and Schulz, G. (1972). Electron transmission spectroscopy: rare gases. *Phys. Rev. A* 5:1672. doi: 10.1103/PhysRevA.5.1672

Santoro, F., and Jacquemin, D. (2016). Going beyond the vertical approximation with time-dependent density functional theory. *Wiley Interdiscipl. Rev. Comput. Mol. Sci*. 6, 460–486. doi: 10.1002/wcms.1260

Schäfer, A., Huber, C., and Ahlrichs, R. (1994). Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys*. 100, 5829–5835. doi: 10.1063/1.467146

Schmidt, W. (1977). Photoelectron spectra of polynuclear aromatics. V. Correlations with ultraviolet absorption spectra in the catacondensed series. *J. Chem. Phys*. 66, 828–845. doi: 10.1063/1.433961

Schreiber, M., Silva-Junior, M. R., Sauer, S. P., and Thiel, W. (2008). Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys*. 128:134110. doi: 10.1063/1.2889385

Schulz, G. J. (1973). Resonances in electron impact on atoms. *Rev. Modern Phys*. 45:378. doi: 10.1103/RevModPhys.45.378

Silva-Junior, M. R., Schreiber, M., Sauer, S. P., and Thiel, W. (2008). Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. *J. Chem. Phys*.129:104103. doi: 10.1063/1.2973541

Skomorowski, W., Gulania, S., and Krylov, A. I. (2018). Bound and continuum-embedded states of cyanopolyyne anions. *Phys. Chem. Chem. Phys*. 20, 4805–4817. doi: 10.1039/C7CP08227D

Slater, J. C. (1951). A simplification of the Hartree-Fock method. *Phys. Rev*. 81:385. doi: 10.1103/PhysRev.81.385

Slavicek, P., and Martinez, T. J. (2010). *Ab initio* floating occupation molecular orbital-complete active space configuration interaction: an efficient approximation to CASSCF. *J. Chem. Phys*. 132, 234102–234102. doi: 10.1063/1.3436501

Stanton, J. F., and Bartlett, R. J. (1993). The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys*. 98, 7029–7039. doi: 10.1063/1.464746

Su, N. Q., Mahler, A., and Yang, W. (2020). Preserving symmetry and degeneracy in the localized orbital scaling correction approach. *J. Chem. Phys. Lett*. 11, 1528–1535. doi: 10.1021/acs.jpclett.9b03888

Thierbach, A., Neiss, C., Gallandi, L., Marom, N., Körzdörfer, T., and Görling, A. (2017). Accurate valence ionization energies from Kohn-Sham eigenvalues with the help of potential adjustors. *J. Chem. Theory Comput*. 13, 4726–4740. doi: 10.1021/acs.jctc.7b00490

Tozer, D. J., and De Proft, F. (2005). Computation of the hardness and the problem of negative electron affinities in density functional theory. *J. Phys. Chem. A* 109, 8923–8929. doi: 10.1021/jp053504y

Tsuneda, T., Song, J.-W., Suzuki, S., and Hirao, K. (2010). On Koopmans' theorem in density functional theory. *J. Chem. Phys*. 133:174101. doi: 10.1063/1.3491272

Vargas, R., Garza, J., and Cedillo, A. (2005). Koopmans-like approximation in the Kohn-Sham method and the impact of the frozen core approximation on the computation of the reactivity parameters of the density functional theory. *J. Phys. Chem. A* 109, 8880–8892. doi: 10.1021/jp052111w

Vosko, S. H., Wilk, L., and Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *UC Irvine* 58, 1200–1211. doi: 10.1139/p80-159

Wang, Z., Tu, Z., and Wang, F. (2014). Equation-of-motion coupled-cluster theory for excitation energies of closed-shell systems with spin-orbit coupling. *J. Chem. Theory Comput*. 10, 5567–5576. doi: 10.1021/ct500854m

Winter, N. O., Graf, N. K., Leutwyler, S., and Hättig, C. (2013). Benchmarks for 0-0 transitions of aromatic organic molecules: DFT/B3LYP, ADC (2), CC2, SOS-CC2 and SCS-CC2 compared to high-resolution gas-phase data. *Phys. Chem. Chem. Phys*. 15, 6623–6630. doi: 10.1039/C2CP42694C

Woon, D. E., and Dunning Jr, T. H. (1993). Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys*. 98, 1358–1371. doi: 10.1063/1.464303

Yang, W., Cohen, A. J., and Mori-Sánchez, P. (2012). Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional theory. *J. Chem. Phys*. 136:204111. doi: 10.1063/1.3702391

Yang, W., and Parr, R. G. (1985). Hardness, softness, and the Fukui function in the electronic theory of metals and catalysis. *Proc. Natl. Acad. Sci. U.S.A*. 82, 6723–6726. doi: 10.1073/pnas.82.20.6723

Yang, W., Parr, R. G., and Pucci, R. (1984). Electron density, Kohn-Sham frontier orbitals, and Fukui functions. *J. Chem. Phys*. 81, 2862–2863. doi: 10.1063/1.447964

Yang, W., Zhang, Y., and Ayers, P. W. (2000). Degenerate ground states and a fractional number of electrons in density and reduced density matrix functional theory. *Phys. Rev. Lett*. 84:5172. doi: 10.1103/PhysRevLett.84.5172

Yang, X., He, Z., and Zheng, X. (2020). Unit cell consistency of maximally localized Wannier functions. *Electron. Struct*. 2:014001. doi: 10.1088/2516-1075/ab5e5a

Zhang, D., Yang, X., Zheng, X., and Yang, W. (2018). Accurate density functional prediction of molecular electron affinity with the scaling corrected Kohn-Sham frontier orbital energies. *Mol. Phys*. 116, 927–934. doi: 10.1080/00268976.2017.1382738

Zhang, D., Zheng, X., Li, C., and Yang, W. (2015). Orbital relaxation effects on Kohn-Sham frontier orbital energies in density functional theory. *J. Chem. Phys*. 142:154113. doi: 10.1063/1.4918347

Zheng, X., Cohen, A. J., Mori-Sánchez, P., Hu, X., and Yang, W. (2011). Improving band gap prediction in density functional theory from molecules to solids. *Phys. Rev. Lett*. 107:026403. doi: 10.1103/PhysRevLett.107.026403

Zheng, X., Li, C., Zhang, D., and Yang, W. (2015). Scaling correction approaches for reducing delocalization error in density functional approximations. *Sci. China Chem*. 58, 1825–1844. doi: 10.1007/s11426-015-5501-z

Zheng, X., Zhou, T., and Yang, W. (2013). A nonempirical scaling correction approach for density functional methods involving substantial amount of Hartree-Fock exchange. *J. Chem. Phys*. 138:174105. doi: 10.1063/1.4801922

Ziegler, T., Rauk, A., and Baerends, E. J. (1977). On the calculation of multiplet energies by the Hartree-Fock-Slater method. *Theor. Chem. Acc*. 43, 261–271. doi: 10.1007/BF00551551

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 10 December 2020
doi: 10.3389/fchem.2020.590184

Solving Coupled Cluster Equations by the Newton Krylov Method

Chao Yang[1]*, *Jiri Brabec*[2], *Libor Veis*[2], *David B. Williams-Young*[1] and *Karol Kowalski*[3]

[1] Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [2] J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences, Prague, Czechia, [3] Pacific Northwest National Laboratory, Richland, WA, United States

We describe using the Newton Krylov method to solve the coupled cluster equation. The method uses a Krylov iterative method to compute the Newton correction to the approximate coupled cluster amplitude. The multiplication of the Jacobian with a vector, which is required in each step of a Krylov iterative method such as the Generalized Minimum Residual (GMRES) method, is carried out through a finite difference approximation, and requires an additional residual evaluation. The overall cost of the method is determined by the sum of the inner Krylov and outer Newton iterations. We discuss the termination criterion used for the inner iteration and show how to apply pre-conditioners to accelerate convergence. We will also examine the use of regularization technique to improve the stability of convergence and compare the method with the widely used direct inversion of iterative subspace (DIIS) methods through numerical examples.

Keywords: couple cluster approximation, Newton-Krylov method, DIIS, precondition, nonlinear solver

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Igor Ying Zhang,
Fudan University, China
Jie Liu,
University of Science and Technology of China, China

***Correspondence:**
Chao Yang
cyang@lbl.gov

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *31 July 2020*
Accepted: *28 September 2020*
Published: *10 December 2020*

Citation:
Yang C, Brabec J, Veis L, Williams-Young DB and Kowalski K (2020) Solving Coupled Cluster Equations by the Newton Krylov Method. Front. Chem. 8:590184. doi: 10.3389/fchem.2020.590184

1. INTRODUCTION

The coupled cluster (CC) theory, introduced to quantum chemistry by Čížek (1966), Paldus and Li (1999), and Bartlett and Musiał (2007), over the past few decades has established itself as one of the most accurate *ab initio* method for electronic structure calculations. The systematic inclusion of higher-rank excitations in the cluster operator allows one to establish a hierarchy of more and more accurate approximations converging toward the full configuration interaction (FCI) limit (Gauss, 1998). These standard approximations also provide a number of unique features such as size-extensivity of the resulting energies, orbital invariance of theory under separate rotations of occupied and virtual orbitals, the possibility of approximating higher excitations by products of lower-rank clusters, which are especially important in proper description of chemical transformations associated with bond forming and bond breaking processes. The CC theory is based on an exponential ansatz acting on the reference wave function, typically Slater determinant obtained from Hartree–Fock (HF), density functional theory, or other independent particle models, which is assumed to provide a reasonable zeroth order description of the correlated ground-state wave function. The CC wave function is determined by the so-called cluster amplitudes obtained by solving nonlinear energy-independent CC equations.

The exponential ansatz ensures the size-extensivity of the CC method, but in contrast to configuration interaction methods, the CC method is not variational (unless all excitations are included). In practice, to make CC approximation numerically feasible, the cluster operator is defined by low-rank excitations. For example, one of the most widely used coupled cluster single and double (CCSD) model includes single and double excitations (Purvis and Bartlett, 1982).

The single-reference CCSD method and non-iterative technique for the inclusion of collective triple excitations [the so-called CCSD(T) approach; Raghavachari et al., 1989] is currently considered as a "gold standard" of high-accuracy computational chemistry. It is available in many program packages and employed widely in a wide variety of chemical applications. The numerical cost scales polynomially with the system size, where the numerical scaling of CCSD is proportional to $\mathcal{O}(N^6)$, whereas for (T) correction is proportional to $\mathcal{O}(N^7)$ (N represents symbolically system size).

For the treatment of the static correlation effects, the multireference CC approach has been introduced, which generalizes the CC exponential parameterization of the wave function (Lyakh et al., 2012). Out of many formulations of MRCC theories, the class of methods relevant to this work is externally corrected CC, which extracts information about the most important higher excitations or active space single and double excitations from an "external" calculation performed by a different method such as complete active space self-consistent field (CASSCF) or multireference configuration interaction (MRCI) (Li and Paldus, 1997; Li, 2001; Kinoshita et al., 2005). In this work, we employed the tailored CCSD (TCCSD) method, where the information for external correction is obtained from a density matrix renormalization group (DMRG) calculation. The TCC approach has been successfully applied (Kinoshita et al., 2005; Lyakh et al., 2011) and generally performs well, although a large active space and CASSCF orbitals might be required for good accuracy. TCC also features the desirable property of being rigorously size extensive.

The CCSD equations correspond to the polynomial set of equations of fourth order, whose solving for large number of cluster amplitudes (often exceeding 10^{10}) in the presence of strong correlation effects may pose a significant challenge and may adversely affect the time to solution associated with solving CC equations, even when efficient implementation of the CCSD method is available. Therefore, the design of fast converging CC solvers is inextricably linked to the effort of enabling CC methods at exa-scale. Currently, CC equations are typically solved via an inexact Newton (IN) method combined with an acceleration scheme called the direct inversion of iterative space (DIIS) (Pulay, 1980), which is also used in many other quantum chemical algorithms to accelerate the convergence, for example in the self-consistent field (SCF) iterations for solving the HF equations. Several other algorithms such as reduced linear equation (Purvis and Bartlett, 1981), quasilinearization of nonlinear terms techniques (Piecuch and Adamowicz, 1994), and multimodel Newton-type algorithms (Kjønstad et al., 2020) have been tested especially in the context of solving CC equations involving high-rank clusters.

In this paper, we describe using the Newton–Krylov (NK) method for solving the projected CC equation. The NK method is a widely used method for solving large-scale nonlinear equations in many fields (Knoll and Keyes, 2004). Its use in quantum chemistry appears to be new. We will describe the basic steps of the method in the context of CCSD in section 3. We compare the method with DIIS in section 4 and discuss the possibility of combining the two methods together. In section 5, we demonstrate the performance of the NK method and compare it with DIIS.

2. COUPLED CLUSTER EQUATIONS

In this section, we briefly discuss the algebraic form of the CC equations for cluster amplitudes. In general, a correlated wave function $|\Psi\rangle$ can be written as,

$$|\Psi\rangle = \Omega|\Phi\rangle \,, \tag{1}$$

where $|\Phi\rangle$ is the reference wave function (typically the HF Slater determinant) and Ω is the wave operator. In the CC method, the wave operator is assumed in an exponential form

$$\Omega = e^T \,, \tag{2}$$

where T is the cluster operator defined by excitations producing excited Slater determinant when acting on the reference function. This property of the cluster operator T assures the so-called intermediate normalization of the CC wave function, i.e.,

$$\langle\Psi|\Phi\rangle = 1 \,, \tag{3}$$

assuming that orthonormal molecular basis set was used to discretize many-body problem of interest. The cluster operator T is a sum of its many-body components

$$T = T_1 + T_2 + \dots \,, \tag{4}$$

where T_n is the linear combination of excitation operators, which corresponds to n-tuple excitations. Thus, for single- and double-excitations one can write

$$T_1 = \sum_{i,j} t_i^a a_a^\dagger a_i \tag{5}$$

$$T_2 = \sum_{i<j;a<b} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \,, \tag{6}$$

$$\dots \tag{7}$$

The coefficients $t_{ij..}^{ab..}$ are the cluster amplitudes, which will be determined by solving CC equations. We use the standard notation, i.e., indices i, j denote occupied, a, b virtual, and p, q general molecular spin orbitals. $a_p^\dagger$ and a_q are fermionic creation or annihilation operators which satisfy set of anticommutation relations

$$\{a_p^\dagger, a_q\} = a_p^\dagger a_q + a_q a_p^\dagger = \delta_{pq} \tag{8}$$

$$\{a_p^\dagger, a_q^\dagger\} = \{a_p, a_q\} = 0 \,. \tag{9}$$

In the context of deriving algebraic form of CC equations, particle-hole formalism is invoked (Shavitt and Bartlett, 2009).

Inserting the CC ansatz into the Schrödinger equation and pre-multiplying from the left by e^{-T} yields

$$e^{-T} H e^T \Phi = E\Phi \,. \tag{10}$$

To employ various diagrammatic techniques to derive CC equations, it is useful to introduce the normal product form of the electronic Hamiltonian (H_N) defined as

$$H_N = H - \langle\Phi|H|\Phi\rangle = \sum_{pq} F_{pq} N\{p^\dagger q\} + \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle N\{p^\dagger q^\dagger sr\}\,, \tag{11}$$

Using the Baker–Campbell–Hausdorff formula, we get

$$\begin{aligned} e^{-T} H_N e^T &= H_N + [H_N, T] + \frac{1}{2}[[H_N, T], T] + \\ &+ \frac{1}{3!}[[[H_N, T], T], T] + \frac{1}{4!}[[[[H_N, T], T], T], T] \\ &\equiv (H_N e^T)_C\,, \end{aligned} \tag{12}$$

where subscript "C" corresponds to a connected part of a given operator expression. Since electronic Hamiltonians are defined by one- and two-body interactions, the above expansion terminates after quadruple commutator (11). For the derivation of the correlation energy expression and amplitude equations, we project the $e^{-T}H_N e^T|\Phi\rangle$ term to the bra vectors $\langle\Phi|$, $\langle\Phi_i^a|$, $\langle\Phi_{ij}^{ab}|$, etc.:

$$\Delta E_{\text{corr}} = \langle\Phi|H_N e^T|\Phi\rangle_C\,, \tag{13}$$

$$\langle\Phi_{ij\ldots}^{ab\cdots}|H_N e^T|\Phi\rangle_C = 0 \tag{14}$$

where $\langle\Phi_i^a|, \langle\Phi_{ij}^{ab}|, \ldots$ represent singly, doubly, etc., excited Slater determinants with respect to the reference function.

The general TCC wave function employs the following split-amplitude ansatz

$$\Omega_{\text{TCC}} = e^{T^{\text{ext}}+T^{\text{act}}}\,, \tag{15}$$

where T^{act} represents the active amplitudes obtained from the active space calculation. These amplitudes are kept constant when solving the amplitude equations, only T^{ext} are iterated. The T^{act} amplitudes are computed from the CI coefficients, extracted from the matrix product states wave function optimized during the DMRG calculation. We will use t to denote, collectively, the CCSD amplitudes in Equations (5) and (6) that are contained in T^{ext}. These amplitudes satisfy a nonlinear equation that can be derived from Equation (14). In the rest of the paper, we will simply write this equation as,

$$r(t) = 0. \tag{16}$$

3. ALGORITHMS FOR SOLVING THE CCSD EQUATION

In this section, we begin with a short description of a general scheme for solving the CCSD nonlinear equation using an inexact Newton's method. We review a commonly used diagonal approximation to the Jacobian, and then describe the NK method for solving the CCSD equation.

3.1. Inexact Newton's Method

Even though Equation (16) is only a second-order nonlinear equation, it is not easy to solve due to the large number of variables contained in t. One should remember that in the state-of-the-art CCSD calculations, the total number of sought cluster amplitudes exceeds 10^{10}. An iterative procedure is generally required to solve the equation numerically. The best known algorithm for solving a general system of nonlinear equations is the Newton's method. In the k+1st iteration of such a method, the approximation to the solution of Equation (16) is updated as

$$t^{(k+1)} = t^{(k)} - \left[J^{(k)}\right]^{-1} r(t^{(k)}), \tag{17}$$

where $t^{(k)}$ is the approximate solution obtained from the kth iteration, and $J^{(k)}$ is the Jacobian of $r(t)$ evaluated at $t^{(k)}$.

Because it is not practical to write down the Jacobian of $r(t)$ or its inverse analytically, we cannot use the Newton's method directly to solve the CC equation. Instead, an IN algorithm of the form

$$t^{(k+1)} = t^{(k)} - \left[\hat{J}^{(k)}\right]^{-1} r(t^{(k)}) \tag{18}$$

where $\hat{J}^{(k)}$ is an approximate Jacobian matrix evaluated at $t^{(k)}$, is often used. In Equations (17) and (18), we view t and $r(t)$ as a column vector with all amplitudes in (5) and (6) enumerated in some specific order, and the contracted tensor amplitudes in $r(t)$ enumerated in the same order.

In CCSD calculation, a common practice is to choose $\hat{J}$ as a diagonal matrix with HF orbital energy difference as the diagonal elements. This is justified because J is known to be diagonal dominant in many cases, and the diagonal matrix of HF orbital energy differences contributes most to the diagonal of J. Replacing J with $\hat{J}$ typically works well when the system is near equilibrium. In this case, the computational cost of the IN method is dominated by the tensor contraction cost for evaluating $r(t^{(k)})$ for each k, which has the complexity of $O(N^6)$ where N is the number of atomic basis used to discretize HF molecular orbitals.

For systems that do not satisfy this property, the diagonal approximation may not be sufficient. As a result, many IN iterations may be required to reach convergence, which is defined by the norm of $r(t)$ being less than a prescribed tolerance level τ. This will lead to extremely long wall clock time.

3.2. Newton–Krylov Method

Even though J is not explicitly available, it is possible to approximate the product of $J(t)$ with any tensor w that has the same dimension as t. This can be done through a finite difference calculation of the form

$$J(t)w \approx \frac{r(t+\delta w) - r(t)}{\delta}, \tag{19}$$

where δ is a small constant.

The possibility to approximate $J(t)w$ by one extra function evaluation allows us to solve the Newton correction equation

$$J(t^{(k)})\Delta = -r(t^{(k)}), \tag{20}$$

by a Krylov subspace-based iterative method such as the GMRES algorithm (Saad and Schultz, 1986) even when $J(t)$ is not explicitly available. The solution Δ is used to update the approximate amplitude via

$$t^{(k+1)} = t^{(k)} + \Delta. \tag{21}$$

This approach is often referred to as the NK method.

In Algorithm 1, we give a description of the simplest Newton-GMRES algorithm for solving the coupled clustered equation. We treat the CC amplitude t and tensors (r, w) of the same dimension as vectors, and denote the inner product of t and r simply as $\langle t, r\rangle$. We treat a set of tensors as a matrix, and use $V(:, j)$, i.e., the jth column of V, to denote the jth tensor in such a set. The vector e_1 used in this algorithm denotes a unit vector of length $j_g + 1$ with 1 in the first entry and 0 elsewhere.

Input: Initial guess $t^{(0)}$ to the solution of the coupled cluster equation; convergence tolerance *tol*; Maximum number of GMRES iterations allowed (j_g);
Output: More accurate approximation t to solution of the coupled cluster equation that satisfies $\|r(t)\| \leq tol$;

1: $k = 0$;
2: **while** $\|r(t^{(k)})\| > tol$ **do**
3: $\beta = \|r(t^{(k)})\|$;
4: $V(:, 1) \leftarrow r(t^{(k)})/\beta$;
5: **for** j = 1, 2,...j_g **do**
6: $w \leftarrow r[t^{(k)} + \delta V(:, j)]$;
7: $w \leftarrow (w - r(t^{(k)}))/\delta$;
8: $H(:, 1 : j) \leftarrow \langle V(:, 1 : j), w\rangle$;
9: $w \leftarrow w - V(:, 1 : j)H(1 : j, j)$;
10: $H(j + 1, j) = \|w\|$;
11: $V(:, j + 1) = w/H(j + 1, j)$;
12: **end for**
Solve the projected linear least squares problem
13: $\min_s \|Hs - e_1\beta\|$
14: $t^{(k+1)} = t^{(k)} - V(:, 1 : j_g)s$;
15: $k \leftarrow k + 1$;
16: Evaluate $r(t^{(k)})$;
17: **end while**

Algorithm 1: A Newton–Krylov method for solving the coupled cluster equation.

The outer k loop of Algorithm 1 performs the Newton update (21) in line 14, using $\Delta = Vs$ as the the approximate solution to the Newton correction equation, where V is an $n \times j_g$ matrix, where n is the total number of CCSD amplitudes, and j_g is the number of inner GMRES iterations. The inner j iteration of Algorithm 1 solves the Newton correction Equation (20) using the GMRES method. The GMRES method performs a Gram–Schmidt process to produce an orthonormal basis of the Krylvo subspace

$$\mathcal{K}(J, r_0) \equiv \{r_0, Jr_0, J^2r_0, ..., J^{j_g}r_0\},$$

where J is the Jacobian evaluated at a particular approximation to the CCSD amplitudes t, and r_0 is the function value of r in Equation (16) defined at such a t. This orthonormal basis is stored in columns of the V matrix. In exact arithmetic, this matrix satisfies the equation

$$JV = \tilde{V}H, \tag{22}$$

where V contains the leading j_g columns of $\tilde{V}$, which are orthonormal, i.e., $V^TV = I$, and H is a $(j_g + 1) \times j_g$ upper Hessenberg matrix. The approximation to the solution of (20) is represented as $\Delta = Vs$ for some vector s of length j_g. This vector can be solved from the least squares problem defined by the Galerkin projection

$$\min_s \|\tilde{V}^T(JVs - r_0)\| \tag{23}$$

It follows from Equation (22), $\tilde{V}^T\tilde{V} = I$, and the fact that the first column of V is $r_0/\|r_0\|$, solving Equation (23) is equivalent to solving

$$\min_s \|Hs - \beta e_1\|, \tag{24}$$

where $\beta = \|r_0\|$. This is the least squares problem solved on line 13 of the algorithm. The solution is used in line 14 to update the CCSD amplitude.

3.3. Precondition

An iterative procedure for computing the solution to the Newton correction Equation (20) can be accelerated by using a pre-conditioner P. Instead solving (20), we solve

$$P^{-1}J^{(k)}\Delta = -P^{-1}r(t^{(k)}), \tag{25}$$

with the hope that $P^{-1}J^{(k)}$ has a reduced conditioner number. The reduced condition number can lead to faster convergence.

It is well-known that the Jacobian associated with the projected coupled cluster equation can be partitioned as

$$J(t^{(k)}) = D + E(t^{(k)}), \tag{26}$$

where D is a diagonal matrix consisting of the difference between virtual and occupied HF orbital energies, and E is a complicated term that depends on the fluctuation potential (Helgaker et al., 2014). When the HF amplitude is relatively large, the first term is dominant. Hence the diagonal matrix can be used as a pre-conditioner for the iterative solver of Equation (20).

Applying such a pre-conditioner only requires adding an extra step before the **while** loop in Algorithm to compute the preconditioned right-hand side in Equation (25) and modifying line 7 of the algorithm to apply P^{-1} to the W tensor.

When D is ill conditioned due to the presence of near degenerate HF orbital energy levels, it may be necessary to

introduce a shift σ and use $D - \sigma I$ as the preconditioner. This *regularization* technique is similar to the level-shifting technique often used in HF SCF calculation (Saunders and Hillier, 1973).

When J is not dominated by D, i.e., when contribution from the HF amplitude is less significant, it is desirable to construct alternative pre-conditioners to accelerate the convergence of the NK method.

3.4. Stopping Criterion

It is well-known that fast convergence of the IN algorithm can be achieved even when the Newton correction Equation (20) is not solved to full accuracy (Eisenstat and Walker, 1994). This is especially true in earlier IN iterations in which the residual norm $\|r(t)\|$ is still relatively large. A general strategy proposed in Eisenstat and Walker (1994) is to solve the correction equation to satisfy the following condition:

$$\|J(t^{(k)})\Delta + r(t^{(k)})\| \leq \eta_k \|r(t^{(k)})\|, \tag{27}$$

for some small constant $0 < \eta_k < 1$. A sophisticated scheme was proposed in Eisenstat and Walker (1994) to ensure the so-called "global convergence." However, such a scheme requires backtracking and additional residual function evaluations and is thus likely to increase the overall cost of the NK method. In this work, we use a simple strategy and set η_k to 10^{-1} with the left-hand side of Equation (27) estimated by the projected residual norm evaluated at the least squares solution to Equation (24). As the residual norm $\|r(t^{(k)})\|$ decreases in the outer iteration, the absolute error in the approximate solution to the correction equation also decreases when Equation (27) is satisfied. Note that there is a trade-off between the number of inner GMRES iterations required to solve the Newton correction equation and the number of NK outer iterations. Setting η_k to a small number may result in too many GMRES iterations performed at each NK iteration and thus increase the overall cost the algorithm. In our implementation, we set a limit on the maximum number of GMRES iterations allowed in each Newton iteration. Our computational experiments show that it is usually sufficient to limit the maximum number of GMRES iterations to 5. Furthermore, η_k can also be chosen in a dynamic way with η_k being relatively large for small k and relatively small for large k.

4. COMPARISON WITH DIIS

The DIIS method (Pulay, 1980) is a commonly used technique to accelerate the convergence of iterative method for solving the CC equation. At the kth iteration, we form a new approximation as

$$\tilde{t}^{(k+1)} = \sum_{j=k-\ell}^{k} \omega_{k-j} \left[t^{(j)} + \Delta^{(j)} \right], \tag{28}$$

for some constant $\ell < k$, where ω_j's are chosen to be the solution to the following constrained minimization problem

$$\min_{\sum_j \omega_j = 1} \| \sum_j \omega_j \Delta^{(k-j)} \|. \tag{29}$$

The $k+1$st amplitude approximation is then computed from

$$t^{(k+1)} = \tilde{t}^{(k+1)} - \tilde{\Delta}^{(k+1)}, \tag{30}$$

where $\tilde{\Delta}^{(k+1)}$ is the approximate solution to the Newton correction Equation (20) or (25).

Note that in some formulations of the DIIS algorithm, the $\Delta^{(k-j)}$ term in the objective of Equation (29) are simply replaced by $r(t^{(k-j)})$. Since $\tilde{\Delta}^{(k-j)}$ is often computed as $D^{-1}r(t^{(k-j)})$, where D is the diagonal component in Equation (26), these two formulations are equivalent up to a scaling matrix D.

In some implementations of the DIIS acceleration method, one performs a fixed number of IN iteration with D in Equation (26) as the approximate Jacobian matrix before the DIIS procedure is used to update $t^{(k+1)}$ according to Equation (30). In other implementations, DIIS is performed in each IN iteration.

The convergence of the DIIS method and its connection with the Broyden's method (Dennis and Schabel, 1996) has been analyzed in Rohwedder and Schneider (2011) and Walker and Ni (2011). The connection between DIIS and Krylov subspace method is made in Harrison (2004) and Ettenhuber and Jrgensen (2015).

One of the practical issues one needs to consider when implementing the DIIS method is the solution of the constrained minimization problem (29). A commonly used approach in existing quantum chemistry software is to write down the linear equation representing the first-order necessary condition of Equation (29) and solve the equation using a Cholesky or LU factorization based method. This approach is not numerically stable, especially when the set of $\{\Delta^{(k-j)}\}$ becomes nearly linearly dependent. A more stable way to solve Equation (29) is to turn it into a unconstrained least squares problem and obtain the optimal solution via a rank-reveal QR factorization of the matrix consisting of $\Delta^{(k-j)}$ as its columns. However, applying rank-reveal QR to $\{\Delta^{(k-j)}\}$ is rather costly. In comparison, performing a rank-reveal QR factorization for solving the projected least squares problem (24) in the NK method is relatively straightforward, and does not introduce significant overhead. To improve numerical stability and computational efficiency, it may be necessary to keep only a subset of $\{\Delta^{(k-j)}\}$'s for a small number of j's.

In terms of memory usage, NK is slightly more efficient. In addition to storing the current approximation to the CC amplitudes and the residual, NK also stores orthonormal basis tensors of the Krylov subspace used to obtain approximate solution to the Newton's correction equation. The DIIS method typically needs to store a set of $\{\Delta^{(k-j)}\}$'s as well as the corresponding set of previous amplitude approximations.

Although one can view the DIIS method as a way to solve the Newton correction Equation (20) (Rohwedder and Schneider, 2011), it is also possible to combine DIIS acceleration with the NK procedure. In such a hybrid scheme, we simply use GMRES to compute $\tilde{\Delta}^{(k+1)}$ correction in Equation (30) after $\tilde{t}^{(k+1)}$ is obtained from a DIIS update.

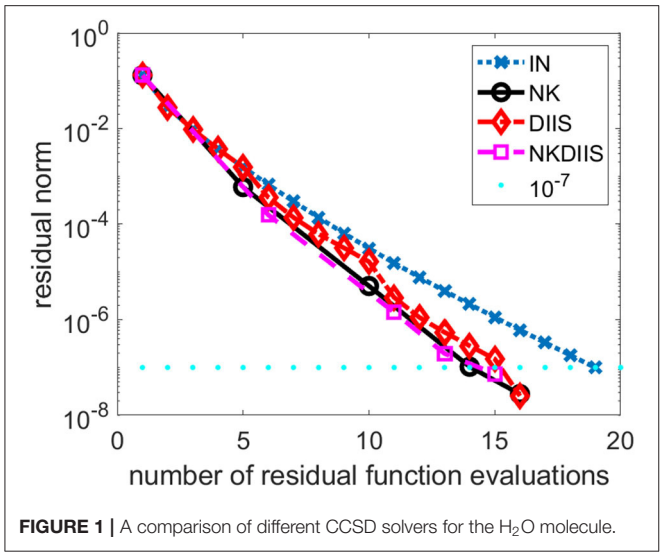

FIGURE 1 | A comparison of different CCSD solvers for the H_2O molecule.

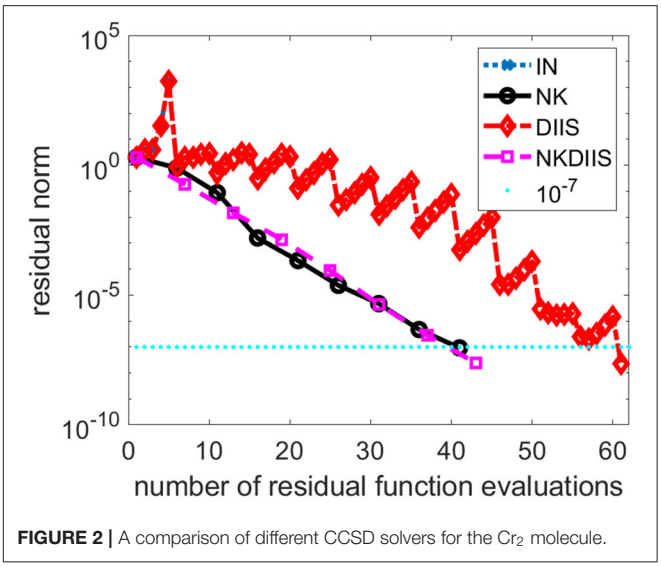

FIGURE 2 | A comparison of different CCSD solvers for the Cr_2 molecule.

5. RESULTS AND DISCUSSION

In this section, we present a few numerical examples to demonstrate the effectiveness of the NK method and compare with the inexact Newton method accelerated by DIIS. All algorithms compared below were implemented using the NWChem software (Valiev et al., 2010) version 6.6.

5.1. H_2O Molecule

In the first example, we show how the NK algorithm behaves when it is applied to a simple water molecule in equilibrium. We use the cc-pvtz basis set to discretize the problem. We compare the NK algorithm with the IN method in which the Jacobian is approximated the diagonal matrix D in Equation (26), and the IN method accelerated by the DIIS procedure (labeled as DIIS). When DIIS is used to accelerate convergence, it is applied every 5 IN iterations. We set the convergence tolerance to 10^{-7}, i.e., we terminate the IN, DIIS, and NK iterations when the Euclidean norm of the residual $r(t)$ falls below 10^{-7}.

In **Figure 1**, we plot the change of residual norm of each method with respect to the cumulative number of CCSD residual function evaluations (tensor contractions). Note that in the IN and DIIS runs, the number of function evaluations is equivalent to the number of IN iterations. However, in the NK run, the number of function evaluations is the total number of inner GMRES iterations and the number of outer NK iterations.

We observe that for this relatively easy problem, the IN method converges without DIIS acceleration. It takes 19 iterations (and 19 function evaluations) to reach convergence. The use of DIIS acceleration reduces the total number function evaluations to 16. This is also the number of function evaluations used in the NK method. By combining NK and DIIS, we reduce the number of functions by 1.

5.2. Cr_2

In this section, we show how NK performs on a Cr_2 molecular. The interatomic distance between two Cr atoms is set to 1.7 angstrom, which is near equilibrium. We use the cc-pvdz basis set to discretized the problem.

This is a relatively difficult problem. As we can see from **Figure 2**, without DIIS acceleration, the IN iteration diverges quickly. Even when the DIIS acceleration is activated, which is applied every 5 IN iterations, the change of residual norm has a zig-zag pattern with the residual norm decreasing only after a DIIS step. It takes a total of 61 residual evaluations before convergence is reached. Both NK and the hybrid NK and DIIS converge rapidly. We used a maximum of 5 GMRES iterations in each NK iteration. The residual norm decreases monotonically in both runs. There is very little difference between the two.

As we indicated earlier, there is a tradeoff between performing more GMRES iterations in the inner loop of the NK method and the number of NK iterations. In **Figure 3**, we compare the total number of NK iterations and the total number of function evaluations for several NK runs in which different numbers of GMRES iterations were performed in each outer NK iteration.

We can see from **Figure 3** that performing 3 GMRES iterations per NK iteration is not enough to achieve rapid convergence. On the other hand, taking too many GMRES iterations does not help either, especially in the first few NK iterations when the residual norm is still relatively large. For this problem, setting the maximum number of GMRES iterations per NK iteration to 5 appears to yield best performance.

In **Figure 4**, we also show the change of the relative GMRES residual norm defined as $\|Hs - \beta e_1\|/\beta$, where H, s, e_1, and β are as defined in Equation (24), with respect to the GMRES iteration number during the first and the 10th NK iterations when the number of GMRES iterations is fixed at 10. We observe that the GMRES iteration converge slowly in the first NK iteration when the CCSD amplitudes are relatively far from the solution. As the

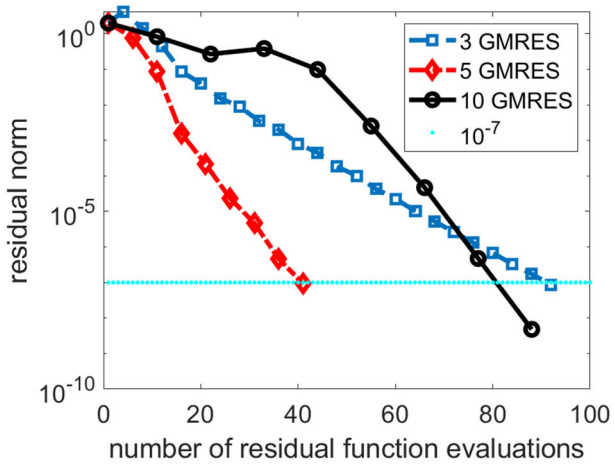

FIGURE 3 | A comparison of three Newton–Krylov (NK) runs that use 3, 5 and 10 Generalized Minimum Residual (GMRES) iterations per NK iteration.

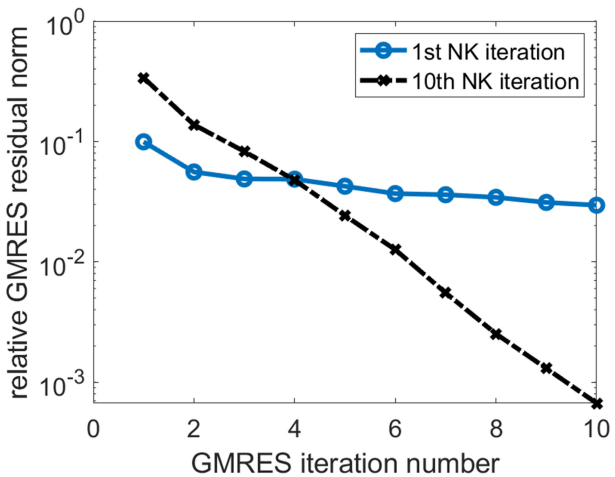

FIGURE 4 | The change of relative GMRES residual norm in the first and 10th NK iterations.

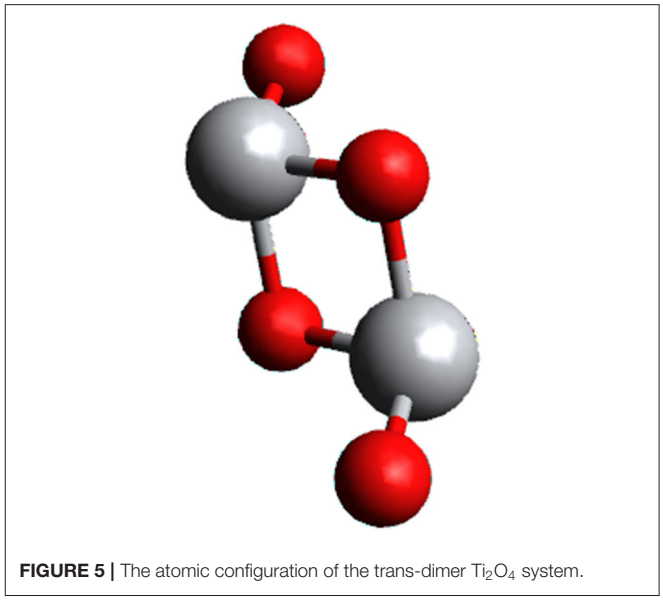

FIGURE 5 | The atomic configuration of the trans-dimer Ti_2O_4 system.

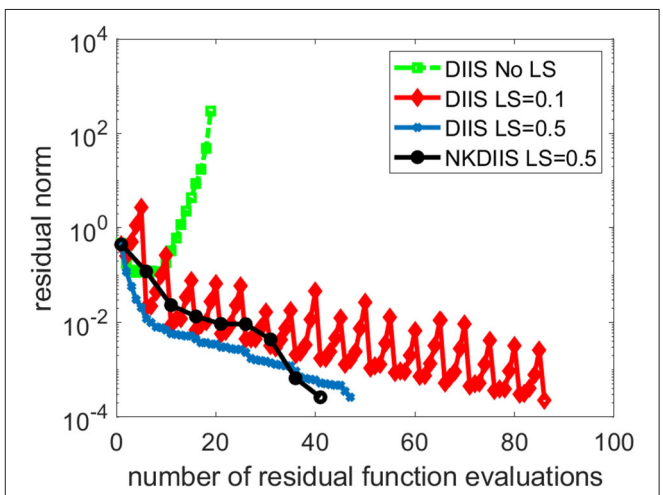

FIGURE 6 | A comparison of different CCSD solvers for the trans-dimer Ti_2O_4 molecule.

amplitudes become closer to the solution, the GMRES iteration converges faster.

5.3. Trans-dimer Ti_2O_4

The next example we use for testing the NK method is a titanium oxide system. Its geometry is shown in **Figure 5**. We use the aug-cc-pvtz basis set to discretize the problem.

This is a difficult problem. **Figure 6** shows that DIIS fails to converge. In fact, the residual norm quickly increases. This is mainly caused by the fact that the D matrix used in the IN method is extremely ill-conditioned. To overcome this difficulty, we regularize the NK calculation by subtracting a constant shift σ from the diagonal of D. The same level-shifting is used in the IN accelerated by DIIS. **Figure 6** shows that DIIS converges when σ is chosen to be 0.1. However, the convergence is rather slow with this choice of level shift. By setting σ to 0.5, we can achieve much faster convergence. The shift can also be applied to in NK when $D - \sigma I$ is used as a preconditioner in the GMRES iteration. For this problem, the NK (combined with DIIS) converges with 41 residual function evaluations, which is fewer than the 47 function evaluations required in DIIS. Similar convergence is observed for NK without DIIS (which we do not plot here) also.

5.4. oxo-Mn (salen)

We now compare the performance of NK and DIIS for an oxo-Mn (salen) molecule shown in **Figure 7**. In practical applications, this system catalyzes enantioselective epoxidation of unfunctional olefines and represents an important substance

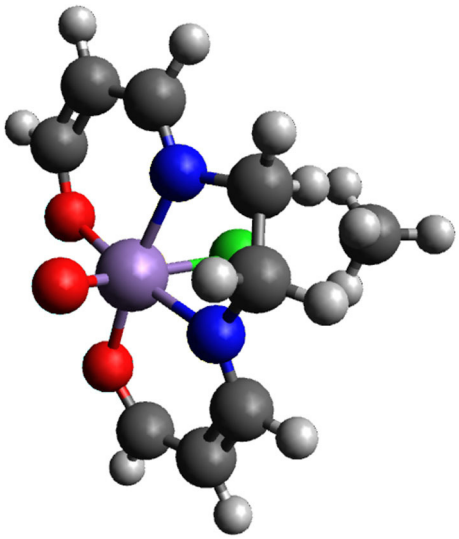

FIGURE 7 | The atomic configuration of the oxo-Mn (salen) molecule.

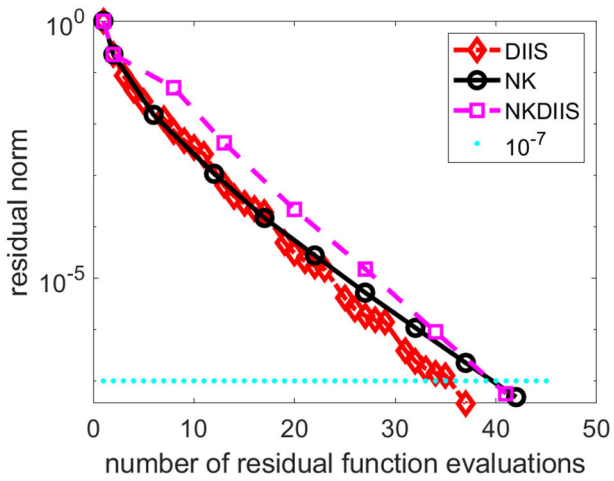

FIGURE 8 | A comparison of direct inversion of iterative subspace (DIIS) and NK combined with DIIS (NKDIIS) convergence for the oxo-Mn (salen) molecule.

in industry. Due to the quasi-degeneracy of the lowest states, this system is a considerable challenge even for multireference methods and has been intensively studied recently (Irie et al., 1990; Zhang et al., 1990; Antalík et al., 2019).

We have performed TCCSD calculations in CAS(28,22) space, where external amplitudes were obtained from DMRG calculation.

Figure 8 shows the convergence history of the IN iteration accelerated by DIIS, NK, and NK combined with DIIS (NKDIIS). For this example, the IN iteration appears to converge very fast. Both the NK and NKDIIS are slightly slower, but converges in 10 NK iterations.This example shows that for systems with multireference features, it is important to use an appropriate model that can treat the multireference character of the molecule effectively. With such a model, the CCSD nonlinear equation can become easier to solve. Although NK may not offer too much advantage in this case, it is still an effective solver for such as model.

5.5. The G2/97 Dataset

In addition to testing the NK algorithm on the above representative molecules, we performed a more extensive testing of the algorithm on a much wider range of molecules randomly selected from the G2/97 dataset (Curtiss et al., 1997). Among 130 systems we tested, NK performs better on 123 of them. On average, NK uses 12% fewer function evaluations when compared with DIIS. In the best case, NK uses 33% fewer function evaluations. In the worst case, NK uses 40% more function evaluations.

6. CONCLUSION

We presented a NK method for solving CC amplitude equations. In such a method, the Newton correction equation is solved by a Krylov subspace iterative method such as the GMRES method. Preconditioners can be applied in the iterative solver to accelerate convergence. We discussed the trade-off between performing more inner (GMRES) iteration and outer Newton iteration, and suggested an adaptive stopping criterion for the inner iteration. We compared the NK method with the widely used DIIS method and showed how the two methods can be combined. We presented several numerical examples to demonstrate the effectiveness and robustness of the NK method not only for standard CCSD calculations but also for tailed CCSD calculations where the information for external correction is obtained from a DMRG calculation. Although the results we presented in this paper are on developments made in an older version NWChem software, the NK has been implemented in the next generation of NWChem software (NWChemEx) (Richard et al., 2019) designed for exa-scale high-performance computing platforms.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of 474 Science and the National Nuclear Security Administration (CY, DW-Y, and KK). This

research used resources of the National Energy Scientific Computing Center (NERSC) and the Oak Ridge Leadership Computing Facility, which are DOE Office of Science User Facilities supported under Contracts DE-AC05-00OR22725. JB acknowledges support from the Czech Science Foundation (Grant No. 19-13126Y).

REFERENCES

Antalík, A., Veis, L., Brabec, J., Demel, O., Örs Legeza, and Pittner, J. (2019). Toward the efficient local tailored coupled cluster approximation and the peculiar case of oxo-mn(salen). *J. Chem. Phys*. 151:084112. doi: 10.1063/1.5110477

Bartlett, R. J., and Musiał, M. (2007). Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys*. 79, 291–352. doi: 10.1103/RevModPhys.79.291

Čížek, J. (1966). On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys*. 45, 4256–4266. doi: 10.1063/1.1727484

Curtiss, L. A., Raghavachari, K., Redfern, P. C., and Pople, J. A. (1997). Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *J. Chem. Phys*. 106, 1063–1079. doi: 10.1063/1.473182

Dennis, J. E., and Schabel, R. B. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Philadelphia, PA: SIAM. doi: 10.1137/1.9781611971200

Eisenstat, S. C. and Walker, H. F. (1994). Globally convergent inexact Newton methods. *SIAM J. Optimiz*. 4, 393–422. doi: 10.1137/0804022

Ettenhuber, P., and Jrgensen, P. (2015). Discarding information from previous iterations in an optimal way to solve the coupled cluster amplitude equations. *J. Chem. Theory Comput*. 11, 1518–1524. doi: 10.1021/ct501114q

Gauss, J. (1998). "Coupled cluster theory," in *The Encyclopedia of Computational Chemistry*, eds V. R. Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer III, and P. R. Scheiner (Wiley: Chichester), 615–636.

Harrison, R. J. (2004). Krylov subspace accelerated inexact newton method for linear and nonlinear equations. *J. Comput. Chem*. 25, 328–334. doi: 10.1002/jcc.10108

Helgaker, T., Jorgensen, P., and Olsen, J. (2014). *Molecular Electronic-Structure Theory*. Chichester: John Wiley & Sons, Ltd.

Irie, R., Noda, K., Ito, Y., Matsumoto, N., and Katsuki, T. (1990). Catalytic asymmetric epoxidation of unfunctionalized olefins. *Tetrahedron Lett*. 31, 7345–7348. doi: 10.1016/S0040-4039(00)88562-7

Kinoshita, T., Hino, O., and Bartlett, R. J. (2005). Coupled-cluster method tailored by configuration interaction. *J. Chem. Phys*. 123:074106. doi: 10.1063/1.2000251

Kjønstad, E. F., Folkestad, S. D., and Koch, H. (2020). Accelerated multimodel newton-type algorithms for faster convergence of ground and excited state coupled cluster equations. *J. Chem. Phys*. 153:014104. doi: 10.1063/5.0010989

Knoll, D., and Keyes, D. (2004). Jacobian-free Newton-Krylov methods: a survey of approaches and applications. *J. Comput. Phys*. 193, 357–397. doi: 10.1016/j.jcp.2003.08.010

Li, X. (2001). Benchmark study of potential energies and vibrational levels using the reduced multireference coupled cluster method. The HF molecule. *J. Mol. Struct*. 547, 69–81. doi: 10.1016/S0166-1280(01)00460-2

Li, X., and Paldus, J. (1997). Reduced multireference CCSD method: an effective approach to quasidegenerate states. *J. Chem. Phys*. 107, 6257–6269. doi: 10.1063/1.474289

Lyakh, D. I., Lotrich, V. F., and Bartlett, R. J. (2011). The tailored CCSD(T) description of the automerization of cyclobutadiene. *Chem. Phys. Lett*. 501, 166–171. doi: 10.1016/j.cplett.2010.11.058

Lyakh, D. I., Musiał, M., Lotrich, V. F., and Bartlett, R. J. (2012). Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev*. 112, 182–243. doi: 10.1021/cr2001417

Paldus, J., and Li, X. (1999). *A Critical Assessment of Coupled Cluster Method in Quantum Chemistry*. Chichester: John Wiley & Sons, Ltd. 1–175. doi: 10.1002/9780470141694.ch1

Piecuch, P., and Adamowicz, L. (1994). Solving the single-reference coupled-cluster equations involving highly excited clusters in quasidegenerate situations. *J. Chem. Phys*. 100, 5857–5869. doi: 10.1063/1.467149

Pulay, P. (1980). Convergence acceleration of iterative sequences: the case of SCF iteration. *Chem. Phys. Lett*. 73, 393–398. doi: 10.1016/0009-2614(80)80396-4

Purvis, G., and Bartlett, R. (1982). A full coupled-cluster singles and doubles model: the inclusion of disconnected triples. *J. Chem. Phys*. 76, 1910–1918. doi: 10.1063/1.443164

Purvis, G. D. III, and Bartlett, R. J. (1981). The reduced linear equation method in coupled cluster theory. *J. Chem. Phys*. 75, 1284–1292. doi: 10.1063/1.442131

Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M. (1989). A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett*. 157, 479–483. doi: 10.1016/S0009-2614(89)87395-6

Richard, R. M., Bertoni, C., Boschen, J. S., Keipert, K., Pritchard, B., Valeev, E. F., et al. (2019). Developing a computational chemistry framework for the exascale era. *Comput. Sci. Eng*. 21, 48–58. doi: 10.1109/MCSE.2018.2884921

Rohwedder, T., and Schneider, R. (2011). An analysis for the DIIS acceleration method used in quantum chemistry calculations. *J. Math. Chem*. 49, 1889–1914. doi: 10.1007/s10910-011-9863-y

Saad, Y., and Schultz, M. H. (1986). GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput*. 7, 856–869. doi: 10.1137/0907058

Saunders, V. R., and Hillier, I. H. (1973). A level-shifting method for converging closed shell Hartree-Fock wavefunctions. *Int. J. Quantum Chem*. 7, 699–705. doi: 10.1002/qua.560070407

Shavitt, I., and Bartlett, R. J. (2009). *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511596834

Valiev, M., Bylaska, E., Govind, N., Kowalski, K., Straatsma, T., Van Dam, H., et al. (2010). NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun*. 181, 1477–1489. doi: 10.1016/j.cpc.2010.04.018

Walker, H. W., and Ni, P. (2011). Anderson acceleration for fixed-point iterations. *SIAM J. Num. Anal*. 49, 1715–1735. doi: 10.1137/10078356X

Zhang, W., Loebach, J. L., Wilson, S. R., and Jacobsen, E. N. (1990). Enantioselective epoxidation of unfunctionalized olefins catalyzed by salen manganese complexes. *J. Am. Chem. Soc*. 112, 2801–2803. doi: 10.1021/ja00163a052

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 10 December 2020
doi: 10.3389/fchem.2020.581058

On the Efficient Evaluation of the Exchange Correlation Potential on Graphics Processing Unit Clusters

David B. Williams-Young[1]*, *Wibe A. de Jong*[1], *Hubertus J. J. van Dam*[2] and *Chao Yang*[1]

[1] Lawrence Berkeley National Laboratory, Computational Research Division, Berkeley, CA, United States, [2] Brookhaven National Laboratory, Computational Science Initiative, Upton, NY, United States

The predominance of Kohn–Sham density functional theory (KS-DFT) for the theoretical treatment of large experimentally relevant systems in molecular chemistry and materials science relies primarily on the existence of efficient software implementations which are capable of leveraging the latest advances in modern high-performance computing (HPC). With recent trends in HPC leading toward increasing reliance on heterogeneous accelerator-based architectures such as graphics processing units (GPU), existing code bases must embrace these architectural advances to maintain the high levels of performance that have come to be expected for these methods. In this work, we purpose a three-level parallelism scheme for the distributed numerical integration of the exchange-correlation (XC) potential in the Gaussian basis set discretization of the Kohn–Sham equations on large computing clusters consisting of multiple GPUs per compute node. In addition, we purpose and demonstrate the efficacy of the use of batched kernels, including batched level-3 BLAS operations, in achieving high levels of performance on the GPU. We demonstrate the performance and scalability of the implementation of the purposed method in the NWChemEx software package by comparing to the existing scalable CPU XC integration in NWChem.

Keywords: density functional theory, graphics processing unit, high-performance computing, parallel computing, quantum chemistry

OPEN ACCESS

Edited by:
Mohan Chen,
Peking University, China

Reviewed by:
Phanish Suryanarayana,
Georgia Institute of Technology, United States
Yu Shen,
University of Science and Technology of China, China
Michele Pavanello,
Rutgers University, Newark, United States

***Correspondence:**
David B. Williams-Young
dbwy@lbl.gov

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *07 July 2020*
Accepted: *14 September 2020*
Published: *10 December 2020*

Citation:
Williams-Young DB, de Jong WA, van Dam HJJ and Yang C (2020) On the Efficient Evaluation of the Exchange Correlation Potential on Graphics Processing Unit Clusters. Front. Chem. 8:581058. doi: 10.3389/fchem.2020.581058

1. INTRODUCTION

Kohn-Sham density functional theory (KS-DFT) (Hohenberg and Kohn, 1964; Kohn and Sham, 1965) is unequivocally the computational workhorse of theoretical chemistry and materials science. With the excellent balance of its computational cost to its ability to accurately predict physical phenomena, KS-DFT is nearly without equal in the routine theoretical treatment of large, experimentally relevant systems (Ratcliff et al., 2017; Wu et al., 2019). A primary factor contributing to the popularity of KS-DFT methods is the existence of highly optimized and scalable software implementations capable of leveraging the latest advances in modern high-performance computing (HPC). The existence of such software enables the treatment of increasingly larger and more complicated systems as computing resources become large enough to accommodate them. Historically, these optimizations have amounted to considering the underlying details of homogeneous computing platforms such as shared and distributed memory multi-core central processing unit (CPU) architectures to exploit memory hierarchies, distributed node topology and interconnection, and computing features such as single-instruction multiple data (SIMD)

instructions, fused multiply-add (FMA), etc. (Belling et al., 1999; Brown et al., 2008; Lasinski et al., 2008; de Jong et al., 2010; Bylaska et al., 2017; Jacquelin et al., 2017; Nguyen et al., 2017; Petrone et al., 2018) However, as we approach the exascale computing era, the emergence of more heterogeneous computing architectures renders non-trivial the direct application of existing algorithms and code bases to target these complex architectures. As such, for KS-DFT to remain relevant in the age of exascale and post-exascale computing, methods developers must be prepared to embrace these emerging architectures to maintain the high standard of computational performance which has come to be expected.

In recent years, the trajectory of HPC has lead to an increasing reliance on the use accelerators, such as graphics processing units (GPU), to perform the majority of the floating point operations (FLOPs) on new and emerging computing resources (Kindratenko et al., 2009; Parnell et al., 2019). For a detailed treatise on the details and challenges presented by these and other emerging architectures and their use in conjunction with electronic structure calculations, we refer to the work of Gordon et al. (2020). In this work, we limit our discussion to the optimization of KS-DFT methods on NVIDIA GPUs (in particular the NVIDIA Tesla V100) using the Compute Unified Device Architecture (CUDA) programming platform (Cook, 2012).

Recently, there has been significant research effort afforded to porting electronic structure software to the GPU (Gordon et al., 2020). In the case of large-scale calculations, much work has gone into the development of massively parallel GPU implementations of methods based on plane wave (Maintz et al., 2011; Wang et al., 2011; Jia et al., 2019), real space (Andrade and Aspuru-Guzik, 2013; Hakala et al., 2013), finite element (Das et al., 2019; Motamarri et al., 2020), and various other discretizations (Genovese et al., 2009; van Schoot and Visscher, 2016; Yoshikawa et al., 2019; Huhn et al., 2020) of the Kohn–Sham equations. In this work, we consider the Gaussian basis set discretization of the Kohn–Sham equations (Pople et al., 1992), which poses a number of challenges for GPU implementations. The majority of these challenges revolve around the computation of molecular integrals over Gaussian basis functions. Of the required integrals, the electron repulsion integrals (ERIs) and the exchange-correlation (XC) potential are among the most costly and the most challenging to port to GPU architectures. Over the years, there has been a considerable amount of research devoted to porting implementations of Gaussian basis set KS-DFT to the GPU (Yasuda, 2008; Brown et al., 2010; Titov et al., 2013; Luehr et al., 2016; Kussmann and Ochsenfeld, 2017; Manathunga et al., 2020; Peters et al., 2020); however, the vast majority of this work has been centered around the evaluation and digestion of the ERIs in the construction of the Fock matrix (Ufimtsev and Martinez, 2008, 2009a,b; Asadchev et al., 2010; Miao and Merz, 2013; Kalinowski et al., 2017; Kussmann and Ochsenfeld, 2017; Laqua et al., 2020). On the other hand, the XC potential has received much less treatment in the literature in this regard (Yasuda, 2008; Luehr et al., 2016; Manathunga et al., 2020). This disparity is understandable due to the fact that for large systems, the ERI-related contributions to the Fock matrix are computationally dominant and the most challenging to parallelize. However, with recent advances in semi-numerical techniques for exact exchange, which have shown great promise in early GPU implementations (Laqua et al., 2020), ERI-dominated calculations are quickly becoming computationally competitive with the evaluation of the XC potential by current methods. Further, current accounts of GPU implementations of the XC integration have been limited to the devices which are accessible within a particular compute node. To the best of the authors' knowledge, there does not exist a GPU accelerated distributed memory evaluation of the XC potential using Gaussian basis sets as of this report. Thus, in this work, we propose a three-level parallelism scheme for the scalable distributed evaluation of the Gaussian basis XC potential on large clusters of GPUs.

In general, there are a number of important features of GPU architectures one must consider in the development of high-performance software:

- GPU architectures exhibit orders of magnitude more computational threads than CPU architectures, allowing for the expression of massive concurrency within a single GPU device.
- The memory space which is directly accessible to GPU devices is much lower in capacity in comparison with their CPU counterparts (O(16–32 GB) on the GPU in comparison to upwards of O(1 TB) on the CPU).
- Memory access within device memory exhibits a much higher bandwidth than CPU memory (O(900 GB/s) on the GPU in comparison to O(20–50 GB/s) on the CPU).
- Data transfers between host and device memory spaces are low bandwidth [O(80 GB/s) with advanced technologies such as NVLink, O(35 GB/s) over PCIe], thus data transfers often pose a non-trivial overhead in GPU applications which require movement of large volumes of data.

A consequence of these features is that, despite the large number of threads that are available to the GPU to perform computation, data locality must be carefully tuned to exploit the low capacity device memory as to allow for the expression of concurrency but also to avoid high cost and inherently serial data transfers between host and device. As such, those algorithms which are able to express massive concurrency on local data without being interrupted by synchronization points such as data transfers and memory allocations are typically the best suited for GPU application. A key aspect of the method proposed in this report is the optimization of data movement within the XC integration as to express massive concurrency using data that resides in device memory without transfers between host and device.

Scientific applications often rely on the existence of highly tuned linear algebra libraries (such as vendor implementations of BLAS and LAPACK) to achieve high levels of performance on contemporary and emerging architectures (Dongarra et al., 1998). Over the years, many areas of matrix computation have achieved significant performance improvements through the use of GPU accelerators (Fatahalian et al., 2004; Kurzak et al., 2012; Herault et al., 2019). However, unless the matrix computations needed by a particular application are large enough as to fully

exploit the resources of the device, it is unlikely that single matrix operation such as matrix–matrix multiplication will be able to achieve high computational occupancy on the device. An important achievement in high-performance numerical linear algebra has been the advent of highly tuned batched implementations of commonly encountered matrix operations, such as matrix–matrix multiplication, triangular factorization, etc. (Haidar et al., 2015; Abdelfattah et al., 2016a). Such batched implementations are provided in both vendor tuned (such as cuBLAS and cuSOLVER provided by NVIDIA) and open source (such as MAGMA, Nath et al., 2010; Tomov et al., 2010; Abdelfattah et al., 2016b) GPU accelerated linear algebra libraries. In these batched implementations, efficiency is achieved by dramatically increasing the throughput of the matrix operations via concurrent execution within a single device. Thus, if an application requires the manipulation of many small matrices in a manner that allows for concurrent execution (such as KS-DFT), large performance improvements can be made by utilizing these batched implementations (see e.g., Das et al., 2019). GPU-accelerated BLAS has previously been used in the context of XC computations (Yasuda, 2008). In this work, we examine the use of batched BLAS to further accelerate these operations to improve overall time-to-solution.

This work will be organized as follows. Sections 2.1 and 2.2 will briefly review the pertinent theory and high-level algorithmic constructs related to the XC integration. Section 2.3 will then describe the proposed method for the scalable, three-level parallelism scheme for the distributed XC integration on clusters of GPUs. Section 3 will demonstrate the performance and scalability of the purposed method in comparison to an existing high-performance CPU implementation using a wide range of molecules, basis sets, and quadrature sizes. Finally, section 4 will conclude this work and offer insight into the impact of the purposed method and briefly discuss future research directions.

2. METHODS

2.1. Kohn–Sham Density Functional Theory

In KS-DFT, the total electronic energy within a particular density functional approximation (DFA) takes the form (Parr and Yang, 1994)

$$\mathcal{E}^{tot} = \mathcal{T}_s + \mathcal{V}_{ne} + \mathcal{J} - c_x\mathcal{K} + \mathcal{E}^{xc}, \tag{1}$$

where $\mathcal{T}_s$ and $\mathcal{V}_{ne}$ are the (non-interacting) kinetic and electron-nuclear attraction energies, and $\mathcal{J}$ and $\mathcal{K}$ are the classical Coulomb and exact exchange energies, respectively. $c_x \in \mathbb{R}$ is a parameter that scales the contribution of exact-exchange to the electronic energy. $c_x = 0$ is used for "pure" DFAs, whereas DFAs that use $c_x \neq 0$ are referred to as "hybrid" DFAs (Becke, 1993). Without loss of generality in the following, we will take $c_x = 0$, though we note that the algorithms presented in the following sections may also be extended to hybrid methods without modification. $\mathcal{E}^{xc}$ is the exchange-correlation (XC) energy which is taken to be a functional of the electron density $\rho : \mathbb{R}^3 \to \mathbb{R}$. In this work, we restrict our discussion to spin-restricted DFAs within the generalized gradient approximation (GGA) (Perdew, 1986; Perdew and Yue, 1986), i.e. $\mathcal{E}^{xc}$ is approximated to only depend on ρ and its gradient $\nabla\rho : \mathbb{R}^3 \to \mathbb{R}^3$. We note for completeness that the information presented in this and the following sections may be extended to both spin-unrestricted and spin-generalized KS-DFT methods as well as more advanced DFAs (such as the meta-GGA) with the addition of only a few intermediates (Egidi et al., 2017; Petrone et al., 2018). As $\nabla\rho$ is a vector valued quantity, and thus dependent on reference frame quantities such as molecular orientation, it is canonical to express $\mathcal{E}^{xc}$ as

$$\mathcal{E}^{xc} = \int_{\mathbb{R}^3} \varepsilon(\{U(\mathbf{r})\})\rho(\mathbf{r})\mathrm{d}^3\mathbf{r}, \tag{2}$$

where ε is an energy density that depends on a set of so-called "U" -variables, $\{U(\mathbf{r})\}$, which are independent of reference frame. Within the GGA, the canonical choice for these variables are $\{U(\mathbf{r})\} = \{\rho(\mathbf{r}), \gamma(\mathbf{r})\}$ with $\gamma(\mathbf{r}) = \|\nabla\rho(\mathbf{r})\|$.

By expanding the density in a finite set of basis functions, $\mathcal{S} = \{\phi_\mu(\mathbf{r})\}_{\mu=1}^{N_b}$,

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}), \tag{3}$$

where $\mathbf{P}$ is the density matrix, the Kohn–Sham Fock matrix takes the form (Parr and Yang, 1994)

$$\mathbf{F} = \mathbf{h} + \mathbf{J} + \mathbf{V}^{xc}. \tag{4}$$

$\mathbf{h}$ is the basis representation of the density-independent core Hamiltonian (e.g., the sum of kinetic energy and external potential operators), and $\mathbf{J}$ is the basis representation of the classical Coulomb operator. Note that we have dropped the exact exchange term in Equation (1) as we have taken $c_x = 0$. $\mathbf{V}^{xc}$ is the XC potential that may be expressed as (Yasuda, 2008; Burow and Sierka, 2011; Petrone et al., 2018)

$$V^{xc}_{\mu\nu} = \int_{\mathbb{R}^3} \phi_\mu(\mathbf{r})Z_\nu(\mathbf{r}) + Z_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\mathrm{d}^3\mathbf{r}, \tag{5}$$

where

$$Z_\mu(\mathbf{r}) = \frac{1}{2}\frac{\partial\varepsilon(\{U(\mathbf{r})\})}{\partial\rho}\phi_\mu(\mathbf{r}) + 2\frac{\partial\varepsilon(\{U(\mathbf{r})\})}{\partial\gamma}\nabla\rho(\mathbf{r})\cdot\nabla\phi_\mu(\mathbf{r}). \tag{6}$$

Note that the partial derivatives of ε are to be evaluated with the U-variables calculated at argument of Z_μ.

Equations (4) to (6) are general to any (real-valued) basis set expansion. In this work, we consider atomically centered contracted Gaussian basis functions of the form

$$\phi_\mu(\mathbf{r}) = (x-R_x)^l(y-R_y)^m(z-R_z)^n \sum_{\xi=1}^{n^\mu_\xi} d^\mu_\xi \exp\left(-\alpha^\mu_\xi(\mathbf{r}-\mathbf{R}_\mu)^2\right), \tag{7}$$

where $\mathbf{R}_\mu = \{R_x, R_y, R_z\}$, n^μ_ξ is the contraction depth, d^μ_ξ is a contraction coefficient, and $L = l + m + n$ is the total angular momentum. Each term in the sum is referred to as a primitive

Gaussian function. Contracted basis functions with the same L, $\{d_\xi^\mu\}$, $\{\alpha_\xi^\mu\}$, and $\mathbf{R}_\mu$ will be referred to as a basis shell. Functions of the form Equation (7) are referred to as Cartesian Gaussian functions, and each Cartesian shell with angular momentum L consists of $L(L+1)$ functions. For $L > 1$, there is often a linear dependency among the functions within each Cartesian shell, which may be addressed by transforming these shells to a set of spherical Gaussian functions (Schlegel and Frisch, 1995). Each spherical Gaussian shell consists of $2L+1$ linearly independent functions. Not all Gaussian basis sets that consist of functions with $L > 1$ require this transformation to be linearly independent, and we will note when such a transformation has taken place.

2.2. Numerical Integration of Molecular Integrands

Even for the simplest forms of ε, neither Equation (2) nor Equation (5) admits analytic expressions, thus these integrations must be performed numerically. For molecular integrands, i.e., integrands with non-trivial behavior in the vicinity of atomic nuclei in polyatomic systems, a particularly attractive approach is to perform the numerical integration as a sum over weighted atomic integrands (Becke, 1988). For a molecular integrand $f : \mathbb{R}^3 \to \mathbb{R}$, we may decompose its integral over $\mathbb{R}^3$ as

$$\int_{\mathbb{R}^3} f(\mathbf{r})\, \mathrm{d}^3\mathbf{r} = \sum_{A=1}^{N_A} I_A[f], \qquad I_A[f] = \int_{\mathbb{R}^3} p_A(\mathbf{r}) f(\mathbf{r})\, \mathrm{d}b^3\mathbf{r}, \quad (8)$$

where N_A is the number of atoms, and $p_A : \mathbb{R}^3 \to \mathbb{R}$ is an atomic partition function that obeys $\sum_A p_A(\mathbf{r}) = 1, \forall \mathbf{r} \in \mathbb{R}^3$. Each atomic integrand $I_A[f]$ may then be approximated by a quadrature rule

$$I_A[f] \approx \sum_{i \in \mathcal{Q}_A} w_i^A f(\mathbf{r}_i^A), \quad w_i^A = p^A(\mathbf{r}_i^A) w_i^q \quad (9)$$

where $\mathcal{Q}_A = \{(w_i^A, \mathbf{r}_i^A)\}_{i=1}^{N_g^A}$ is a set of quadrature points indexed by i centered around the A-th nucleus with atomically scaled quadrature weights w_i^A. $\{w_i^q\}$ is the set of unmodified weights associated with the base quadrature around a particular nucleus. For convenience in the following, we define the total quadrature

$$\mathcal{Q} = \bigcup_A \mathcal{Q}_A = \{(w_i, \mathbf{r}_i)\}_{i=1}^{N_g},$$

where $N_g = \sum_A N_g^A$ is the total number of grid points needed to perform the numerical integration over the molecular integrand. Note that w_i is assumed to have the proper atomic scaling per Equation (9).

There are many possible choices for both the atomic partitioning scheme (Becke, 1988; Stratmann et al., 1996; Laqua et al., 2018; Aprà et al., 2020) and base quadratures around each atomic center (Becke, 1988; Murray et al., 1993; Treutler and Ahlrichs, 1995; Mura and Knowles, 1996; Gill and Chien, 2003; Aprà et al., 2020). In this work, we will use the following:

- For the atomic partition function, we will use the scheme proposed by of Stratmann, Scuseria, and Frisch (SSF) (Stratmann et al., 1996).
- For the base atomic quadrature, we will use a spherical product grid consisting of the Mura-Knowles (MK) quadrature (Mura and Knowles, 1996) for the radial integration and the Lebedev–Laikov quadrature (Lebedev, 1976) for the angular integration.

These schemes are chosen in part for the simplicity and robustness, as well as their standard use in industry KS-DFT software. Further, while it is standard practice to perform angular grid pruning to reduce the number of grid points in these product quadratures (Gill et al., 1993; Chien and Gill, 2006; Laqua et al., 2018), we perform no such procedure here. We note that the methodological details presented in this work are largely independent of such choices.

It is well-known that a naive application of Equations (8) and (9) to evaluate $\mathbf{V}^{xc}$ and $\mathcal{E}^{xc}$ is very inefficient (Stratmann et al., 1996). This is due to the fact that while Gaussian functions of the form Equation (7) do not admit compact support, their exponential character yields numerically negligible contributions when evaluated far from their center. As such, Gaussians of this form may be approximated to have compact support on a sphere centered at their $\mathbf{R}_\mu$ with cutoff radius (Burow and Sierka, 2011)

$$r_\mu^{cut} = \max_\xi \sqrt{\frac{1}{\alpha_\xi^\mu}\left(\frac{\ln \alpha_\xi^\mu}{2} - \ln \eta\right)}, \quad (10)$$

where η is a tolerance for which $|\phi_\mu| < \eta$ for all points outside of the sphere. In this work, we have chosen $\eta = 10^{-10}$. Remark that the cutoff radius only depends on the exponential coefficients, and thus may be calculated at the level of basis shell rather than individual functions for $L > 0$. Given this cutoff criteria, one may form a list of basis shells that are non-negligible for each quadrature point. Rather than check each individual quadrature points against r^{cut} for each basis shell's cutoff radius, it is canonical to group quadrature points that are spatially close into batches and perform the coarse-grained screening for non-negligible basis shells at the batch level rather than the quadrature points themselves. This procedure is known as micro-batching (Stratmann et al., 1996) and is one of the primary mechanisms by which linear scaling (with respect to system size) is achieved in the evaluation of the XC potential. Given quadrature micro-batches with a sufficiently small spatial extent, basis screening via Equation (10) produces an approximately constant number of basis functions per quadrature batch, thus leading to an overall scaling that depends only on the number of quadrature points. There are several ways to obtain the quadrature batches (Stratmann et al., 1996; Burow and Sierka, 2011; Manathunga et al., 2020). In this work, we recursively subdivide the domain spanned by the quadrature points into cuboids until the number of quadrature points within each cuboid is below a certain threshold. In this work, we have chosen this threshold to be 512 quadrature points. In practice, this partitioning scheme produces batches similar to the octree method of Manathunga et al. (2020). However, rather than bisecting every domain into octants, cuboids that contain an atomic center are partitioned into 27

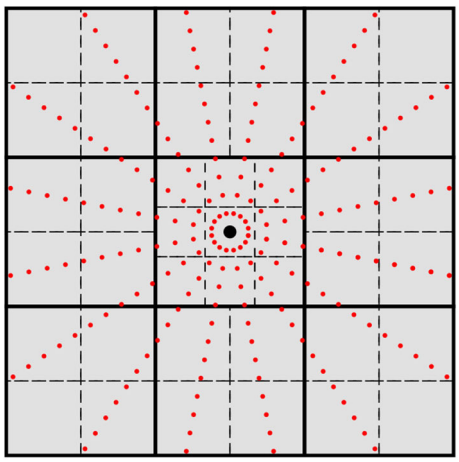

FIGURE 1 | 2-D cross-section of the grid batching scheme used in this work. The large black dot represents an atomic center and the small red dots represent quadrature points for spherical integration. Thick solid lines represent the initial cuboid partition, and dashed lines represent the next partition level. Atomic centered cuboids are partitioned into 27 cubical domains while off-center cuboids are partitioned into octants.

cuboids as shown in **Figure 1**. Our experiments show that this procedure produces fewer batches with the same non-negligible shell list, which in turn improves the performance of the load balancing scheme discussed later in this section. However, much like the choice of atomic quadrature and partition functions, the choice of batching scheme does not affect the methodological details presented in this work just as long as the batches produced are able to produce sufficiently short lists of non-negligible basis shells. For a total quadrature $\mathcal{Q}$, we denote the set of quadrature batches produced by this procedure as $\mathcal{B} = \{\mathcal{B}_j\}$ such that

$$\mathcal{Q} = \bigcup_{\mathcal{B}_j \in \mathcal{B}} \mathcal{B}_j, \quad \text{s.t.} \quad \mathcal{B}_j \cap \mathcal{B}_k = \emptyset, \text{for } j \neq k. \tag{11}$$

In the case where the batches are defined by non-overlapping cuboids surrounding an atomic center, basis shell screening may be accomplished by calculating the point of closest approach between the cuboid defining the batch and the spheres defined by center $\mathbf{R}_\mu$ and radius r_μ^{cut} (Arvo, 2013). A description of this procedure is given in Algorithm 1. For $\mathcal{B}_j \in \mathcal{B}$, we define the list of non-negligible basis functions for $\mathcal{B}_j$ as $\mathcal{S}_j$, the number of non-negligible basis functions as $N_b^j = |\mathcal{S}_j|$, and the number of quadrature points in the batch as $N_g^j = |\mathcal{B}_j|$.

Another advantage of quadrature batching is the ability to cast the evaluation of $\mathbf{V}^{xc}$ and $\mathcal{E}^{xc}$ in terms of efficient level-1 BLAS operations such as dot products (DOT) and level-3 BLAS operations such as matrix–matrix multiplication (GEMM) and symmetric rank-2K updates (SYR2K). For a particular batch $\mathcal{B}_j$, we may define a batch collocation matrix ($\mathbf{\Phi}^j$) and a local density matrix ($\mathbf{P}^j$) as

$$\Phi^j_{\mu i} = \begin{cases} \phi_\mu(\mathbf{r}_i), & \text{for } i \in \mathcal{B}_j \text{ and } \mu \in \mathcal{S}_j \\ 0, & \text{otherwise.} \end{cases} \tag{12a}$$

Algorithm 1: Basis shell screening via cuboid–sphere intersection.

Input : Sphere center $\mathbf{R}_\mu = \{R_x, R_y, R_z\}$, sphere radius r_μ^{cut}, minimum (maximum) vertex defining the cuboid $V = \{V_x, V_y, V_z\}$ ($W = \{W_x, W_y, W_z\}$).

Output: **True** if the cuboid and sphere spatially intersect, **False** otherwise.

$d \leftarrow \left(r_\mu^{cut}\right)^2$

for $p \in \{x, y, z\}$ **do**
  if $R_p < V_p$ **then** $d \leftarrow d - (R_p - V_p)^2$
  else if $R_p > W_p$ **then** $d \leftarrow d - (R_p - W_o)^2$
end

return $(d < 0)$

$$P^j_{\mu\nu} = \begin{cases} P_{\mu\nu}, & \text{for } \mu, \nu \in \mathcal{S}_j \\ 0, & \text{otherwise.} \end{cases} \tag{12b}$$

In the following, we will refer to the extent to which $\mathbf{\Phi}^j$ and $\mathbf{P}^j$ are numerically zero due to basis function screening as their local sparsity. This yields the following expressions for the density and its gradient evaluated on the quadrature points within $\mathcal{B}_j$,

$$\rho_i^j = \sum_{\mu \in \mathcal{S}_j} \Phi^j_{\mu i} X^j_{\mu i}, \quad \text{(DOT)} \tag{13}$$

$$\nabla \rho_i^j = 2 \sum_{\mu \in \mathcal{S}_j} \nabla \Phi^j_{\mu i} X^j_{\mu i}, \quad \text{(DOT)} \tag{14}$$

$$\mathbf{X}^j = \mathbf{P}^j \mathbf{\Phi}^j. \quad \text{(GEMM)} \tag{15}$$

It should be understood from the context that the free index i is restricted to quadrature points in $\mathcal{B}_j$. Given these expressions, we may now express the XC-related quantities as (Petrone et al., 2018)

$$\mathcal{E}^{xc} = \sum_{\mathcal{B}_j \in \mathcal{B}} \sum_{i \in \mathcal{B}_j} \varepsilon_i^j \rho_i^j, \quad \text{(DOT)} \tag{16}$$

$$V^{xc}_{\mu\nu} = \sum_{\mathcal{B}_j \in \mathcal{B}} V^j_{\mu\nu}, \tag{17}$$

$$\mathbf{V}^j = \mathbf{Z}^j \mathbf{\Phi}^{j,T} + \mathbf{\Phi}^j \mathbf{Z}^{j,T}, \quad \text{(SYR2K)} \tag{18}$$

with

$$\varepsilon_i^j = w_i \varepsilon(\{U(\mathbf{r}_i)\}), \quad \frac{\partial \varepsilon_i^j}{\partial \rho} = w_i \frac{\partial \varepsilon(\{U(\mathbf{r}_i)\})}{\partial \rho}$$
$$\frac{\partial \varepsilon_i^j}{\partial \gamma} = w_i \frac{\partial \varepsilon(\{U(\mathbf{r}_i)\})}{\partial \gamma}, \tag{19}$$

$$Z^j_{\mu i} = \frac{1}{2} \frac{\partial \varepsilon_i^j}{\partial \rho} \Phi^j_{\mu i} + 2 \frac{\partial \varepsilon_i^j}{\partial \gamma} \left(\nabla \rho_i^j \cdot \nabla \Phi^j_{\mu i} \right). \tag{20}$$

For brevity in the following, we define for $i \in \mathcal{B}_j$

$$\boldsymbol{\rho}^j = \left\{ \rho_i^j \right\}, \quad \nabla \boldsymbol{\rho}^j = \left\{ \nabla \rho_i^j \right\}, \quad \boldsymbol{\varepsilon}^j = \left\{ \varepsilon_i^j \right\},$$

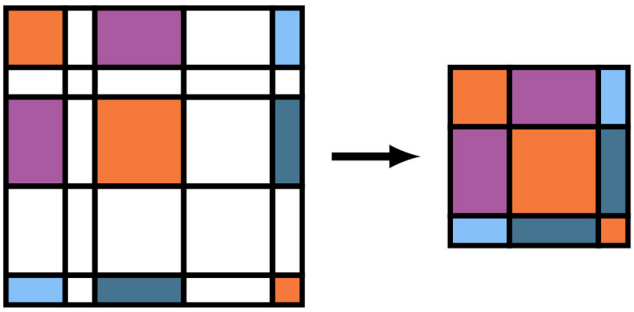

FIGURE 2 | Batch matrix compression scheme for operator basis representations relative to non-negligible function indices. Colored tiles represent matrix elements that are to be included in the compressed matrix, and white tiles represent matrix elements that are to be neglected. Note that these do not necessarily correspond to zeros/non-zeros in the original matrix.

$$\boldsymbol{\varepsilon}^j_\rho = \left\{ \frac{\partial \varepsilon^j_i}{\partial \rho} \right\}, \quad \boldsymbol{\varepsilon}^j_\gamma = \left\{ \frac{\partial \varepsilon^j_i}{\partial \gamma} \right\}. \tag{21}$$

As written, the GEMM and SYR2K given in Equations (15) and (18) are block sparse level-3 BLAS operations, i.e. BLAS operations involving matrices which contain many blocks which are numerically zero. To avoid performing unnecessary FLOPs in the evaluation of these intermediates, it is possible to store the batch local matrices in Equations (12b), (15), and (18) in a compressed format which stores the blocks corresponding to non-negligible basis shells contiguously and explicitly removes the zeros from related computation (Stratmann et al., 1996). A pictorial representation of this matrix compression for the density matrix is given in **Figure 2**. We note for completeness that the forms of Equations (15) and (18) do not change under this compression, but the sizes of the free indices (as well as the contracted index in the case of Equation 15) are reduced. To avoid a full decompression of the batched $\mathbf{V}^j$ intermediates, Equation (17) may be implemented by simply incrementing the blocks of the full dimensional $\mathbf{V}^{xc}$ by the corresponding blocks of $\mathbf{V}^j$ for each j. Note that compression of $\mathbf{\Phi}^j$, $\mathbf{X}^j$, and $\mathbf{Z}^j$ need not be explicit in that they may be evaluated directly in compressed form.

2.3. Distributed Parallel Implementation on Clusters of GPU Accelerators

In this section, we propose a three-level parallelism scheme for the distributed evaluation of $\mathbf{V}^{xc}$ and $\mathcal{E}^{xc}$. A schematic representation of this procedure is given in Algorithm 2. For simplicity in the following discussion, we will assume MPI message passing for distributed computation. Parallelism will be expressed at the following levels:

1. Concurrent evaluation of the quadrature batches between independent computing ranks;
2. Concurrent evaluation of the quadrature batches assigned to a particular computing rank;
3. Concurrency within the evaluation of a particular quadrature batch to evaluate terms such as the atomically scaled quadrature weights, batch collocation and local density matrices, the level-3 BLAS operations of Equations (15) and (18), etc.

In the context of the batching scheme discussed in section 2.2, ensuring proper local sparsity in the batch local $\mathbf{P}^j$ and $\mathbf{\Phi}^j$ typically generates a large number of relatively small batches that must be evaluated. As the work required to evaluate a single $\mathcal{B}_j$ is typically small, distributing its evaluation would be inefficient. Given that $\mathbf{P}$ and $\mathbf{V}^{xc}$ can be replicated in the memory spaces accessible to each the compute rank, the evaluation of each quadrature batch requires no communication. Thus, the fully distributed numerical integration of the XC quantities may be performed with only a single distributed reduction (`MPI_Reduce` or `MPI_Allreduce`) following completely independent local computation. We note for posterity that this replication need not constitute a unique copy of these matrices for each compute rank, only that these matrices are accessible from each rank, e.g. in the case of partitioned global address space (PGAS) distributed memory models such as the one provided by the `GlobalArrays` library, it would be sufficient to keep a single copy of these matrices within the memory accessible to a single compute node. However, in this work, we do not explore the use of PGAS memory models, thus the replication will be performed at the rank level.

2.3.1. Distributed Load Balance in the XC Integration

Despite this embarrassingly parallel integration procedure, care must be taken to ensure load balance among the independent ranks as the variance in the computational work required between different batches is often quite large due to differences in local sparsity and batch sizes. The simplest choice to distribute this work would be to distribute the batches at the atomic quadrature level, i.e. each rank receives the quadrature batches generated from a particular atomic quadrature. However, this scheme can lead to load imbalance as the local sparsity of the atoms far from the center of mass can often be much larger than those that are surrounded by other atoms. In this work, we choose to distribute the work at the individual batch level by approximating the FLOPs incurred by each batch,

$$W_j = N^j_g \left(N^2_A + 9N^j_b + 2(N^j_b)^2 + 3 \right) + (N^j_b)^2. \tag{22}$$

Each term in Equation (22) accounts for a rough estimate of the number of operations (FLOPs or otherwise) required for specific algorithmic kernels in the digestion of $\mathcal{B}_j$ for the XC integration. The first four terms accounts for (1) the atomic weight partitioning, (2) Equations 13, 14, and 20 and the collocation matrix (and its gradient), (3) the level-3 BLAS operations in Equations 15 and 18, and (4) Equations 16 and 19. The final term in Equation (22) accounts for the packing of Equation (12b) and the increment of Equation (17). Note that W_j does not represent the true number of FLOPs required to evaluate intermediates associated with $\mathcal{B}_j$, e.g., we do not

Algorithm 2: Parallelism scheme for the evaluation of the XC potential and XC energy.

Input : Density matrix $\mathbf{P}$, basis functions $\mathcal{S}$ and atomic centers $\mathcal{A} = \{\mathbf{R}_A\}$.

Output: XC potential $\mathbf{V}^{xc}$, XC energy $\mathcal{E}^{xc}$.

2.1 $\mathcal{B}_{local} \leftarrow$ Form balanced local batches according to Algorithm 3. (host)

2.2 Perform device allocation. (host/device)

2.3 Send constant data (e.g., $\mathbf{P}$, $\mathcal{S}$, and $\mathcal{A}$) to the device. (host/device)

2.4 $\mathbf{V}_{local} \leftarrow 0$; $\mathcal{E}_{local} \leftarrow 0$. (device)

do

2.5 $\quad \mathcal{B}_{device} \leftarrow$ Determine subset of $\mathcal{B}_{local}$ to saturate device memory (host)

2.6 $\quad \mathcal{B}_{local} \leftarrow \mathcal{B}_{local} \setminus \mathcal{B}_{device}$ (host)

2.7 $\quad$ Pack $\mathcal{B}_{device}$ contiguously on host and send to device. (host/device)

2.8 $\quad$ Update $\mathbf{V}_{local}$ and $\mathcal{E}_{local}$ by $\mathcal{B}_{device}$ according to Algorithm 4. (device)

while $\mathcal{B}_{local} \neq \emptyset$;

2.9 Retrieve $\mathbf{V}_{local}$ and $\mathcal{E}_{local}$ from device (host/device)

2.10 (All) reduce $\mathcal{E}^{xc} \leftarrow \mathcal{E}_{local}$ (host)

2.11 (All) reduce $\mathbf{V}^{xc} \leftarrow \mathbf{V}_{local}$ (host)

return $(\mathbf{V}^{xc}, \mathcal{E}^{xc})$

consider FLOP estimates for evaluation of the exponential in Equation (7), nor screening in the evaluation of the atomic weight scaling, etc. However, W_j has empirically sufficed to produce balanced distributed computation for all problems considered. A schematic for the load balance scheme used in this work is given in Algorithm 3. There are two important remarks that should be understood from Algorithm 3. The first is that it requires no communication between independent ranks, i.e., the load balance is replicated on each processor. The second is that once the set of local batches $\mathcal{B}_{local}$ has been determined for each processor, batches with the same $\mathcal{S}_j$ are merged into a single batch (Line 3.11). The rationale behind this step is to avoid polluting the device memory with redundant copies of $\mathbf{P}^j$ and $\mathbf{V}^j$.

While Algorithm 3 could be implemented on the GPU, as has been discussed in the context of batch generation in related work (Manathunga et al., 2020), we do not explore such implementations in this work. To improve the performance of the CPU implementation of Algorithm 3, the loop around the atomic quadrature batches may be parallelized using shared memory parallelism schemes such as `OpenMP`. Further, as has been suggested by others (Yasuda, 2008), the cost of grid generation may be amortized in calculations involving many Fock matrix formations with the same nuclear geometry by forming it once for the formation of the first Fock matrix and reusing it for subsequent formations. As will be demonstrated in section 3, Algorithm 3 only becomes a computational bottleneck in the strong scaling limit for medium-to-large molecular systems.

2.3.2. Local XC Integration on the GPU

Up to this point, the discussed work distribution scheme has been largely independent of whether or not the evaluation of local quadrature batches is to be performed on the host or the device. In this work, we only consider the case where a single MPI rank is driving a single device (one-to-one), i.e. we do not consider device affinities of multiple MPI ranks driving a single device (many-to-one) nor a single MPI rank driving multiple devices (one-to-many). The method proposed could be extended to one-to-many device affinities through an additional invocation of Algorithm 3 to produce balanced quadrature batches which are to be executed on a particular device. However, in the strong scaling limit, it would be unlikely that this affinity would be resource efficient due to a decrease in work assigned to any particular compute rank.

2.3.2.1. Architecture of NVIDIA Tesla V100

The GPU targeted in this work is the NVIDIA Tesla V100-SXM2 using the CUDA programming environment. However, the methodological developments described in this work may be extended to any GPU device given a software stack which provides batched BLAS functionality. The V100 is equipped with 16 GB high-bandwidth global memory and 80 streaming multiprocessors (SM). Within the CUDA model, independent tasks are launched in the form of kernels and concurrency on the device is expressed in a four-level parallelism scheme:

- At the lowest level is the GPU thread that executes instructions issued by the SM.

Algorithm 3: Quadrature batch load balance for distributed XC integration.

Input : Basis functions $\mathcal{S}$ and atomic centers $\mathcal{A} = \{\mathbf{R}_A\}$.

Output: Local quadrature batches $\mathcal{B}_{local}$.

3.1 $myRank \leftarrow$ Current MPI rank.

3.2 Compute $\{r_\mu^{cut}\}$ via Equation (10) for $\phi_\mu \in \mathcal{S}$.

3.3 $\mathcal{W} \leftarrow$ Allocate an array of size of the number MPI ranks.

3.4 $\mathcal{W} \leftarrow 0$; $\mathcal{B}_{local} \leftarrow \emptyset$

for $\mathbf{R}_A \in \mathcal{A}$ **do**

3.5 $\quad \mathcal{Q}_A \leftarrow$ Form spherical quadrature around $\mathbf{R}_A$.

3.6 $\quad \mathcal{B}_A \leftarrow$ Generate batches from $\mathcal{Q}_A$.

$\quad$ **for** $\mathcal{B}_j \in \mathcal{B}_A$ **do**

3.7 $\quad\quad \mathcal{S}_j \leftarrow$ Select from $\mathcal{S}$ the non-negligible basis functions via Algorithm 1 with the cuboid enclosing $\mathcal{B}_j$ and the spheres defined by $\{\mathbf{R}_\mu\}$ and $\{r_\mu^{cut}\}$.

3.8 $\quad\quad W_j \leftarrow$ Compute work estimate for $\mathcal{B}_j$ via Equation (22).

3.9 $\quad\quad I \leftarrow$ Find rank with minimum workload from $\mathcal{W}$.

3.10 $\quad\quad \mathcal{W}_I \leftarrow \mathcal{W}_I + W_j$.

$\quad\quad$ **if** $I = myRank$ **then** $\mathcal{B}_{local} \leftarrow \mathcal{B}_{local} \cup \{\mathcal{B}_j\}$.

$\quad$ **end**

end

3.11 $\mathcal{B}_{local} \leftarrow$ Merge $\mathcal{B}_j \in \mathcal{B}_{local}$ with the same $\mathcal{S}_j$.

return $\mathcal{B}_{local}$

- In contrast to CPU architectures, where all threads may execute more or less independently, the overhead of instruction issuance is mitigated on GPU devices in part by issuing a single instruction to multiple threads which execute in lock step. This is known as single-instruction multiple thread (SIMT) concurrency, and the collection of threads which execute in this manner is known as a *warp* in the CUDA vernacular. On the V100, a warp consists of 32 threads.
- Warps are then collected into groups called thread blocks, which may share data and be mutually synchronized. Thread blocks are typically comprised of 256–1024 threads which execute independently at the warp level.
- Thread blocks are further grouped into process grids which are specified at the time that the kernel is launched. A kernel has completed once all the thread blocks in its specified process grid have finished executing.

For a kernel launched with a particular process grid, thread blocks are scheduled and executed concurrently among the different SMs. Ordering of kernel execution on CUDA devices is achieved by a software construct known as a stream: kernels launched on the same stream are guaranteed to be executed in the order with which they were specified. For kernels which are designed not to achieve full occupancy within the SM, it is possible to overlap independent kernel invocations on separate streams. In this work, however, the kernels developed are designed to achieve high occupancy within each SM, thus the potential for overlap of independent kernels is minimal. Another consideration one must account for within the SIMT execution model is the concept of warp divergence, i.e. kernels that execute different instructions within a particular warp. Due to the SIMT execution model, instructions must be executed at the warp level, thus if branch logic causes the warp to diverge into N unique instructions, the execution time of this kernel will be roughly the sum of the execution times for the individual instructions, thus reducing the parallel efficiency of the particular kernel. Such divergence can lead to significant performance degradation. As such, one must carefully design GPU kernels such that unique instructions that are desired to execute concurrently are executed along (or near) warp boundaries to avoid such degradation.

2.3.2.2. Data Locality

The algorithm presented in this work aims to maximize the potential for concurrency in the evaluation of the local quadrature batches by minimizing synchronization points, such as data transfers and memory allocations, which hinder the ability to express concurrency. As the computational work required to evaluate any particular quadrature batch is small, concurrency is achieved by batching the evaluation of the quadrature batches on the GPU. This approach has been inspired by GPU accelerated batched BLAS operations, which achieve high throughput by batching the evaluation of small matrix operations into a single kernel launch (Haidar et al., 2015; Abdelfattah et al., 2016a). Given that the data associated with a particular $\mathcal{B}_j$ must reside in device memory for it to be processed (quadrature points and weights, $\mathcal{S}_j$, $\mathbf{\Phi}^j$, $\mathbf{P}^j$, $\mathbf{Z}^j$, etc.), the approach taken in this work is to saturate the device memory with as many quadrature batches as possible as to allow for their concurrent evaluation. Note that this approach does not change the amount of data that must be transferred between host and device throughout the XC integration, but it does reduce the frequency and improve the performance of these data transfers by saturating the bandwidth between host and device while allowing for the expression of more concurrency on the device between data transfers. In the case when all of the quadrature batches are unable to simultaneously occupy the device memory, subsets of the local quadrature batches which saturate device memory are chosen to be executed concurrently until all batches have been processed. A depiction of this procedure is given in Lines 2.5 to 2.8. The performance of these data transfers may be further improved in Line 2.7 by packing the batch data contiguously into page-locked memory (as is produced by `cudaMallocHost` in the CUDA SDK) on the host. In addition, rather than perform numerous memory allocations and deallocations between processing subsets of local quadrature batches, the cost of device memory allocation may be amortized by preallocating a large fraction of available device memory at the beginning of the XC integration and manually managing memory allocation throughout the calculation (Line 2.2). Note

that a vast majority of the data associated with a particular $\mathcal{B}_j$ need not be referenced on the host nor transferred between host and device. In essence, the only batch-specific data that need be transferred between host and device for a particular $\mathcal{B}_j$ are its quadrature points and weights, the information pertaining to the atomic center which generated that batch (for the evaluation of the atomic partition function), and the information describing $\mathcal{S}_j$. All other data may be allocated and manipulated directly on the device.

In addition to batch-specific data that must reside in device memory, there are a number of other quantities that are unrelated to a particular batch that are useful to store in device memory to avoid host-device transfers and to exploit the high-bandwidth memory, which is common on contemporary devices. These quantities include $\mathbf{P}$, $\mathcal{S}$, and things such as the atomic positions, inter-nuclear distances, etc. For example, in cases where $\mathbf{P}$ can reside in memory, the packing of batch local $\mathbf{P}^j$ may be made very efficient by limiting data transfers to be internal to the device memory (i.e. device memory copies). In addition, it is also advantageous to store local contributions to $\mathbf{V}^{xc}$ and $\mathcal{E}^{xc}$ on the device as to avoid communication of intermediate data between the evaluation of batch subsets on the device. We note that even for the largest problem considered in this work [1,231 atoms, $N_b = \mathrm{O}(10{,}000)$], both $\mathbf{V}^{xc}$ and $\mathbf{P}$ may reside simultaneously in device memory while leaving enough additional memory for batch-specific data as to allow for enough concurrency to be resource efficient on the device. For hypothetical problems for which this is not possible, the packing of $\mathbf{P}^j$ and the increment of $\mathbf{V}^j$ can be performed on the host at the cost of significant performance degradation. We do not explore such implementations here.

2.3.2.3. Batch Execution of Quadrature Batches on the GPU

Given a set of quadrature batches that saturate device memory, Algorithm 4 depicts a general outline of the concurrency pattern for their simultaneous evaluation on a single device. Algorithm 4 exhibits a number of important features that warrant brief discussion. The first is the utilization of batched level-3 BLAS primitives for the concurrent evaluation of Equations (15) and (18) for all batches that reside in device memory (Algorithm 4). An important remark related to this batched BLAS invocation is that the batch local matrices are often not of uniform dimension for all batches in device memory. As such, they may not be implemented by uniform batched BLAS implementations, such as those provided by `cuBLAS`. In this work, we have used the variable-dimension batched (or "vbatched") GEMM (VB-GEMM) and SYR2K (VB-SYR2K) implementations from the MAGMA (Nath et al., 2010; Tomov et al., 2010; Abdelfattah et al., 2016b) library to perform these batched evaluations. Another important feature of Algorithm 4 is that, while the order of operations within the various **parallel for** loops are indicative of the order with which the various tasks are executed at a high level, each of these tasks represent individual kernels for which concurrency between the separate $\mathcal{B}_j$'s occurs at the thread block level. That is to say that each kernel invocation performs the **parallel for** loop as a batched invocation for each task individually. As has been discussed in similar work (Laqua et al., 2020), these operations could also be scheduled on different streams to achieve concurrency in batch execution. We do not explore such implementations in this work. Finally, much like the batched BLAS invocations, which are designed to express concurrency both within a matrix operation and between matrix operations themselves, each kernel invocation for the XC-specific tasks in Algorithm 4 is designed to express concurrency within each task as well. Each batch-local task is designed to occupy a subset of the process grid while evaluation of each batch local task is performed independently on separate subsets within the same kernel launch. In practice, this may implemented using multi-dimensional kernel launches within the CUDA framework.

While GPU-accelerated BLAS functionality may be provided by optimized third-party libraries, as of this work there does not exist standard GPU implementations of the remainder of the operations required for the XC integration. As such, they

Algorithm 4: Concurrent evaluation of quadrature batches on a GPU device.

Input : Quadrature batches $\mathcal{B}$, density matrix $\mathbf{P}$, XC potential $\mathbf{V}^{xc}$, and XC energy $\mathcal{E}^{xc}$ all in device memory.

Output: $\mathbf{V}^{xc}$ and $\mathcal{E}^{xc}$ updated by quadrature contributions from $\mathcal{B}$

parallel for $\mathcal{B}_j \in \mathcal{B}$ **do**

4.1 Update quadrature weights by atomic partition function.

4.2 $\mathbf{P}^j \leftarrow$ Compress batch local density matrix from $\mathbf{P}$.

4.3 $(\mathbf{\Phi}^j, \nabla\mathbf{\Phi}^j) \leftarrow$ Evaluate compressed batch local collocation matrix and its gradient given $\mathcal{S}_j$.

end

4.4 $\{\mathbf{X}^j\} \leftarrow$ Concurrent evaluation of Equation (15) for all $\mathbf{\Phi}^j$ and $\mathbf{P}^j$ via VB-GEMM.

parallel for $\mathcal{B}_j \in \mathcal{B}$ **do**

4.5 $(\boldsymbol{\rho}^j, \nabla\boldsymbol{\rho}^j) \leftarrow$ Evaluate ρ and $\nabla\rho$ via Equations (13) and (14).

4.6 $(\boldsymbol{\epsilon}^j, \boldsymbol{\epsilon}^j_\rho, \boldsymbol{\epsilon}^j_\gamma) \leftarrow$ Evaluate XC functional and its derivatives according to Equation (19).

4.7 Update $\mathcal{E}^{xc}$ according to Equation (16).

4.8 $\mathbf{Z}^j \leftarrow$ Equation (20).

end

4.9 $\{\mathbf{V}^j\} \leftarrow$ Concurrent evaluation of Equation (18) for all $\mathbf{\Phi}^j$ and $\mathbf{Z}^j$ via VB-SYR2K.

parallel for $\mathcal{B}_j \in \mathcal{B}$ **do**

4.10 Update $\mathbf{V}^{xc}$ by $\mathbf{V}^j$ via Equation (17).

end

must be implemented and optimized by hand. The details of such implementations are outside the scope of this work as they are largely dependent on the data structures used in a particular software. However, there are a few important details related to the algorithmic choices used in this work, which warrant brief discussion. In the context of the evaluation of $\mathbf{\Phi}^j$ on the device, we adopt a simple strategy that assigns the evaluation of a single contracted basis shell at a particular point to a single thread, i.e., we do no express concurrency in the evaluation of the exponential factors of the primitive Gaussians. Care is taken in the implementation presented in this work to minimize the chance of warp divergence by assigning evaluations of the same basis shell at various quadrature points to the same warp (i.e., to minimize the frequency of divergence in the sum of Equation 7 with functions of differing n_ξ^μ). We will demonstrate the efficacy of this simple strategy in section 3.

A major difference in the work presented here relative to existing methods for GPU XC integration (Yasuda, 2008; Manathunga et al., 2020) is the strategy for the evaluation of $\boldsymbol{\epsilon}^j$ and its functional derivatives on the device. On the CPU, there are several standard libraries, such as `Libxc` (Lehtola et al., 2018) and `XCFun` (Ekström, 2020), which implement a vast number of XC functionals that are commonly used in KS-DFT calculations. Some work (Manathunga et al., 2020) has been dedicated to porting all or portions of these libraries to the GPU, including an initial implementation of porting `Libxc` to CUDA in the development version of the library itself. However, there does not exist a mature, high-performance GPU interface for these libraries at this time. To ensure the highest performance possible, the approach taken in this work has been to develop an open-source library, `ExchCXX` (Williams-Young, 2020), which provides the necessary functionality. `ExchCXX` is a modern C++ library that implements a commonly used subset of XC functionals for evaluation on the host or device though a simple, common API. We note that the numerical expressions for the XC functionals implemented in `ExchCXX` have been taken directly from `Libxc` and have been demonstrated to produce numerically indistinguishable results.

We note for posterity that, in previous work (Yasuda, 2008), the use of single precision and mixed precision arithmetic has been shown to further improve the performance of GPU-accelerated XC integration. However, as the performance gap between single and double precision arithmetic on GPU hardware has been closing in recent years (Cook, 2012), all calculations performed in this work use strictly double-precision arithmetic.

3. RESULTS

In essence, the method proposed and implemented in this work (Algorithm 2) is composed of three computationally dominant phases:

1. A load balancing phase which is replicated on all MPI ranks (Algorithm 3);
2. A local integration phase which is executed on the device (Algorithm 4);
3. A reduction phase that combines the locally computed XC quantities in distributed memory to produce the final integration results.

TABLE 1 | Molecule sizes and basis dimensions.

Molecule	N_A	N_b/6-31G(d)	N_b/cc-pVDZ
Taxol	110	1,013	1,099
Valinomycin	168	1,350	1,542
Olestra	453	3,181	3,840
Ubiquitin	1,231	10,292	11,577

TABLE 2 | Atomic quadrature sizes.

Grid	N_{ang}	N_{rad}	N_g^A
FG	302	75	22,650
UFG	590	99	58,410
SFG	974	175	170,450

In this section, we examine various performance characteristics of these phases as implemented in the open-source NWChemEx software package (Kowalski et al., 2020). In addition, we compare the performance and scaling of this implementation to that of an analogous scalable CPU implementation in the open-source NWChem software package (Aprà et al., 2020). We have chosen to examine the performance of the purposed method as applied to 4 molecules: Taxol, Valinomycin, Olestra, and Ubiquitin; and 2 basis sets: 6-31G(d) (Ditchfield et al., 1971; Hehre et al., 1972; Hariharan and Pople, 1973; Francl et al., 1982; Gordon et al., 1982) and cc-pVDZ (Dunning, 1989; Woon and Dunning, 1993), to provide a performance characterization for systems with a wide range of size, spacial extent, and basis dimension. The geometries and references for this structures are included in the **Supplementary Material**. All calculations were performed using the PBE GGA XC functional (Perdew et al., 1996). Calculations involving the 6-31G(d) basis set were performed using Cartesian Gaussian functions, while those involving cc-pVDZ were performed using spherical Gaussian functions. A list of data relevant to the performance of calculations involving these systems can be found in **Table 1**. In addition, we have examined the use of 3 commonly encountered atomic quadrature sizes: the fine (FG), ultra-fine (UFG), and super-fine (SFG) grids, as described in **Table 2**.

All calculations have been performed on the Summit supercomputer at the Oak Ridge Leadership Computing Facility (OLCF). Each Summit node consists of 2 IBM POWER9 CPUs (2x21 @ 3.8 GHz) and 6 NVIDIA Tesla V100 GPUs. Further, the Summit supercomputer leverages an NVLINK host-device interconnect that drastically improves the bandwidth of data transfers in this work. To enable a fair comparison between NWChem and NWChemEx, each Summit node has been subdivided into 6 equally sized "resource sets" consisting of 7 CPU cores and 1 GPU. For calculations involving NWChemEx, concurrency in the CPU execution will be performed in shared

memory to adhere to the one-to-one CPU-to-GPU affinity previously discussed, i.e., 1 MPI rank with 7 shared memory threads driving a single GPU. Note that CPU parallelism is only utilized in the generation of the local quadrature batches as discussed in section 2.3.1, and the launching of kernels to execute Algorithm 4 on the GPU is performed in serial.

Calculations involving NWChem were performed using a locally modified copy of release version 7.0.0. Code modifications were limited to ensuring that the radial scaling factors of the MK radial quadrature produced identical atomic quadratures to those in NWChemEx. Further, NWChem DFT calculations were performed with grid pruning disabled and using the SSF atomic partitioning scheme. Note that while the quadratures are identical between the two codes, NWChem exhibits a number of algorithmic differences with those presented in this work. These include additional density and weight screening techniques within each quadrature batch. However, these steps only improve the observed performance in NWChem, thus they do not detract from the performance comparisons made in this work. To ensure that we are comparing with consistent, replicatable performance in NWChem, all calculations have been performed using converged density matrices. Each resource set will consist of 7 MPI ranks for calculations involving NWChem as, with the exception of the atomic weight scaling, its implementation of the XC integration does not exploit shared memory parallelism. Further, we note that the use of the `GlobalArrays` library (Nieplocha et al., 2006; Krishnan et al., 2012) in NWChem yields that one MPI rank per physical node will be used as a progress rank for remote memory access rather than performing computation related to the XC integration.

Both NWChem and NWChemEx were compiled using the GNU 8.1.0 compiler suite (`gcc`, `g++`, `gfortran`) to compile host code using high levels of compiler optimization (`-O3 -mcpu=native -mtune=native -ffast-math`). The device code in NWChemEx was compiled using the NVIDIA CUDA compiler (`nvcc`) as provided in the CUDA SDK (version 10.1.105). Analogous optimization flags (`-O3 --use-fast-math`) as well as architecture specific flags to generate optimized binaries for CUDA compute capability 7.0 (`-gencode sm_70,compute_70`) were used in the compilation of device code. NWChem was linked to the serial version of the IBM Engineering Scientific Software Library (ESSL version 6.1.0) for POWER9 optimized BLAS functionality. GPU accelerated batched BLAS was provided by the MAGMA library (version 2.5.1) while non-batched BLAS for operations such as dot products was provided by the `cuBLAS` library from the NVIDIA CUDA SDK.

3.1. Integration Performance on GPU Devices

First, we examine the performance characteristics of Algorithm 2 on a single Summit node. This treatment allows us to examine the effects of molecule size, basis dimension, and quadrature size on overall GPU performance separately from scaling in a distributed setting. Strong scaling of the purposed method as well as its comparison to NWChem will be presented in the following subsection. An overall component analysis of the timings on a single Summit node is given in **Table 3**. The wall times presented in **Table 3** are aggregated over the entire XC integration, i.e. for the local integration, the times presented are representative of the sum of all invocations that saturate device memory (N_{sat}). Further, we note that these times also include the contiguous host packing and host-device transfer of batch data (i.e., all operations contained in the loop over quadrature batches in Algorithm 2). In addition, the times presented for load balancing include all operations in Algorithm 3, i.e. batch generation and the course-grained screening of basis shells at the batch level. As these calculations were performed within a single Summit node, the reduction phase is not explicitly considered in **Table 3**, but its contributions are included in the times labeled "Other." As expected, although Algorithm 3 is executed on the host in this work, the dominant computational phase for these calculations was the local integration. Further, we note that the overall cost of Algorithm 3 for a particular molecule/grid pair is largely independent of basis size but scales linearly with respect to grid size for a particular molecule/basis pair. The result of this is that the relative cost of load balancing is reduced as basis size increases. However, while this cost is not dominant at low processor counts, it will be demonstrated to be dominant in the strong scaling limit in the following subsection.

In this work, we focused on two algorithmic motifs that are important for the XC integration on the GPU:

1. Optimizing data locality to minimize the overhead of low-bandwidth data transfers between host and device and to maximize the potential to express concurrency without synchronization, and
2. Batching together the evaluation of small tasks on the device through the use of kernels that express concurrency both within a quadrature batch and between batches to improve throughput on the device.

To demonstrate the efficacy of these motifs, we examine the relative costs of the various compute and memory intensive operations incurred by the various kernels during the local integration on the device. Due to the fact that GPU computation is generally asynchronous with respect to host computation, care must be taken in accruing accurate performance data relating to individual kernels as to not impede computational progress on the device. For this purpose, we have utilized the NVIDIA profiler `nvprof` to obtain kernel level performance metrics. A summary of the overall time spent on various operations involving the GPU for the UFG basis and 6-31G(d) basis set is provided in **Figure 3**.

There are a number of important features exemplified in the results presented in **Figure 3**. The first is that saturating the device memory to ensure data locality all but removes the cost of host-to-device (H2D) and device-to-host (D2H) data transfers, yielding < 1% of the overall computational cost combined for all problems considered. For the smaller test cases (Taxol and Valinomycin), the GPU implementation is dominated by the evaluation of ρ / $\nabla\rho$ and device-to-device (D2D) memory transfers. For the larger test cases (Olestra and Ubiquitin), the integration is dominated by the evaluation of the SSF atomic

TABLE 3 | Aggregate wall times for computationally intensive operations of XC integration for the various problems considered.

Molecule	Basis	Grid	N_{sat}	Load balance	(%)	Local integration	(%)	Other	(%)	Total
Taxol	6-31G(d)	FG	1	0.073	(17.49)	0.310	(73.99)	0.036	(8.52)	0.419
		UFG	2	0.145	(15.50)	0.746	(79.59)	0.046	(4.91)	0.937
		SFG	3	0.252	(15.76)	1.30	(80.84)	0.055	(3.41)	1.60
	cc-pVDZ	FG	1	0.075	(14.62)	0.399	(77.59)	0.040	(7.79)	0.514
		UFG	2	0.153	(13.70)	0.918	(82.12)	0.047	(4.18)	1.12
		SFG	3	0.268	(13.26)	1.68	(83.24)	0.071	(3.50)	2.02
Valinomycin	6-31G(d)	FG	1	0.128	(14.74)	0.685	(79.14)	0.053	(6.12)	0.865
		UFG	3	0.259	(15.79)	1.33	(80.95)	0.054	(3.26)	1.64
		SFG	5	0.446	(14.98)	2.45	(82.21)	0.084	(2.81)	2.98
	cc-pVDZ	FG	2	0.136	(12.17)	0.916	(82.27)	0.062	(5.55)	1.11
		UFG	3	0.274	(11.99)	1.96	(85.74)	0.052	(2.27)	2.29
		SFG	6	0.474	(11.09)	3.70	(86.61)	0.098	(2.30)	4.27
Olestra	6-31G(d)	FG	2	0.433	(23.60)	1.20	(65.45)	0.201	(10.95)	1.84
		UFG	5	0.872	(23.48)	2.65	(71.39)	0.191	(5.13)	3.72
		SFG	9	1.49	(21.79)	5.14	(75.13)	0.211	(3.08)	6.84
	cc-pVDZ	FG	3	0.481	(19.87)	1.68	(69.48)	0.258	(10.66)	2.42
		UFG	6	0.953	(19.59)	3.63	(74.57)	0.284	(5.83)	4.87
		SFG	11	1.63	(18.54)	6.92	(78.53)	0.259	(2.94)	8.82
Ubiquitin	6-31G(d)	FG	22	3.12	(10.94)	22.5	(78.90)	2.89	(10.15)	28.5
		UFG	45	6.01	(10.84)	47.5	(85.70)	1.92	(3.46)	55.4
		SFG	84	10.2	(9.94)	90.2	(87.82)	2.30	(2.24)	103
	cc-pVDZ	FG	30	3.44	(7.83)	38.2	(86.96)	2.29	(5.21)	43.9
		UFG	61	6.64	(7.50)	79.6	(89.80)	2.40	(2.71)	88.6
		SFG	111	11.2	(7.04)	145	(90.90)	3.30	(2.07)	160

All times are given in seconds and N_{sat} is the number of times the device memory was saturated in Algorithm 2 to complete the integration.

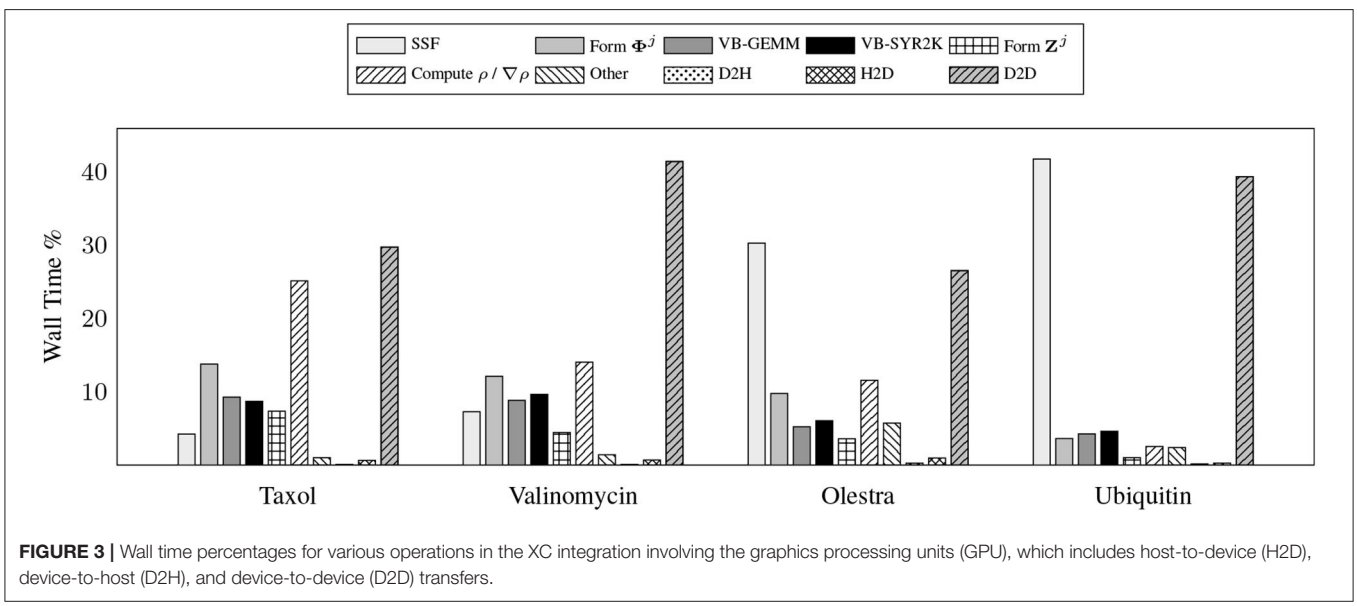

FIGURE 3 | Wall time percentages for various operations in the XC integration involving the graphics processing units (GPU), which includes host-to-device (H2D), device-to-host (D2H), and device-to-device (D2D) transfers.

partition weights and D2D memory transfers. We note for clarity that the D2D transfers are intra-GPU device memory copies, not inter-GPU communication. The times for the evaluation of the XC functional on the device are not explicitly shown in **Figure 3** as they are negligibly small. They are however included in the "Other" timing accumulations.

A somewhat unexpected result is the dominant cost posed by intra-GPU D2D transfers for all problems considered. The D2D timings including the packing of Equation (12b), the incrementing of Equation (17), and various other small D2D transfers such as those involving storage of the basis functions. This result is unexpected due to the high-bandwidth of memory transfers within device memory. To further examine the details of this unexpected dominant cost, **Figure 4** shows the achieved memory read and write throughputs for the intra-GPU data transfers incurred by the batch kernels that implement Equations (12b) and (17). These achieved throughputs are compared to the peak bandwidth of DDR4 (CPU) memory: 50 GB/s. For these kernels, we are able to achieve a memory throughput of O(100 GB/s) for data writes and between 50 and 70 GB/s for data reads, with the throughput for data reads decreasing with increasing system size. This decrease in data read throughout with system size is likely due to memory bank conflicts arising from multiple GPU threads accessing the same memory address simultaneously. Although these kernels are not able to achieve memory throughput reflective of peak device memory bandwidth (900 GB/s) due to their access of non-coalesced, non-contiguous memory, they far exceed the throughput that would be achievable in CPU memory. Further, as the memory footprint of these packed matrices are among the largest in the purposed method, exploiting intra-GPU memory transfers avoids additional H2D and D2H transfers which would pose non-trivial costs due to their low bandwidth.

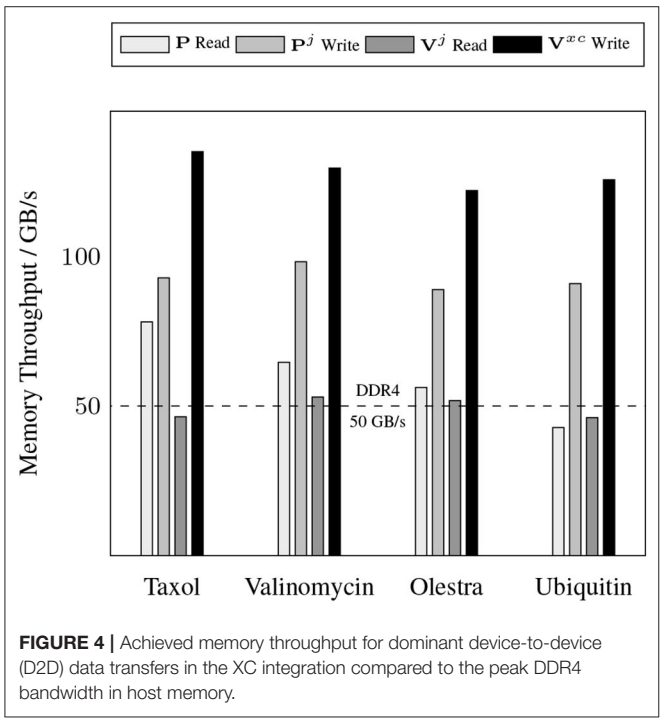

FIGURE 4 | Achieved memory throughput for dominant device-to-device (D2D) data transfers in the XC integration compared to the peak DDR4 bandwidth in host memory.

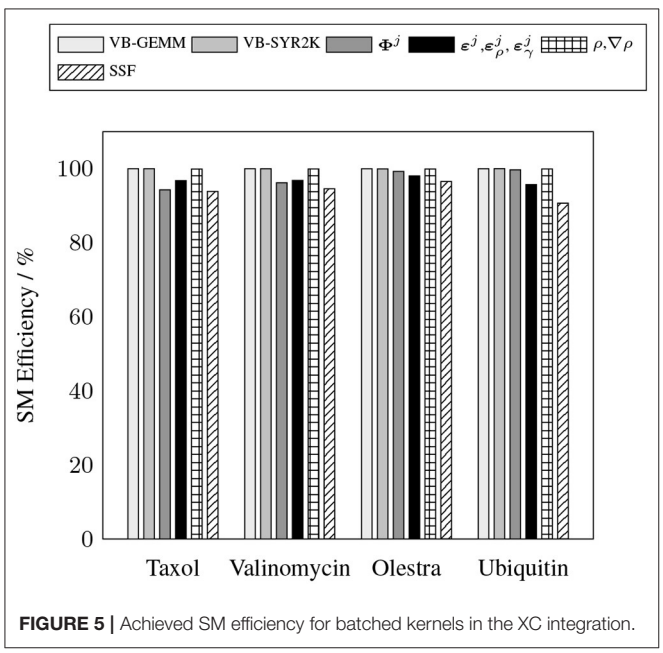

FIGURE 5 | Achieved SM efficiency for batched kernels in the XC integration.

To demonstrate the efficacy of the batched kernels proposed in this work, **Figures 5**, **6** illustrate the capability of these kernels to efficiently exploit the resources of the device. These figures present the efficiency of the batched kernels in two regimes. The SM efficiency (**Figure 5**) illustrates the efficiency of the kernels at the SM level by calculating the percentage of time each SM has at least one active warp executing instructions. The warp execution efficiency (**Figure 6**) illustrates their efficiency at the warp level by calculating the percentage of active threads within each warp in the issuance of any particular instruction in the kernel execution. Deviations from 100% in the SM efficiency indicate that the SM is sitting idle due to some sort of contention, e.g. warp divergence, while deviations in the warp execution efficiency indicate that some warps have diverged such that the SM is only able to execute instructions to some subset of the threads within these diverged warps, reducing overall parallel efficiency. These performance measurements were obtained by the `nvprof` profiler metrics `sm_efficiency` and `warp_execution_efficiency`, respectively. As we can see, both the MAGMA provided batched BLAS and the hand optimized XC integration kernels developed for this work are able to achieve high SM efficiency, i.e. the SM is occupied and issuing instructions a high percentage of the time. With the exception of the SSF weights kernel, each of the batched kernels also exhibits an excellent warp execution efficiency (>90%), which means that there are not typically a large number of warp divergences in the execution of these kernels. The relatively low (60–70%) warp execution efficiency of the SSF kernels is due to the screening of weight partitions by the SSF scheme, i.e. adjacent quadrature points often follow different branch logic in the screening procedure. Note that the high SM and warp execution efficiencies for the kernel responsible for the batched evaluation of $\mathbf{\Phi}^j$ by the simple method proposed in this

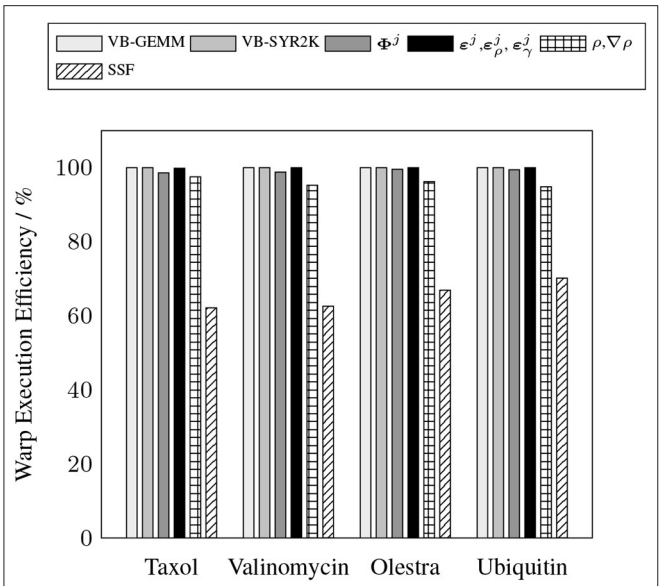

FIGURE 6 | Achieved warp execution efficiency for batched kernels in the XC integration.

work, combined with its relatively low cost percentage (>20%) for all problems considered, indicate that further optimization of this kernel by more advanced techniques would likely not yield a large impact on overall wall time.

3.2. Strong Scaling

The primary goal of this work has been to provide a scalable implementation of the XC integration. As such, we examine the strong scaling the proposed method in comparison with the CPU implementation in NWChem. Strong scaling results for the CPU and GPU XC integrations using the 6-31G(d) basis and UFG quadrature are given in **Figure 7**. The wall times presented in **Figure 7** only include those operations that are required to perform the XC integration; wall times for the allocation of device memory in the NWChemEx results have been removed. For the smallest problems (Taxol and Valinomycin), both NWChem and NWChemEx exhibit near linear strong scaling out to 4 Summit nodes (168 MPI ranks in the case of NWChem, and 24 GPUs in the case of NWChemEx). For largest problems (Olestra and Ubiquitin), linear strong scaling is exhibited up to 8 Summit notes (48 GPUs) in the case of NWChemEx and 16 nodes (336 MPI ranks) in the case of NWChem. The relative speedups of NWChemEx over NWChem for the considered systems in the 6-31G(d) basis set are given in **Figure 8**. For all but the largest problem (Ubiquitin), speedups over 10x are observed over the CPU implementation at all resource set counts. For the smallest problems with the smallest grid size (FG), speedups of ∼100x are observed when run on a small number of resource sets. The degradation in speedup as a function in quadrature size is due to the aforementioned differences in weight and density screening techniques between NWChem and NWChemEx. The magnitude of these speedups decrease as the amount of resources increase. This is especially the case for ubiquitin, where a speedup of ∼10x is observed at a single Summit node, but this speed up falls to nearly 2.5x in the strong scaling limit. To better understand the stagnation of strong scaling in this case, it is necessary to examine the scaling of the individual components of the XC integration.

Figure 9 shows the timings for various components of the GPU XC integration for considered systems. Rather than examine the scaling for each of the considered systems, we choose to profile the largest of the small sized problems (Valinomycin), and the largest problem (Ubiquitin) as representative test cases. As can be seen in **Figure 9**, the local integration scales linearly for all processor counts considered. As the local integration scales linearly, stagnation is not due to a lack of sufficient work to occupy the GPU, but rather due to the increasing cost of the MPI reduction and the constant cost of replicating Algorithm 3 on all resource sets. This scaling behavior could be further improved by porting Algorithm 3 to the GPU, however, in the case of large processor counts, the reduction becomes competitive with Algorithm 3, thus it would be unlikely to demonstrate any qualitatively different scaling behavior in this regime.

4. CONCLUSION

In this work, we have proposed and implemented a three-level, GPU-based parallelism scheme for the distributed numerical integration of the XC potential and energy required for the evaluation of the Fock matrix in the Gaussian basis discretization of KS-DFT. In addition to the development of a simple load balancing scheme, the method proposed in this work for the evaluation of local integration quantities emphasizes the use of batched kernel invocations to achieve high throughput in the evaluation of localized quadrature batches on the GPU. This approach was motivated by the recent advent of GPU-accelerated batched BLAS kernels, which have seen wide adoption in many GPU applications. We have demonstrated that the proposed load balancing scheme produces linear strong scaling in the local integration of XC quantities for the problems considered. Further we have validated the efficacy of the use of batched kernels, including the use of batched GEMM and SYR2K, by demonstrating the ability of these kernels to achieve excellent efficiency on the NVIDIA Tesla V100 for a wide range of molecular systems, basis sets, and quadrature sizes.

The largest deficiency in the current work is the restricted implementation of the GPU-related techniques to NVIDIA GPUs and the CUDA SDK. As of this work, emerging architectures are increasingly relying upon other GPU vendors (AMD, Intel, etc.), which would render direct application of the current implementation impossible. However, the principles of batched kernel evaluation may be extended to many if not all GPU devices. Thus, as has been explored in the context of related implementations of seminumerical exchange calculations (Laqua et al., 2020), future work will focus on the *portable* implementation of the scalable GPU method presented in this work.

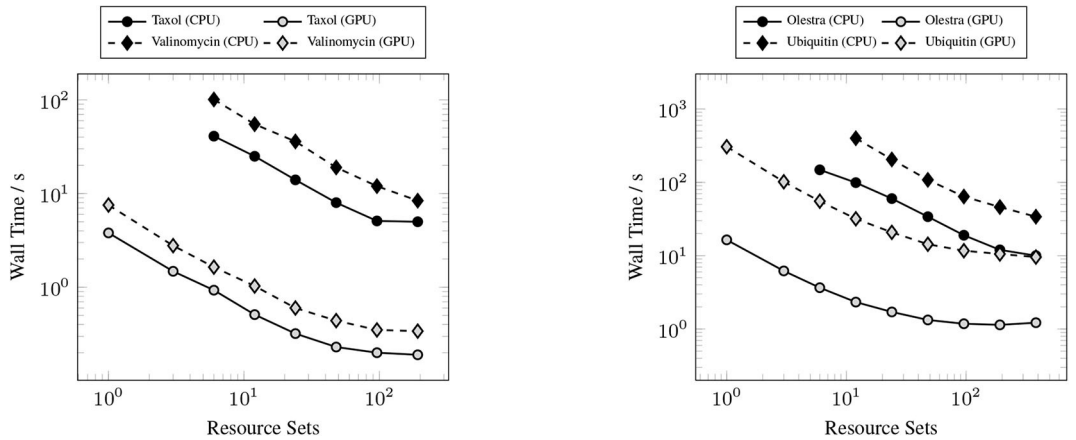

FIGURE 7 | Strong scaling comparisons for the CPU (NWChem) and GPU (NWChemEx) implementations of the XC integration. Timings for both NWChem and NWChemEx include all steps in the XC integration (batch generation, weight scaling, local integration, and reduction).

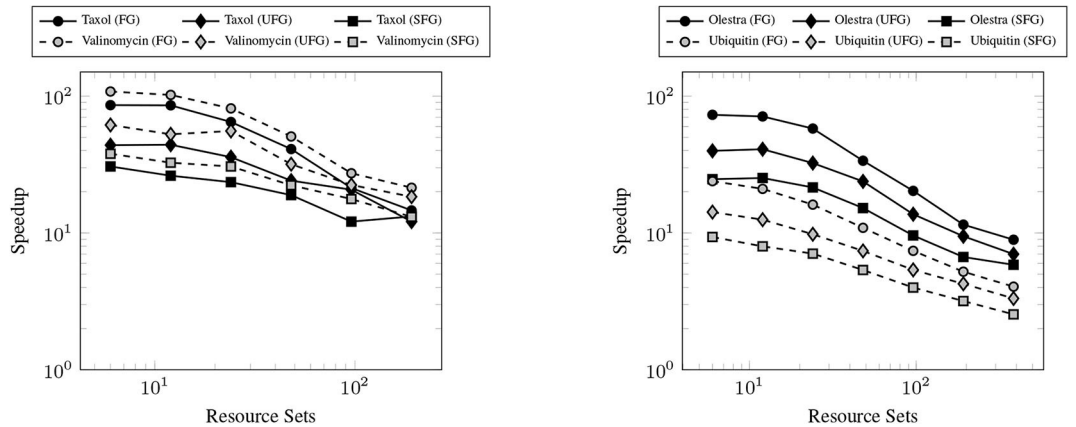

FIGURE 8 | Achieved speedups of the GPU (NWChemEx) implementation over the CPU (NWChem) implementation of the XC integration for the 6-31G(d) basis set.

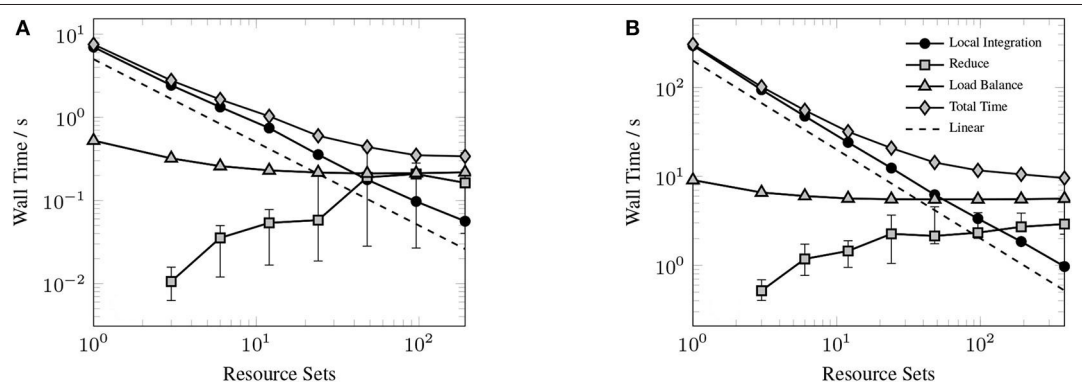

FIGURE 9 | Strong scaling of individual components of the XC integration for valinomycin **(A)** and ubiquitin **(B)** in comparison to total execution time. Error bars represent min/max times and solid markers represent average wall time over all resource sets.

We have implemented the proposed method in the open-source NWChemEx software package and have demonstrated speedups between 10x and 100x over the analogous CPU implementation in NWChem. However, in the strong scaling limit, the proposed replicated load balance scheme and distributed reduction of XC integrands become computationally dominant, which causes early stagnation relative to the linearly scaling local integration on the GPU. As has been demonstrated in related work (Manathunga et al., 2020), porting the batch generation and screening procedure to the GPU would help mitigate the strong scaling stagnation, though the asymptotic bottleneck of the distributed reduction would still remain. With the one-to-one CPU-to-GPU affinity discussed in this work, the computational cost of the MPI reduction could be reduced through the use of remote memory access (RMA) to exploit shared memory spaces and void explicit data communication. As the local integration scales linearly out to very large processor and GPU counts, further improvements in these non-GPU aspects of the XC integration would drastically improve the strong scaling of the proposed methods. Such improvements will be explored in future work.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

The software implementation of the purposed XC integration method is due to DW-Y. The algorithmic development of the load balance and distributed parallelism schemes is due to DW-Y, WJ, and CY. The development of software for the evaluation of the XC functional on the device as well as selection of the molecular test systems is due to DW-Y and HD. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by 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.

ACKNOWLEDGMENTS

The authors would also like to thank Eduardo Aprà and Ajay Panyala of the Pacific Northwest National Laboratory (PNNL) for insightful discussions regarding DFT calculations and code modifications related to the NWChem software package to enable meaningful comparisons with the methods presented in this work. Part of this research was carried out at the Brookhaven GPU Hackathon 2019. Brookhaven Hackathon is a collaboration between and used resources of Brookhaven National Laboratory, University of Delaware, and the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory. We thank Piotr Luszczek for serving as our mentor at the hackathon. This research was supported by 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. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.581058/full#supplementary-material

REFERENCES

Abdelfattah, A., Baboulin, M., Dobrev, V., Dongarra, J., Earl, C., Falcou, J., et al. (2016b). *High-Performance Tensor Contractions for GPUs*. Technical Report UT-EECS-16-738.

Abdelfattah, A., Haidar, A., Tomov, S., and Dongarra, J. (2016a). "Performance, design, and autotuning of batched GEMM for GPUs," in *High Performance Computing*, eds J. M. Kunkel, P. Balaji, and J. Dongarra (Springer International Publishing), 21–38. doi: 10.1007/978-3-319-41321-1_2

Andrade, X., and Aspuru-Guzik, A. (2013). Real-space density functional theory on graphical processing units: Computational approach and comparison to Gaussian basis set methods. *J. Chem. Theory Comput*. 9, 4360–4373. doi: 10.1021/ct400520e

Aprà, E., Bylaska, E. J., de Jong, W. A., Govind, N., Kowalski, K., Straatsma, T. P., et al. (2020). NWChem: past, present, and future. *J. Chem. Phys*. 152:184102. doi: 10.1063/5.0004997

Arvo, J. (2013). *Graphics Gems II*. Amsterdam: Elsevier.

Asadchev, A., Allada, V., Felder, J., Bode, B. M., Gordon, M. S., and Windus, T. L. (2010). Uncontracted RYS quadrature implementation of up to G functions on graphical processing units. *J. Chem. Theory Comput*. 6, 696–704. doi: 10.1021/ct9005079

Becke, A. D. (1988). A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys*. 88, 2547–2553. doi: 10.1063/1.454033

Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys*. 98, 5648–5652. doi: 10.1063/1.464913

Belling, T., Grauschopf, T., Krüger, S., Mayer, M., Nörtemann, F., Staufer, M., et al. (1999). "Quantum chemistry on parallel computers: Concepts and results of a density functional method," in *High Performance Scientific and Engineering Computing*, eds H. J. Bungartz, F. Durst, and C. Zenger (Berlin, Heidelberg: Springer), 441–455. doi: 10.1007/978-3-642-60155-2_37

Brown, P., Woods, C., McIntosh-Smith, S., and Manby, F. R. (2008). Massively multicore parallelization of Kohn-Sham theory. *J. Chem. Theory Comput*. 4, 1620–1626. doi: 10.1021/ct800261j

Brown, P., Woods, C. J., McIntosh-Smith, S., and Manby, F. R. (2010). A massively multicore parallelization of the Kohn-Sham energy gradients. *J. Comput. Chem*. 31, 2008–2013. doi: 10.1002/jcc.21485

Burow, A. M., and Sierka, M. (2011). Linear scaling hierarchical integration scheme for the exchange-correlation term in molecular and periodic systems. *J. Chem. Theory Comput*. 7, 3097–3104. doi: 10.1021/ct200412r

Bylaska, E. J., Jacquelin, M., de Jong, W. A., Hammond, J. R., and Klemm, M. (2017). Performance evaluation of nwchem *ab-initio* molecular dynamics (AIMD) simulations on the Intel® Xeon phi™ processor," in *High Performance Computing*, eds J. M. Kunkel, R. Yokota, M. Taufer, and J. Shalf (Cham: Springer International Publishing), 404–418. doi: 10.1007/978-3-319-67630-2_30

Chien, S.-H., and Gill, P. M. W. (2006). SG-0: a small standard grid for DFT quadrature on large systems. *J. Comput. Chem*. 27, 730–739. doi: 10.1002/jcc.20383

Cook, S. (2012). *CUDA Programming: A Developer's Guide to Parallel Computing with GPUs, 1st Edn*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Das, S., Motamarri, P., Gavini, V., Turcksin, B., Li, Y. W., and Leback, B. (2019). "Fast, scalable and accurate finite-element based *ab initio* calculations using mixed precision computing: 46 PFLOPS simulation of a metallic dislocation system," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '19* (New York, NY: Association for Computing Machinery). doi: 10.1145/3295500.3357157

de Jong, W. A., Bylaska, E., Govind, N., Janssen, C. L., Kowalski, K., Müller, T., et al. (2010). Utilizing high performance computing for chemistry: parallel computational chemistry. *Phys. Chem. Chem. Phys*. 12, 6896–6920. doi: 10.1039/c002859b

Ditchfield, R., Hehre, W. J., and Pople, J. A. (1971). Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys*. 54. doi: 10.1063/1.1674902

Dongarra, J. J., Duff, I. S., Sorensen, D. C., and van der Vorst, H. A. (1998). *Numerical Linear Algebra for High-Performance Computers*. Philadelphia, PA: *Society for Industrial and Applied Mathematics*. doi: 10.1137/1.9780898719611

Dunning, T. H. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys*. 90:1007. doi: 10.1063/1.456153

Egidi, F., Sun, S., Goings, J. J., Scalmani, G., Frisch, M. J., and Li, X. (2017). Two-component noncollinear time-dependent spin density functional theory for excited state calculations. *J. Chem. Theory Comput*. 13, 2591–2603. doi: 10.1021/acs.jctc.7b00104

Ekström, U. (2020). *XCFun: A Library of Exchange-Correlation Functionals With Arbitrary-Order Derivatives (Version v2.0.1)*. Zenodo. doi: 10.5281/zenodo.3796573

Fatahalian, K., Sugerman, J., and Hanrahan, P. (2004). "Understanding the efficiency of GPU algorithms for matrix-matrix multiplication," in *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS Conference on Graphics Hardware*, 133–137. doi: 10.1145/1058129.1058148

Francl, M. M., Pietro, W. J., Hehre, W. J., Binkley, J. S., Gordon, M. S., DeFrees, D. J., et al. (1982). Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys*. 77:3654. doi: 10.1063/1.444267

Genovese, L., Ospici, M., Deutsch, T., Mehaut, J.-F., Neelov, A., and Goedecker, S. (2009). Density functional theory calculation on many-cores hybrid central processing unit-graphic processing unit architectures. *J. Chem. Phys*. 131:034103. doi: 10.1063/1.3166140

Gill, P. M., Johnson, B. G., and Pople, J. A. (1993). A standard grid for density functional calculations. *Chem. Phys. Lett*. 209, 506–512. doi: 10.1016/0009-2614(93)80125-9

Gill, P. M. W., and Chien, S.-H. (2003). Radial quadrature for multiexponential integrands. *J. Comput. Chem*. 24, 732–740. doi: 10.1002/jcc.10211

Gordon, M. S., Barca, G., Leang, S. S., Poole, D., Rendell, A. P., Galvez Vallejo, J. L., et al. (2020). Novel computer architectures and quantum chemistry. *J. Phys. Chem. A* 124, 4557–4582. doi: 10.1021/acs.jpca.0c02249

Gordon, M. S., Binkley, J. S., Pople, J. A., Pietro, W. J., and Hehre, W. J. (1982). Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J. Am. Chem. Soc*. 104, 2797–2803. doi: 10.1021/ja00374a017

Haidar, A., Dong, T., Luszczek, P., Tomov, S., and Dongarra, J. (2015). Batched matrix computations on hardware accelerators based on GPUs. *Int. J. High Perform. Comput. Appl*. 29, 193–208. doi: 10.1177/1094342014567546

Hakala, S., Havu, V., Enkovaara, J., and Nieminen, R. (2013). "Parallel electronic structure calculations using multiple graphics processing units (GPUs)," in *Applied Parallel and Scientific Computing*, eds P. Manninen and P. Öster (Berlin; Heidelberg: Springer), 63-76. doi: 10.1007/978-3-642-36803-5_4

Hariharan, P. C., and Pople, J. A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* 28. doi: 10.1007/BF00533485

Hehre, W. J., Ditchfield, R., and Pople, J. A. (1972). Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys*. 56, 213–222. doi: 10.1063/1.1677527

Herault, T., Robert, Y., Bosilca, G., and Dongarra, J. (2019). "Generic matrix multiplication for multi-GPU accelerated distributed-memory platforms over parsec," in *2019 IEEE/ACM 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA)* (IEEE), 33–41. doi: 10.1109/ScalA49573.2019.00010

Hohenberg, P., and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev*. 136, B864–B871. doi: 10.1103/PhysRev.136.B864

Huhn, W. P., Lange, B., zhe Yu, V. W., Yoon, M., and Blum, V. (2020). GPU acceleration of all-electron electronic structure theory using localized numeric atom-centered basis functions. *Comput. Phys. Commun*. 254:107314. doi: 10.1016/j.cpc.2020.107314

Jacquelin, M., de Jong, W., and Bylaska, E. (2017). "Towards highly scalable *ab initio* molecular dynamics (AIMD) simulations on the Intel knights landing manycore processor," in *2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 234–243. doi: 10.1109/IPDPS.2017.26

Jia, W., Wang, L.-W., and Lin, L. (2019). "Parallel transport time-dependent density functional theory calculations with hybrid functional on summit," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '19* (New York, NY: Association for Computing Machinery). doi: 10.1145/3295500.3356144

Kalinowski, J., Wennmohs, F., and Neese, F. (2017). Arbitrary angular momentum electron repulsion integrals with graphical processing units: application to the resolution of identity Hartree-Fock method. *J. Chem. Theory Comput*. 13, 3160–3170. doi: 10.1021/acs.jctc.7b00030

Kindratenko, V. V., Enos, J. J., Shi, G., Showerman, M. T., Arnold, G. W., Stone, J. E., et al. (2009). "GPU clusters for high-performance computing," in *2009 IEEE International Conference on Cluster Computing and Workshops* (IEEE), 1–8. doi: 10.1109/CLUSTR.2009.5289128

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev*. 140, A1133–A1138. doi: 10.1103/PhysRev.140.A1133

Kowalski, K., Apra, E., Bair, R., Bertoni, C., et al. (2020). NWChemEx – computational chemistry for the exascale era. *Chem. Rev*. (Submitted).

Krishnan, M., Palmer, B., Vishnu, A., Krishnamoorthy, S., Daily, J., and Chavarria, D. (2012). *The Global Arrays User Manual*. Pacific Northwest National Laboratory, Richland, WA.

Kurzak, J., Luszczek, P., Faverge, M., and Dongarra, J. (2012). *Lu Factorization With Partial Pivoting for a Multi-CPU, Multi-GPU Shared Memory System*. Technical Report 266, LAPACK Working Note. doi: 10.2172/1173291

Kussmann, J., and Ochsenfeld, C. (2017). Employing opencl to accelerate *ab initio* calculations on graphics processing units. *J. Chem. Theory Comput*. 13, 2712–2716. doi: 10.1021/acs.jctc.7b00515

Laqua, H., Kussmann, J., and Ochsenfeld, C. (2018). An improved molecular partitioning scheme for numerical quadratures in density functional theory. *J. Chem. Phys*. 149:204111. doi: 10.1063/1.5049435

Laqua, H., Thompson, T. H., Kussmann, J., and Ochsenfeld, C. (2020). Highly efficient, linear-scaling seminumerical exact-exchange method for graphic processing units. *J. Chem. Theory Comput*. 16, 1456–1468. doi: 10.1021/acs.jctc.9b00860

Lasinski, M. E., Romero, N. A., Yau, A. D., Kedziora, G., Blaudeau, J., and Brown, S. T. (2008). "Optimization and parallelization of DFT and TDDFT in GAMESS on DoD HPC machines," in *2008 DoD HPCMP Users Group Conference*, 437–441. doi: 10.1109/DoD.HPCMP.UGC.2008.7

Lebedev, V. (1976). Quadratures on a sphere. *USSR Comput. Math. Math. Phys*. 16, 10–24. doi: 10.1016/0041-5553(76)90100-2

Lehtola, S., Steigemann, C., Oliveira, M. J., and Marques, M. A. (2018). Recent developments in libxc—A comprehensive library of functionals for density functional theory. *SoftwareX* 7, 1–5. doi: 10.1016/j.softx.2017.11.002

Luehr, N., Sisto, A., and Martinez, T. J. (2016). "Gaussian basis set Hartree-Fock, density functional theory, and beyond on GPUs," in *Electronic Structure Calculations on Graphics Processing Units: From Quantum Chemistry to Condensed Matter Physics*, eds R. C. Walker and A. W. Gotz (Hoboken, NJ: John Wiley and Sons, Ltd.), 67–100. doi: 10.1002/9781118670712.ch4

Maintz, S., Eck, B., and Dronskowski, R. (2011). Speeding up plane-wave electronic-structure calculations using graphics-processing units. *Comput. Phys. Commun*. 182, 1421–1427. doi: 10.1016/j.cpc.2011.03.010

Manathunga, M., Miao, Y., Mu, D., Götz, A. W., and Merz, K. M. (2020). Parallel implementation of density functional theory methods in the quantum interaction computational kernel program. *J. Chem. Theory Comput*. 16, 4315–4326. doi: 10.1021/acs.jctc.0c00290

Miao, Y., and Merz, K. M. (2013). Acceleration of electron repulsion integral evaluation on graphics processing units via use of recurrence relations. *J. Chem. Theory Comput*. 9, 965–976. doi: 10.1021/ct300754n

Motamarri, P., Das, S., Rudraraju, S., Ghosh, K., Davydov, D., and Gavini, V. (2020). DFT-FE—A massively parallel adaptive finite-element code for large-scale density functional theory calculations. *Comput. Phys. Commun*. 246:106853. doi: 10.1016/j.cpc.2019.07.016

Mura, M. E., and Knowles, P. J. (1996). Improved radial grids for quadrature in molecular density-functional calculations. *J. Chem. Phys*. 104, 9848–9858. doi: 10.1063/1.471749

Murray, C. W., Handy, N. C., and Laming, G. J. (1993). Quadrature schemes for integrals of density functional theory. *Mol. Phys*. 78, 997–1014. doi: 10.1080/00268979300100651

Nath, R., Tomov, S., and Dongarra, J. (2010). An improved MAGMA GEMM for fermi graphics processing units. *Int. J. High Perform. Comput. Appl*. 24, 511–515. doi: 10.1177/1094342010385729

Nguyen, T., Cicotti, P., Bylaska, E., Quinlan, D., and Baden, S. (2017). Automatic translation of MPI source into a latency-tolerant, data-driven form. *J. Parallel Distrib. Comput*. 106, 1–13. doi: 10.1016/j.jpdc.2017.02.009

Nieplocha, J., Palmer, B., Tipparaju, V., Krishnan, M., Trease, H., and Aprà, E. (2006). Advances, applications and performance of the global arrays shared memory programming toolkit. *Int. J. High Perform. Comput. Appl*. 20, 203–231. doi: 10.1177/1094342006064503

Parnell, L. A., Demetriou, D. W., Kamath, V., and Zhang, E. Y. (2019). "Trends in high performance computing: exascale systems and facilities beyond the first wave," in *2019 18th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, 167–176. doi: 10.1109/ITHERM.2019.8757229

Parr, R. G., and Yang, W. (1994). *Density Functional Theory of Atoms and Molecules*. New York, NY: Oxford University Press.

Perdew, J. P. (1986). Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* 33, 8822–8824. doi: 10.1103/PhysRevB.33.8822

Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett*. 77, 3865–3868. doi: 10.1103/PhysRevLett.77.3865

Perdew, J. P., and Yue, W. (1986). Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation. *Phys. Rev. B* 33, 8800–8802. doi: 10.1103/PhysRevB.33.8800

Peters, L. D. M., Kussmann, J., and Ochsenfeld, C. (2020). Combining graphics processing units, simplified time-dependent density functional theory, and finite-difference couplings to accelerate nonadiabatic molecular dynamics. *J. Phys. Chem. Lett*. 11, 3955–3961. doi: 10.1021/acs.jpclett.0c00320

Petrone, A., Williams-Young, D. B., Sun, S., Stetina, T. F., and Li, X. (2018). An efficient implementation of two-component relativistic density functional theory with torque-free auxiliary variables. *Eur. Phys. J. B* 91:169. doi: 10.1140/epjb/e2018-90170-1

Pople, J. A., Gill, P. M., and Johnson, B. G. (1992). Kohn-Sham density-functional theory within a finite basis set. *Chem. Phys. Lett*. 199, 557–560. doi: 10.1016/0009-2614(92)85009-Y

Ratcliff, L. E., Mohr, S., Huhs, G., Deutsch, T., Masella, M., and Genovese, L. (2017). Challenges in large scale quantum mechanical calculations. *WIREs Comput. Mol. Sci*. 7:e1290. doi: 10.1002/wcms.1290

Schlegel, H. B., and Frisch, M. J. (1995). Transformation between Cartesian and pure spherical harmonic Gaussians. *Int. J. Quant. Chem*. 54, 83–87. doi: 10.1002/qua.560540202

Stratmann, R. E., Scuseria, G. E., and Frisch, M. J. (1996). Achieving linear scaling in exchange-correlation density functional quadratures. *Chem. Phys. Lett*. 257, 213–223. doi: 10.1016/0009-2614(96)00600-8

Titov, A. V., Ufimtsev, I. S., Luehr, N., and Martinez, T. J. (2013). Generating efficient quantum chemistry codes for novel architectures. *J. Chem. Theory Comput*. 9, 213–221. doi: 10.1021/ct300321a

Tomov, S., Dongarra, J., and Baboulin, M. (2010). Towards dense linear algebra for hybrid GPU accelerated manycore systems. *Parallel Computing* 36, 232–240. doi: 10.1016/j.parco.2009.12.005

Treutler, O., and Ahlrichs, R. (1995). Efficient molecular numerical integration schemes. *J. Chem. Phys*. 102, 346–354. doi: 10.1063/1.469408

Ufimtsev, I. S., and Martinez, T. J. (2008). Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput*. 4, 222–231. doi: 10.1021/ct700268q

Ufimtsev, I. S., and Martinez, T. J. (2009a). Quantum chemistry on graphical processing units. 2. Direct self-consistent-field implementation. *J. Chem. Theory Comput*. 5, 1004–1015. doi: 10.1021/ct800526s

Ufimtsev, I. S., and Martinez, T. J. (2009b). Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput*. 5, 2619–2628. doi: 10.1021/ct9003004

van Schoot, H., and Visscher, L. (2016). "GPU acceleration for density functional theory with slater-type orbitals," in *Electronic Structure Calculations on Graphics Processing Units: From Quantum Chemistry to Condensed Matter Physics*, eds R. C. Walker and A. W. Gotz (Hoboken, NJ: John Wiley and Sons, Ltd.), 101–114. doi: 10.1002/9781118670712.ch5

Wang, L., Wu, Y., Jia, W., Gao, W., Chi, X., and Wang, L. (2011). "Large scale plane wave pseudopotential density functional theory calculations on GPU clusters," in *SC '11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, 1–10. doi: 10.1145/2063384.2063479

Williams-Young, D. (2020). *ExchCXX: A Modern C++ Library for the Evaluation of XC Functionals (Version v0.1.0)*. Zenodo. doi: 10.5281/zenodo.3929758

Woon, D. E., and Dunning, T. H. (1993). Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys*. 98:1358. doi: 10.1063/1.464303

Wu, X., Kang, F., Duan, W., and Li, J. (2019). Density functional theory calculations: a powerful tool to simulate and design high-performance energy storage and conversion materials. *Prog. Nat. Sci. Mater. Int*. 29, 247–255. doi: 10.1016/j.pnsc.2019.04.003

Yasuda, K. (2008). Accelerating density functional calculations with graphics processing unit. *J. Chem. Theory Comput*. 4, 1230–1236. doi: 10.1021/ct8001046

Yoshikawa, T., Komoto, N., Nishimura, Y., and Nakai, H. (2019). Gpu-accelerated large-scale excited-state simulation based on divide-and-conquer time-dependent density-functional tight-binding. *J. Comput. Chem*. 40, 2778–2786. doi: 10.1002/jcc.26053

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership