
















RESEARCH ARTICLE | MAY 06 2020

The ONETEP linear-scaling density functional theory program

Special Collection: [Chemical Physics Software Collection](#) , [Electronic Structure Software](#)

Joseph C. A. Prentice  ; Jolyon Aarons  ; James C. Womack  ; Alice E. A. Allen; Lampros Andrinopoulos; Lucian Anton; Robert A. Bell; Arihant Bhandari; Gabriel A. Bramley; Robert J. Charlton  ; Rebecca J. Clements  ; Daniel J. Cole; Gabriel Constantinescu; Fabiano Corsetti; Simon M.-M. Dubois; Kevin K. B. Duff; José María Escartín  ; Andrea Greco; Quintin Hill; Louis P. Lee; Edward Linscott; David D. O'Regan  ; Maximilian J. S. Phipps; Laura E. Ratcliff; Álvaro Ruiz Serrano; Edward W. Tait; Gilberto Teobaldi  ; Valerio Vitale; Nelson Yeung; Tim J. Zuehlsdorff  ; Jacek Dziedzic  ; Peter D. Haynes  ; Nicholas D. M. Hine  ; Arash A. Mostofi  ; Mike C. Payne; Chris-Kriton Skylaris  



J. Chem. Phys. 152, 174111 (2020)

<https://doi.org/10.1063/5.0004445>



CrossMark



APL Energy

Latest Articles Online!

Read Now



The ONETEP linear-scaling density functional theory program

Cite as: *J. Chem. Phys.* **152**, 174111 (2020); doi: 10.1063/5.0004445

Submitted: 12 February 2020 • Accepted: 2 April 2020 •

Published Online: 6 May 2020



View Online



Export Citation



CrossMark

Joseph C. A. Prentice,^{1,2,3,4} Jolyon Aarons,⁵ James C. Womack,^{6,7,8} Alice E. A. Allen,⁹
Lampros Andrinopoulos,² Lucian Anton,¹⁰ Robert A. Bell,^{9,11} Arihant Bhandari,^{6,8} Gabriel A. Bramley,⁶
Robert J. Charlton,^{1,3} Rebecca J. Clements,⁶ Daniel J. Cole,¹² Gabriel Constantinescu,⁹ Fabiano Corsetti,^{1,2,3}
Simon M.-M. Dubois,¹³ Kevin K. B. Duff,^{9,14} José María Escartín,^{9,15} Andrea Greco,² Quintin Hill,⁶
Louis P. Lee,⁹ Edward Linscott,^{9,16} David D. O'Regan,¹⁷ Maximillian J. S. Phipps,^{6,18} Laura E. Ratcliff,^{1,3}
Álvaro Ruiz Serrano,^{6,19} Edward W. Tait,⁹ Gilberto Teobaldi,^{6,20,21,22} Valerio Vitale,^{1,3} Nelson Yeung,⁵
Tim J. Zuehlsdorff,²³ Jacek Dziedzic,^{6,8,24,a)} Peter D. Haynes,^{1,2,3,a)} Nicholas D. M. Hine,^{5,a)}
Arash A. Mostofi,^{1,2,3,a)} Mike C. Payne,^{9,a)} and Chris-Kriton Skylaris^{6,8,a),b)}

AFFILIATIONS

¹ Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

² Department of Physics, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

³ Thomas Young Centre, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

⁴ Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom

⁵ Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

⁶ School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

⁷ School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom

⁸ The Faraday Institution, Quad One, Becquerel Avenue, Harwell Campus, Didcot OX11 0RA, United Kingdom

⁹ TCM Group, Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom

¹⁰ UKAEA, Culham Science Centre, Abingdon OX14 3DB, United Kingdom

¹¹ Mind Foundry Ltd., Oxford, United Kingdom

¹² School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom

¹³ Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

¹⁴ Enthougt, 21 J. J. Thomson Avenue, Cambridge CB3 0FA, United Kingdom

¹⁵ Institut de Química Teòrica i Computacional, Universitat de Barcelona, Calle Martí i Franquès 1, 08028 Barcelona, Spain

¹⁶ Theory and Simulation of Materials (THEOS), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

¹⁷ School of Physics, AMBER, and CRANN Institute, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland

¹⁸ Capco, 77-79 Great Eastern Street, Hackney, London EC2A 3HU, United Kingdom

¹⁹ The Boston Consulting Group, Calle Alcalá 95, Madrid 28009, Spain

²⁰ Scientific Computing Department, STFC-UKRI, Rutherford Appleton Laboratory, Didcot OX11 0QX, United Kingdom

²¹ Beijing Computational Science Research Centre, Beijing 100193, China

²² Stephenson Institute for Renewable Energy and Department of Chemistry, University of Liverpool, Liverpool L69 3BX, United Kingdom

²³ Chemistry and Chemical Biology, University of California Merced, Merced, California 95343, USA

²⁴ Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland

Note: This article is part of the JCP Special Topic on Electronic Structure Software.

^{a)} Members of the ONETEP Developers Group, who are Responsible for the Long-Term Maintenance and Sustainability of ONETEP.

^{b)} **Author to whom correspondence should be addressed:** c.skylaris@soton.ac.uk

ABSTRACT

We present an overview of the ONETEP program for linear-scaling density functional theory (DFT) calculations with large basis set (plane-wave) accuracy on parallel computers. The DFT energy is computed from the density matrix, which is constructed from spatially localized orbitals we call Non-orthogonal Generalized Wannier Functions (NGWFs), expressed in terms of periodic sinc (psinc) functions. During the calculation, both the density matrix and the NGWFs are optimized with localization constraints. By taking advantage of localization, ONETEP is able to perform calculations including thousands of atoms with computational effort, which scales linearly with the number of atoms. The code has a large and diverse range of capabilities, explored in this paper, including different boundary conditions, various exchange–correlation functionals (with and without exact exchange), finite electronic temperature methods for metallic systems, methods for strongly correlated systems, molecular dynamics, vibrational calculations, time-dependent DFT, electronic transport, core loss spectroscopy, implicit solvation, quantum mechanical (QM)/molecular mechanical and QM-in-QM embedding, density of states calculations, distributed multipole analysis, and methods for partitioning charges and interactions between fragments. Calculations with ONETEP provide unique insights into large and complex systems that require an accurate atomic-level description, ranging from biomolecular to chemical, to materials, and to physical problems, as we show with a small selection of illustrative examples. ONETEP has always aimed to be at the cutting edge of method and software developments, and it serves as a platform for developing new methods of electronic structure simulation. We therefore conclude by describing some of the challenges and directions for its future developments and applications.

© 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0004445>

I. PRINCIPLES OF LINEAR-SCALING DENSITY FUNCTIONAL THEORY

A common starting point for performing density functional theory (DFT) calculations with linear-scaling computational cost is to use the one-particle density matrix

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_i f_i \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}'), \quad (1)$$

which provides a complete description of the fictitious Kohn–Sham (KS) system, where $\{\psi_i\}$ are the KS orbitals and $\{f_i\}$ are their occupancies. KS DFT calculation methods working directly with the density matrix have been developed such as the approach by Li, Nunes, and Vanderbilt.¹ The density matrix is a quantity that decays exponentially with the distance $|\mathbf{r} - \mathbf{r}'|$ (for gapped systems at zero temperature and metals at a finite temperature), as formulated by Walter Kohn in his “nearsightedness of electronic matter” principle.^{2,3} To take practical advantage of this principle and develop methods with reduced or linear-scaling computational cost, we need to truncate the exponentially decaying tail of the density matrix. As the system size (number of atoms) increases, we will eventually reach the point where the remaining amount of information increases linearly with the size of the system. This is not straightforward to implement if we use the KS orbitals directly. A more practical approach is to work with a set of localized orbitals $\{\phi_\alpha\}$ that, in general (but not necessarily), are non-orthogonal as non-orthogonality can result in better localization.⁴ These localized orbitals can be considered to be connected to the KS orbitals via a linear transformation \mathbf{M} as follows:

$$\phi_\alpha(\mathbf{r}) = \sum_i \psi_i(\mathbf{r}) M_{i\alpha} \quad \text{and} \quad \psi_i(\mathbf{r}) = \phi_\alpha(\mathbf{r}) M_i^\alpha, \quad (2)$$

where we have used Greek indices for non-orthogonal components and Latin indices for orthogonal components. We also use the Einstein implicit summation convention for repeated greek indices. We will follow these conventions throughout this paper. It is easy to show that the density matrix can be expressed in terms of the local

orbitals in the following form:^{5,6}

$$\rho(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r}) \mathbf{K}^{\alpha\beta} \phi_\beta^*(\mathbf{r}'), \quad (3)$$

where the matrix \mathbf{K} , which we call the density kernel, is the representation of the density matrix in the basis of the localized orbitals as follows:

$$K^{\alpha\beta} = \sum_i M_{i\alpha}^\alpha f_i M_i^{*\beta}. \quad (4)$$

Linear-scaling schemes can be developed by following this framework. Calculations with such schemes need to obey the idempotency property for the density kernel

$$K^{\alpha\beta} = K^{\alpha\gamma} S_{\gamma\zeta} K^{\zeta\beta}, \quad S_{\gamma\zeta} = \langle \phi_\gamma | \phi_\zeta \rangle, \quad (5)$$

which is equivalent to the necessary requirements of KS orbital orthonormality and integer occupancies of 1 or 0 (for calculations in materials with a gap).

II. ONETEP: LINEAR-SCALING DFT WITH LARGE-BASIS SET ACCURACY

One class of linear-scaling approach typically uses a fixed basis set of localized functions, such as Gaussian atomic orbitals or numerical atomic orbitals. However, it is difficult to systematically improve the quality of such basis sets toward convergence while simultaneously maintaining linear-scaling behavior, as this requires the inclusion of a large number of extended atomic basis functions, which can make it difficult to impose the necessary localization constraints on the density matrix.

To overcome this limitation, methods have emerged, which aim to achieve linear-scaling cost with near-complete basis set accuracy or “plane wave accuracy.” The main idea in these methods is to optimize both the non-orthogonal localized orbitals $\{\phi_\alpha\}$ and the density kernel \mathbf{K} , both subject to localization constraints, so as to achieve the variational freedom of a large basis set while retaining the localization and small matrix sizes of a minimal AO basis. ONETEP⁷ belongs

to this class of methods; other codes in this category include `CONQUEST`⁸ and `BIGDFT`.⁹ In `ONETEP`, the localized orbitals $\{\phi_\alpha\}$ are called non-orthogonal generalized Wannier functions (NGWFs) and are optimized self-consistently during the calculation.¹⁰

In `ONETEP`, we aim to achieve controllable accuracy equivalent to that of plane-wave pseudopotential calculations, so the basis set we use to expand the local orbitals is a set of periodic sinc (psinc) basis functions $\{D_i\}$,^{10,11} which are unitary transformations of plane waves. The quality of the psinc set is controlled by a single parameter, the psinc kinetic energy cut-off, in analogy with plane wave codes. Thus, the psinc basis set provides a uniform description of space, similar to plane waves, and the kinetic energy cutoff controls the resolution of this description, providing a finer resolution with an increase in the cut-off value. Typical kinetic energy cutoff values in routine calculations are in the range of 500–1000 eV, although higher or lower values are possible and have been used in specialized cases.

Another advantage of using a psinc basis set is that many operations on the NGWFs that are needed for the construction of the Hamiltonian in the NGWF representation can be performed efficiently using Fast Fourier Transforms (FFTs), as in plane wave codes. This involves transforming the NGWFs to reciprocal space, applying the appropriate operator in diagonal form (e.g., the kinetic energy operator), and transforming back into real space. However, unlike in plane wave codes, these operations need to be performed with computational effort that is both relatively small and independent of the size of the system (i.e., independent of the number of atoms). To achieve this, we perform the FFTs on NGWFs in a miniature simulation cell, which we call the “FFT box.”¹² In order to preserve important properties of the integrals [such as hermiticity and uniformity of the quantum mechanical (QM) operators] computed with the FFT box, particular care must be taken in how this box is defined and used, as we have described in early publications.¹³ These publications also present the use of the FFT box for the efficient construction of the electronic density and the NGWF gradient.¹⁰

The ground state KS energy can now be considered as the minimum of a functional of the NGWFs and the density kernel

$$E[\{\phi_\alpha\}, \mathbf{K}] = K^{\alpha\beta} \left[\langle \phi_\beta | -\frac{1}{2} \nabla^2 | \phi_\alpha \rangle + \langle \phi_\beta | \hat{V}_{\text{loc}} | \phi_\alpha \rangle + \langle \phi_\beta | \hat{V}_{\text{nl}} | \phi_\alpha \rangle \right] + E_{\text{xc}}[n], \quad (6)$$

where the local potential $\hat{V}_{\text{loc}} = \hat{V}_{\text{ps,loc}} + \hat{V}_{\text{H}}$ is the sum of the Hartree potential and the local pseudopotential, \hat{V}_{nl} is the non-local part of the pseudopotential in the Kleinman and Bylander representation, and $n = n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r})$ is the electronic density.

As the NGWFs are optimized to minimize the ground state energy, it turns out that they can provide a good description of the valence bands, but a very poor description of the conduction bands. For this reason, in order to provide a set of functions suitable for describing the conduction bands (e.g., for excited state calculations), we can optimize a second set of NGWFs $\{\chi_\alpha\}$, after the completion of the ground state calculation, which provide an accurate description of a set of low-lying conduction bands (see Sec. IV E 1).

III. PARALLELIZATION WITHIN ONETEP

`ONETEP` was designed from its inception as a parallel code,¹⁴ and the extent and efficiency of its parallelism have continued to improve

during its development.^{15,16} Originally created within the paradigm of the Message Passing Interface (MPI), the code is now able to extensively harness “hybrid” parallelism enabled by OpenMP multithreading. MPI divides the available computational resources into processes, each of which can then be subdivided by OpenMP into threads, each occupying a single physical core. Each of the principal data structures and computational tasks is distributed in a manner designed to best exploit the locality on which the linear-scaling formalism relies. Data and tasks associated with atoms are distributed according to an algorithm designed to balance the competing demands of locality, aimed at ensuring nearby atoms are on the same or nearby processes to minimize communication, and load-balance, aimed at ensuring all processes have nearly equal numbers of NGWFs. This algorithm first sorts all the atoms based on a Peano space-filling curve¹⁷ and then assigns them to processes based on a modified greedy algorithm,¹⁸ looping over sets of atoms with decreasing numbers of NGWFs so that the largest are assigned first.

Data associated with each atom, particularly the coefficients of each atom-centered NGWF, are distributed and held locally on the process of that atom. The functions of each atom are then ideal for a second level of parallelism, implemented using OpenMP threading, whereby each thread treats a single function at a time, for example, when applying an operator to the set of functions. As demonstrated in Ref. 16, this leads to efficient scaling with respect to both MPI processes and threads, enabling the code to take advantage of very large-scale HPC (high-performance computing) architectures, with efficient runs on as many as 16 384 parallel cores. Subsequent work has extended the threaded parallelism to the implicit solvent model and exact exchange functionalities, detailed in Secs. IV D 1 and IV B 1. However, more optimization is always possible, and a current limitation is that MPI communications within threaded regions are handled with CRITICAL regions that are only entered by one thread

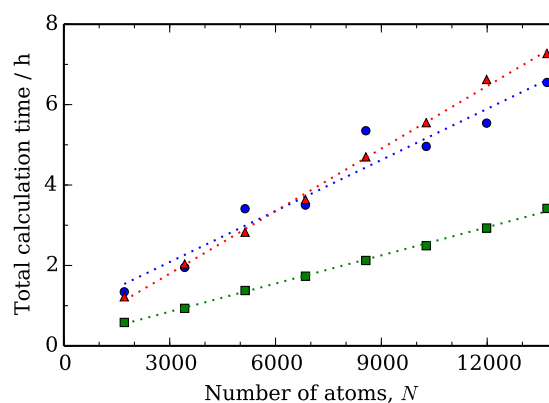


FIG. 1. Total execution time for single point energy calculations on amyloid fibril segments with N atoms ($N = 1712$ – $13\,696$). Timings are for calculations using a GGA (PBE, green squares) and two meta-GGAs (PKZB, red triangles; B97M-rV, blue circles), with dotted lines representing linear least-squares fits to the data. The calculations were run with 128 MPI processes with 4 OpenMP threads per process. Figure reprinted with permission from Womack *et al.*, *J. Chem. Phys.* **145**, 204114 (2016). Copyright 2016 AIP Publishing LLC (see the original article for further calculation details).

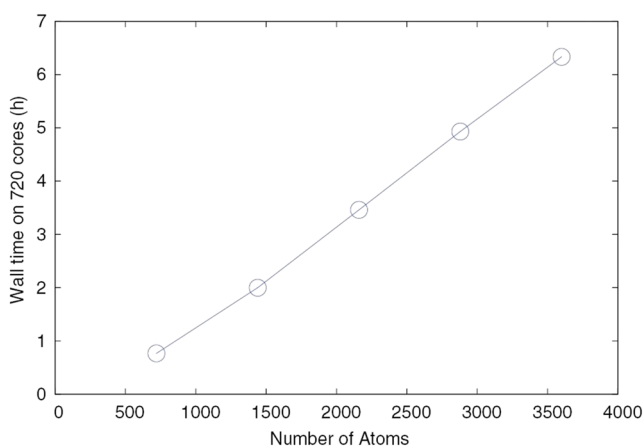


FIG. 2. Demonstration of linear scaling of total computational effort with system size for a slab of the [010] surface of the anatase-structure TiO_2 . For further details on the calculations, see Ref. 20. Figure adapted with permission (cropped) from N. D. M. Hine, *J. Phys.: Condens. Matter* **29**, 024001 (2016). Copyright 2016 IOP Publishing.

at a time, which causes whole-cell grid operations to scale relatively poorly with thread count beyond around 8–10 threads.

Figures 1 and 2 demonstrate the linear-scaling capabilities of ONETEP in two very different systems. Figure 1 presents results for amyloid fibril segments of varying sizes using several different functionals, demonstrating linear-scaling behavior at least up to 14 000 atoms.¹⁹ Figure 2 presents results for slabs of anatase-structure TiO_2 of varying sizes and shows linear-scaling behavior at least up to 3500 atoms.²⁰

IV. FEATURES AND CAPABILITIES OF ONETEP

In this section, we describe the various capabilities of ONETEP. Each feature has its own subsection, which includes a brief introduction to the functionality as well as relevant citations to other publications for those interested in either the theory or the application of the functionality. The capabilities are grouped thematically—the themes are “Improvements to core algorithms and functionality,” “Beyond-DFT methods,” “Dynamics and structure optimization,” “Including environmental effects,” “Excitations and spectroscopy,” and “Partitioning the density/energy.”

A. Improvements to core algorithms and functionality

1. Projector augmented wave methods

Blöchl’s projector augmented wave (PAW) approach enables all-electron (AE) calculations with the efficiency, systematic accuracy, and transferability of plane-wave pseudopotential calculations.²¹ PAW works by relating the soft pseudo-(PS-)wavefunctions, easily represented by plane-waves or on a grid, to AE wavefunctions, varying rapidly near the nuclei, by means of a linear transformation augmenting the soft part of the wavefunction with partial waves near each atom. PAW increases the accuracy feasible at a given computational cost, enabling efficient treatment of, for example,

transition metal oxides, lanthanides, and actinides, which would be highly challenging for norm-conserving pseudopotentials (NCPPs). PAW also allows calculation of quantities requiring knowledge of wavefunctions, densities, and electric fields in the regions near to the atomic nuclei, which are not correctly represented by norm-conserving methods. Hine²⁰ introduced a novel adaptation of the PAW formalism suited to linear-scaling DFT within ONETEP. The nomenclature here follows that of Ref. 22, which also contains an introduction to the PAW method as applied to traditional DFT.

The fundamental PAW transformation relates AE orbitals $|\psi_n\rangle$ to PS orbitals $|\tilde{\psi}_n\rangle$,

$$|\psi_n\rangle = |\tilde{\psi}_n\rangle + (|\varphi_v\rangle - |\tilde{\varphi}_v\rangle)\langle\tilde{P}^v|\tilde{\psi}_n\rangle. \quad (7)$$

$|\tilde{\psi}_n\rangle$ are the projectors, and $|\varphi_v\rangle$ and $|\tilde{\varphi}_v\rangle$ are the AE and PS partial waves (see Refs. 22 and 20), respectively.

Within the linear-scaling DFT approach, we do not have direct access to the eigenstates, so we can instead use an equivalent transformation to relate the AE density matrix to the PS density matrix,

$$\rho = \tilde{\rho} + \sum_{\nu\mu} (|\varphi_\nu\rangle\langle\varphi_\mu| - |\tilde{\varphi}_\nu\rangle\langle\tilde{\varphi}_\mu|)\langle\tilde{P}^\nu|\tilde{\rho}|\tilde{P}^\mu\rangle. \quad (8)$$

The NGWFs in ONETEP, $\{\phi_\alpha(\mathbf{r})\}$, which are constructed out of psinc functions and thus equivalent to the plane waves of standard PAW, are an ideal means to represent this “soft” part of the density matrix (DM), so one can define

$$\tilde{\rho} = |\phi_\alpha\rangle K^{\alpha\beta} \langle\phi_\beta|. \quad (9)$$

Notably, this allows one to retain the same basic forms of almost all the optimization algorithms for the kernel and NGWFs, having first redefined the overlap matrix between NGWFs to account for the PAW overlap operator,

$$S_{\alpha\beta} = \langle\phi_\alpha|[1 + \tilde{P}^\dagger](\langle\varphi_\nu|\varphi_\mu\rangle - \langle\tilde{\varphi}_\nu|\tilde{\varphi}_\mu\rangle)\langle\tilde{P}^\mu|\phi_\beta\rangle. \quad (10)$$

Further complications associated with the representation of the compensation density, non-local potential energy terms, NGWF gradient, forces, and preconditioning must also be dealt with: these are discussed in detail in Ref. 20. The result is a method that has demonstrably superior convergence properties compared to NCPPs (see Fig. 2 of Ref. 20) and thus can achieve greater accuracy for a given computational cost. At fixed parameters such as NGWF radius and cutoff energy, the overhead of a PAW calculation is relatively low, at around 10%–20% compared to NCPPs. Linear-scaling computational cost remains just as achievable as for NCPPs, since the majority of the terms in PAW appear naturally on a per-atom basis. Reference 20 also shows excellent parallel scaling, demonstrating efficient simulation of 1824 atoms on 1920 cores.

2. Ensemble density functional theory

Ensemble density functional theory (EDFT)^{23–25} calculations are available in ONETEP. Within this formalism, the KS states have fractional occupancies, which can be determined in several different ways, including the Methfessel–Paxton,²⁶ Gaussian smearing,^{27,28} and Fermi–Dirac²⁴ schemes. In ONETEP, the occupancies are determined by the Fermi–Dirac distribution, regulated by a constant electronic temperature. The macroscopic state function is the Helmholtz free energy, which obeys the variational principle of

quantum mechanics. ONETEP directly minimizes the Helmholtz free energy functional to self-consistently find the KS states and their fractional occupancies.²⁹

The ONETEP approach requires a Hamiltonian diagonalization step in each SCF (self-consistent field) iteration, which makes the overall cost scale with the cube of the system size. However, the cost of the Hamiltonian diagonalization is kept to a minimum as a consequence of using a minimal set of NGWFs. In this way, the dimension of the Hamiltonian matrix is reduced to the minimum possible, and the prefactor due to diagonalization is reduced significantly. The remaining steps of the method can be performed at a linear-scaling cost due to the sparsity of the matrices required in the NGWF representation.

Additionally, parallel diagonalization techniques can be used to scale the diagonalization of the Hamiltonian to many computational cores and to distribute the associated memory requirements. As a result, simulations on metallic systems with thousands of atoms can be achieved using the ONETEP EDFT implementation.

In free-spin EDFT calculations, the charge and net spin of a system do not necessarily have to be integers, and the net spin can additionally be relaxed during a calculation. This makes the exploration of systems with unknown magnetic properties much easier.

3. Fermi operator expansion

As in linear-scaling methods for insulators, the idea behind the Fermi operator expansion (FOE)^{30,31} is to be able to perform an SCF calculation without using a cubic-scaling diagonalization step. FOE, however, also works for metallic systems where we want to apply fractional occupancies to the KS states.

In FOE, the density matrix, with finite-temperature electronic occupancies, is constructed as a polynomial expansion of the matrix form of the Fermi–Dirac occupancy formula,

$$f(\mathbf{H}, \mathbf{S}, \mu, \beta) = \left(\mathbf{S} + e^{(\mathbf{H}\mathbf{S}^{-1} - \mu\mathbf{I})\beta} \mathbf{S} \right)^{-1}, \quad (11)$$

where \mathbf{I} is the identity matrix, \mathbf{H} is the Hamiltonian matrix, \mathbf{S} is the overlap matrix, μ is the chemical potential, and $\beta = 1/k_B T$. Equation (11) cannot be applied directly as the condition number of the matrix to be inverted will be too large in general. Many options for expansions have been proposed over the last three decades.^{32–38} In ONETEP, we have implemented the AQuA-FOE method described in detail in Ref. 39. AQuA-FOE takes advantage of the trigonometric mapping between the Fermi–Dirac occupancy formula and the hyperbolic tangent function. In this way, we can construct the density kernel at an electronic temperature many times that of the target temperature. The target temperature density kernel can then be recovered by repeatedly halving the temperature of the density kernel with the matrix analog of the hyperbolic tangent double angle formula. A fixed expansion length is always employed at the high temperature to generate the density matrix, requiring fewer terms to reach a given accuracy than at lower temperature.

To build the high temperature density kernel, we use a Chebyshev expansion of Eq. (11) using the fast re-summing algorithm of Liang *et al.*³³ As an alternative option, this algorithm can be used as the sole FOE method in ONETEP without the annealing and quenching steps.

The AQuA-FOE method in ONETEP can also find the chemical potential that conserves electron number by using a safeguarded

Newton's algorithm and the analytic derivative of the number of electrons with respect to the chemical potential. The chemical potential of the density kernel is then updated using the addition theorem of the hyperbolic tangent function so that the expansion need not be recalculated from scratch.

The AQuA-FOE method is presently invoked within the EDFT method (Sec. IV A 2) in ONETEP in order to produce finite temperature density kernels for each trial Hamiltonian matrix. In the future, this method could be used within density-mixing type or direct minimization algorithms.

4. Hybrid NGWFs

ONETEP exploits the nearsightedness of the electronic structure (Sec. I) by using NGWFs that are strictly localized within spherical regions and which give rise to sparse matrix representations that may be truncated in real space in a systematically controllable manner and enable scalability of the method to very large systems. There are cases, however, in which it is advantageous to use a hybrid approach, in which the representation of electronic structure is localized along certain directions (Wannier-like) and delocalized along others (Bloch-like). Use cases of such a hybrid approach include layered systems and surface slabs, in which localization is applied only in the direction normal to the layer or slab, and nanowires, in which localization is applied in the two directions normal to the nanowire axis.

The concept of such so-called “hybrid Wannier functions” was first introduced by Sgiarovello *et al.*⁴⁰ This idea, within the formalism of maximally localized Wannier functions,⁴¹ was subsequently used, for example, to study the dielectric properties of layered materials^{42,43} and topological insulators.^{44,45} In these applications, the hybrid Wannier functions were obtained by post-processing the Bloch eigenstates from a traditional DFT calculation.^{46,47} In ONETEP, we have introduced an approach in which hybrid NGWFs are optimized directly *in situ*, avoiding the system-size bottleneck associated with the $O(N^3)$ scaling of the traditional DFT calculation. We have also included the flexibility for the user to choose the lattice vector directions that are represented by a localized description and those that are extended, providing the full range of options in going from fully localized NGWFs to hybrid NGWFs that are localized in either two directions or one direction to a fully extended Bloch-like representation (see Fig. 3). Just as in the case of fully localized NGWFs, the hybrid NGWFs are constrained to be strictly zero outside of their localization regions.

Figure 4 shows the total energy per atom, calculated with hybrid NGWFs that are localized along the [001] direction (“2D-HNGWFs”) with different localization lengths r_{cut} , for a $2 \times 2 \times 10$ slab supercell of an eight-atom cubic diamond carbon conventional unit cell. The slab is terminated with hydrogen atoms on its {001} surfaces, and periodic replicas in the [001] direction are separated by 15 Å vacuum regions. We use a kinetic energy cutoff of 1200 eV for the psinc basis, norm-conserving pseudopotentials for the C and H atoms, the local-density approximation (LDA) for exchange and correlation, and the Brillouin zone is sampled at the Γ -point only. The carbon diamond lattice parameter is set to its equilibrium bulk value within LDA-DFT, which we calculate to be 3.539 Å. The energies shown in Fig. 4 are referenced to an equivalent calculation with the traditional plane-wave code CASTEP.⁴⁹ As can be seen, with hybrid NGWFs, it is possible to attain equivalence with

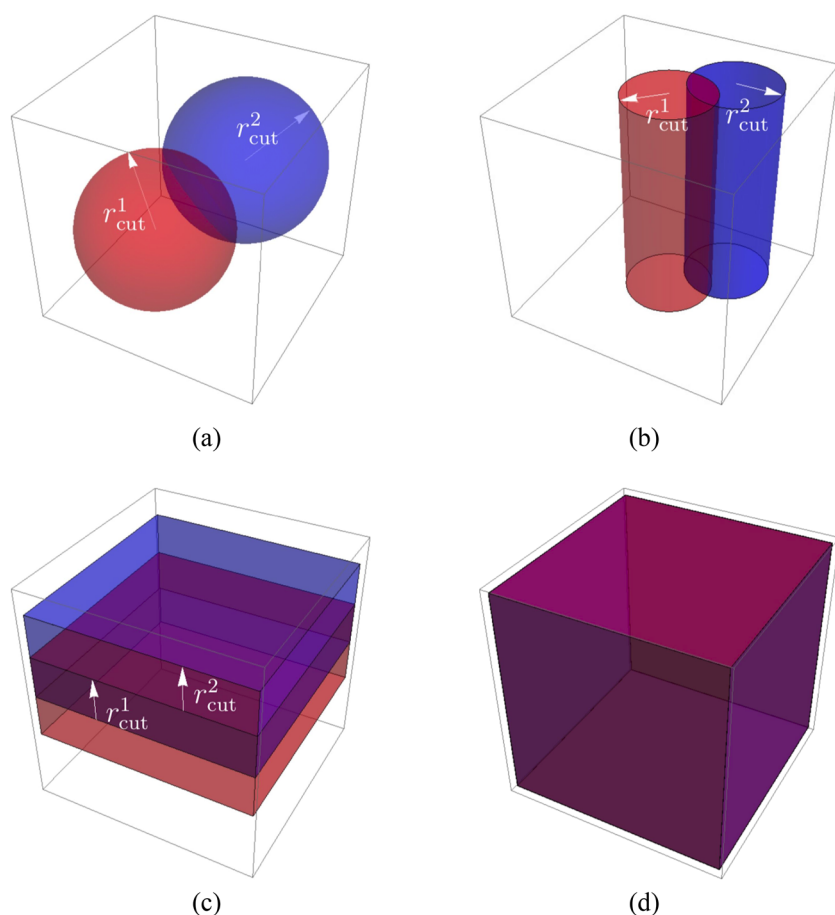


FIG. 3. Different representations of the electronic structure: (a) fully localized NGWFs; [(b) and (c)] hybrid NGWFs that are localized in two directions and one direction, respectively, and extended in the other direction(s); and (d) fully extended Bloch-like functions. Two functions are shown in each case (red and blue). A single parameter r_{cut}^i , for each function i , controls the localization: it represents the radius of the sphere in case (a), the radius of the cylinder in case (b), and the half-width of the localization length in case (c). Figures adapted with permission from Fig. 4.2 of A. Greco, “Development and application of first-principles methods for complex oxide surfaces and interfaces,” Ph.D. thesis, Imperial College London, 2017. Copyright 2017 Author(s), licensed under a Creative Commons Attribution 3.0 Unported License.

a traditional plane-wave approach using a localization length of around 5 Å.

While the relaxation of the localization constraints results in a less favorable scaling of computational effort with system-size than using fully localized orbitals, the computational cost still only scales linearly with the size of the system along the directions in which the localization constraints are still applied (e.g., the layer/slab normal in a layered/slab system). Furthermore, in our experience, there are additional advantages that offset the nominal increase in computational cost per self-consistent iteration: the use of hybrid NGWFs has a tendency to improve the conditioning of the energy minimization, resulting in faster convergence; and the accuracy of the ionic forces tends to be better with a reduction in the “egg-box” effect^{50–54} that is usually present in real-space local-orbital methods. For further details on these issues, refer to Ref. 48.

5. Poisson equation multigrid solver (DL_MG)

DL_MG⁵⁵ is a library that solves variants of the Poisson equation with the general form

$$\nabla \cdot (\epsilon(\mathbf{r}) \nabla v(\mathbf{r})) = \alpha n_{\text{tot}}(\mathbf{r}) + \gamma \sum_i \tilde{c}_i q_i \lambda(\mathbf{r}) \exp(-\beta z_i v(\mathbf{r})) \quad (12)$$

in 3D over a cuboid domain with periodic, Dirichlet or mixed boundary conditions. A regular discretization grid is used over which the derivatives are approximated with finite differences of adjustable order. The library is developed in close coordination with ONETEP and is distributed as part of the package.

In Eq. (12), $\epsilon(\mathbf{r})$ is the relative permittivity, $v(\mathbf{r})$ is the electrostatic potential (ESP), and $n_{\text{tot}}(\mathbf{r})$ is the fixed charge density. These terms feature in the generalized Poisson equation (GPE)

$$\nabla \cdot (\epsilon(\mathbf{r}) \nabla v(\mathbf{r})) = \alpha n_{\text{tot}}(\mathbf{r}). \quad (13)$$

The second term on the right-hand side of Eq. (12) appears in the Poisson–Boltzmann equation (P–BE), representing the charge density of mobile ions (e.g., electrolyte). In this term, z_i and \tilde{c}_i are the electric charge and the average bulk concentration (N_i/V) for the mobile ion of type i , respectively, $\beta = 1/k_B T$ is the inverse temperature, and $0 \leq \lambda(\mathbf{r}) \leq 1$ is the ion accessibility function that accounts for the short range repulsion effect between mobile ions and fixed ionic cores. α and γ are constants that depend on the used unit systems.

In addition to solving the non-linear Poisson–Boltzmann equation [NLP–BE, Eq. (12)], DL_MG employs specialized algorithms to solve both the GPE [Eq. (13)] and linearized Poisson–Boltzmann

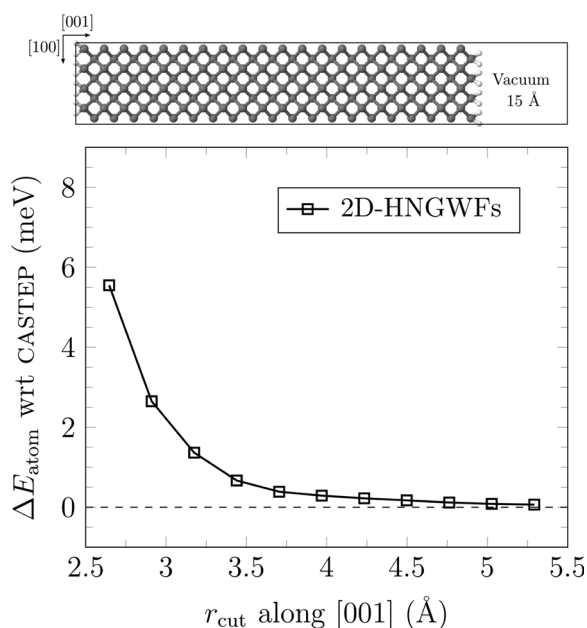


FIG. 4. Total energy per atom, referenced to CASTEP, for a $2 \times 2 \times 10$ H-terminated diamond carbon slab, obtained with hybrid NGWFs localized along [001], as a function of the hybrid NGWF localization length r_{cut} . Figure adapted with permission from Fig. 4.2 of A. Greco, "Development and application of first-principles methods for complex oxide surfaces and interfaces," Ph.D. thesis, Imperial College London, 2017. Copyright 2017 Author(s), licensed under a Creative Commons Attribution 3.0 Unported License.

equation (LP-BE),

$$\nabla \cdot (\epsilon(\mathbf{r}) \nabla v(\mathbf{r})) + \gamma \beta \sum_i \tilde{c}_i z_i^2 \lambda(\mathbf{r}) v(\mathbf{r}) = \alpha n_{\text{tot}}(\mathbf{r}). \quad (14)$$

In brief, DL_{MG} finds the solution for the GPE or P-BE in two stages:

1. A solution is found for the second order finite difference discretization using a multigrid algorithm for the GPE/LP-BE or a combination of the inexact Newton method and multigrid for the NLP-BE.
2. The solution corresponding to the chosen finite difference order (if larger than two) is found with a defect correction procedure that starts from the second order solution. The largest finite difference order implemented is 12.

See Ref. 55 for a detailed description of the theory and methods employed in DL_{MG} .

DL_{MG} underpins several functionalities of ONETEP , including implicit solvent (Sec. IV D 1) and open boundary conditions (Sec. IV D 4), which utilize the flexible Poisson-solving capabilities of the library.

B. Beyond-DFT methods

1. Orbital-dependent exchange-correlation functionals

On the well-known Jacob's ladder of density functional approximations^{56,57} (Fig. 5), orbital-dependent functionals occupy the

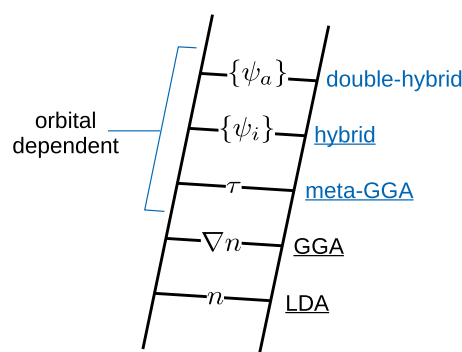


FIG. 5. The Jacob's ladder of density functional approximations.^{56,57} Families of exchange–correlation functionals currently supported in ONETEP are underlined, and the orbital-dependent families are labeled in blue. Figure reprinted with permission from Womack *et al.*, J. Chem. Phys. **145**, 204114 (2016). Copyright 2016 AIP Publishing LLC.

higher rungs above the local density approximation (LDA) and the generalized gradient approximation (GGA). Dependence on the form of the KS orbitals is added on top of the usual dependence on density n in LDA and density gradient ∇n in GGA. This increases the flexibility available in the functional forms, enabling the satisfaction of additional physical constraints and/or better fitting of empirical data. The upshot of this is that improved accuracy is possible, though not guaranteed.

Two families of orbital-dependent functionals are currently supported in ONETEP : meta-GGAs and hybrids, which depend on the KS orbitals via the kinetic energy density, τ , and the Hartree–Fock exchange energy, E_{x}^{HF} , respectively. The primary challenge of supporting both families in ONETEP is maintaining the ONETEP formalism's $\mathcal{O}(N)$ scaling when evaluating quantities with dependence on the KS orbitals.

a. Meta-GGA functionals. The orbital dependence of the meta-GGAs is via τ , i.e.,

$$E_{\text{xc}}^{\text{MGGGA}}[n] = \int d\mathbf{r} \epsilon_{\text{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r}), \tau(\mathbf{r})) \quad (15)$$

with

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_i^{N_{\text{occ}}} |\nabla \psi_i(\mathbf{r})|^2, \quad (16)$$

where the summation is over the occupied KS orbitals. In ONETEP , τ is evaluated in terms of the NGWFs and density kernel,

$$\tau(\mathbf{r}) = \frac{1}{2} (\nabla \phi_{\alpha}(\mathbf{r})) \cdot (K^{\alpha\beta} \nabla \phi_{\beta}(\mathbf{r})). \quad (17)$$

This quantity has the same structure as the charge density n , but with $\nabla \phi_{\alpha}$ in place of ϕ_{α} , and can be evaluated using the same procedure,¹⁴ subject to the subtle approximation that extra overlapping pairs of NGWFs $\alpha\beta$ that would result from the delocalization of $\{\phi_{\alpha}\}$ upon application of the gradient operator are neglected.¹⁹

Once τ is constructed, evaluation of $E_{\text{xc}}^{\text{MGGGA}}$ is simple, since the semi-local exchange–correlation energy density $\epsilon_{\text{xc}}(\mathbf{r})$ depends only on n , ∇n , and τ at point \mathbf{r} . The difficulty arises in evaluating the exchange–correlation potential,

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}, \quad (18)$$

and the quantities that depend on this, such as the gradient of $E[\{\phi_\alpha\}, \mathbf{K}]$ [Eq. (6)] with respect to \mathbf{K} and the NGWFs. Since E_{xc}^{MGGA} is a functional of τ , but the dependence of τ on n is unknown, evaluation of the functional derivative is not straightforward. ONETEP's implementation of meta-GGAs sidesteps this issue by employing the "functional derivatives of τ -dependent functionals with respect to the orbitals" (FDO) method,^{58,59} which yields a generalized Kohn-Sham approach⁶⁰ with an XC potential comprised of a local GGA-like part and a non-multiplicative orbital-specific τ -dependent part. This can be expressed in a "per-NGWF" form

$$\hat{V}_{xc}^{FDO} \phi_\alpha(\mathbf{r}) = V_{xc}^{GGA} \phi_\alpha(\mathbf{r}) + \{\hat{V}_{xc}^\tau \phi_\alpha(\mathbf{r})\}, \quad (19)$$

which can be easily inserted wherever the product of the XC potential and an NGWF arises (e.g., in Hamiltonian matrix elements and energy gradient expressions).

In ONETEP v5.2, two meta-GGA XC functionals are available: PKZB^{61,62} and B97M-rV.^{63,64} The former is an early meta-GGA form (1999), implemented primarily for testing and validation. The latter is a modern and very accurate⁶⁵ semi-empirical functional developed through a combinatorial approach, where the τ -dependent functional form itself was obtained by fitting to empirical data. B97M-rV additionally accounts for dispersion effects via a contribution from the (revised) VV10 non-local correlation functional^{66,67} (see also Sec. IV B 2).

Energies computed with meta-GGA XC in ONETEP using a minimal basis of NGWFs show excellent agreement with large basis set calculations in QCHEM, and the cost of these calculations scales linearly with system size, N .¹⁹ Figure 1 shows that single point energy calculations on amyloid fibril segments up to 13 696 atoms are only a factor of two to three times more costly for meta-GGAs compared to Perdew–Burke–Ernzerhof (PBE) (a GGA functional). Consequently, even very large systems ($\sim 10^4$ atoms) are accessible with meta-GGA XC in ONETEP.

b. Hybrid functionals. Hybrid functionals include some fraction of the Hartree–Fock exchange (HFx) energy, E_x^{HF} . For example, the widely used Becke three-parameter hybrids^{68,69} (e.g., B3LYP⁷⁰) have the form

$$E_{xc}^{B3hyb} = E_{xc}^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{GGA}, \quad (20)$$

mixing gradient-corrected exchange and correlation (from E_x^{LDA} , ΔE_x^{B88} , E_c^{LDA} , and ΔE_c^{GGA}) and HFx in fractions determined by the following three fitted parameters: a_0 , a_x , and a_c .

Whereas $\tau(\mathbf{r})$ is a local quantity, depending on the orbitals at \mathbf{r} , HFx is non-local, i.e.,

$$E_x^{HF} = \int d\mathbf{r} \epsilon_x^{HF}[\{\psi_i\}](\mathbf{r}), \quad (21)$$

with an exchange energy density that depends on the form of the orbitals over all space,

$$\epsilon_x^{HF}[\{\psi_i\}](\mathbf{r}) = - \sum_{i,j=1}^{N_{occ}} \psi_i(\mathbf{r})\psi_j(\mathbf{r}) \int d\mathbf{r}' \frac{\psi_i(\mathbf{r}')\psi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (22)$$

The non-locality of HFx presents a challenge to ONETEP's linear-scaling computational framework, which is premised on the

"nearsightedness" of electron–electron interactions (Sec. I). Nevertheless, ONETEP is able to perform hybrid functional DFT with $\mathcal{O}(N)$ scaling computational cost.⁷¹ This is achieved through a combination of a distance-based exchange interaction cutoff and the use of an auxiliary basis of atom-centered truncated spherical waves (SWs).

The HFx energy can be expressed in terms of the NGWFs and density kernel, i.e.,

$$E_x^{HF} = -K^{\beta\alpha} X_{\alpha\beta}, \quad (23)$$

with exchange matrix elements

$$X_{\alpha\beta} = (\phi_\alpha \phi_\delta | \phi_\gamma \phi_\beta) K^{\delta\gamma}, \quad (24)$$

where, for the sake of brevity, we have introduced the notation

$$(g|h) = \iint g^*(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} h(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (25)$$

$$(ab|cd) = \iint a^*(\mathbf{r}) b(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} c^*(\mathbf{r}') d(\mathbf{r}') d\mathbf{r} d\mathbf{r}'.$$

The distance-based cutoff simply ensures that the number of non-zero elements in \mathbf{X} increases linearly with system size by setting $X_{\alpha\beta} = 0$ where the distance between the centers of NGWFs ϕ_α and ϕ_β is greater than the cutoff, r_X (Fig. 6).

The SW basis reduces the cost of evaluating the electron repulsion integrals (ERIs) in Eq. (24) in two ways. First, inserting a resolution-of-the-identity in the SW basis (SWRI) into Eq. (24) converts the four-index ERIs into products of more manageable two- and three-index objects, i.e.,

$$X_{\alpha\beta} = (\phi_\alpha \phi_\delta | f_p) V^{pq} (f_q | \phi_\gamma \phi_\beta) K^{\delta\gamma}, \quad (26)$$

where $\{f_p\}$ are SWs, V^{pq} is the inverse of the Coulomb metric matrix in the SW basis, and $(\phi_\alpha \phi_\delta | f_p)$ and $(f_q | \phi_\gamma \phi_\beta)$ are electrostatic integrals over NGWFs and SWs. Second, the evaluation of the two- and three-index objects themselves is simplified by the availability of analytic expressions for the Coulomb potentials of the SWs (SWpots).⁷²

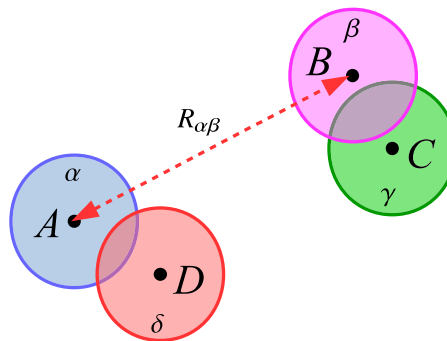


FIG. 6. Diagram illustrating the distance-based cutoff used in ONETEP's linear-scaling HFx implementation. NGWF localization spheres centered on atoms A to D are shown as shaded circles. When $R_{\alpha\beta} > r_X$, $X_{\alpha\beta} = 0$. Additionally, integrals where there is no overlap in either of the NGWF pairs (α, δ) or (β, γ) are zero and thus do not contribute to $X_{\alpha\beta}$ in Eq. (24).

The SWs are products of spherical Bessel functions, $j_l(x)$, and real spherical harmonics, $Z_{lm}(\hat{\mathbf{r}})$, localized within spheres of radius a ,

$$f_p(\mathbf{r}) = \begin{cases} j_{l_p}(q_p r) Z_{l_p m_p}(\hat{\mathbf{r}}), & r < a \\ 0, & r \geq a. \end{cases} \quad (27)$$

SWs are a convenient basis in which to expand quantities used in ONETEP, since, like NGWFs, they are both strictly localized within a sphere (radius a , set to the NGWF localization radius) and are systematically improvable with respect to a kinetic energy cut-off parameter.⁷³ Since SWs are angular-momentum-resolved, they have also found utility as a basis for ONETEP's distributed multipole analysis (DMA, Sec. IV F 1) and projected density of states (p-DOS, Sec. IV E 4) functionality.

Linear-scaling Hartree-Fock exchange in ONETEP was originally described in 2013.⁷¹ Although the initial implementation of ERI evaluation was not parallelized, $\mathcal{O}(N)$ scaling of computational cost was demonstrated for short polyethylene and polyacetylene chains (~ 10 s of atoms).

Recent developments have accelerated $\mathcal{O}(N)$ Hartree-Fock exchange in ONETEP, making hybrid functional calculations feasible on much larger systems, with ~ 2000 atoms becoming accessible with routinely available supercomputing hardware (Fig. 7). This dramatic improvement in performance was achieved by resolving bottlenecks in two major components of the implementation. First, a novel mixed numerical-analytic approach was developed for evaluating elements of the Coulomb metric matrix in the SW basis [V in Eq. (26)], reducing the computational cost (memory and execution time) of this step by orders of magnitude.⁷⁵ Second, the parallel implementation of the evaluation of Hartree-Fock exchange

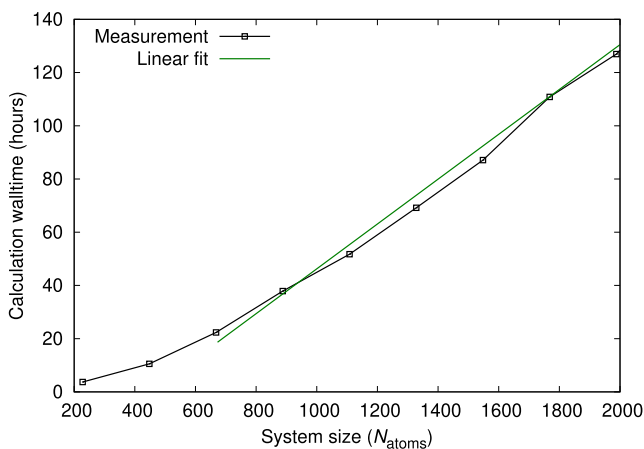


FIG. 7. Measured walltime of a calculation with a hybrid XC functional (black squares) as a function of system size for stacked polymer chains of increasing length [four chains of PBTZT-stat-BDTP-8 polymer analog (as studied in Ref. 74)]. The linear fit excludes the two smallest systems, where the linear-scaling regime has not yet been reached. Calculations were performed on the Iridis5 cluster using 32 compute nodes with 40 Intel Skylake CPUs each, running at 2.0 GHz, for a total of 64 MPI processes, each spawning 20 OMP threads. Calculation parameters—kinetic energy cutoff: 827 eV, NGWF localization radius: $8.0 a_0$, exchange cutoff: $20.0 a_0$.

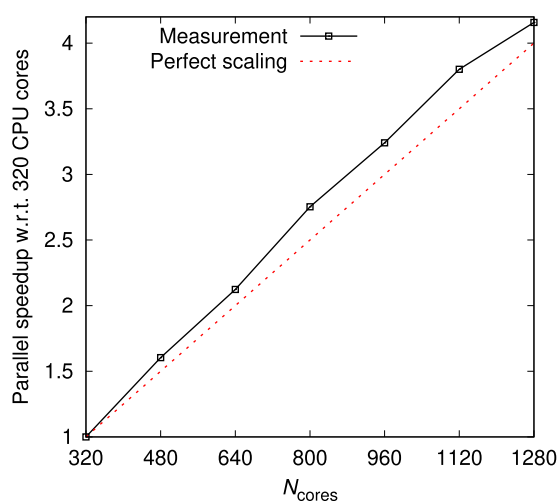


FIG. 8. Illustration of excellent strong scaling of parallel speedup with the number of CPU cores. The test case is an 888-atom stacked polymer chain system. The reference calculation uses 320 cores. The measured parallel speedup exceeds the “theoretical perfect scaling” in the entire range of CPU cores studied. This is an advantage of the memory-for-speed tradeoffs employed—adding further compute nodes increases not only the available CPU power but also the available RAM that can be used to cache intermediate results. Calculation parameters and CPU specifications are the same as in Fig. 7.

was completely overhauled with efficient data distribution among MPI processes, extensive use of OpenMP threading, and intelligent caching of computed quantities in memory.⁷⁶ The excellent parallel scalability of the overhauled implementation is illustrated in Fig. 8.

While great strides have been made in reducing the computational effort associated with hybrid XC calculations in ONETEP, using semi-local functionals remains considerably less costly. ONETEP can perform semi-local XC calculations on systems of $>10^4$ atoms with routinely available computational resources (see, e.g., Fig. 1). However, applying hybrid XC to systems of this size remains prohibitively expensive. A pragmatic solution to this issue is to employ embedded mean-field theory (EMFT, Sec. IV D 3), wherein hybrid XC may be applied in a small active region (<1000 atoms) that is embedded in a larger environment, treated with a semi-local XC functional.

The release of ONETEP (v5.2) at the time of writing supports global hybrid functionals in ground-state DFT calculations and conduction NGWF optimization (Sec. IV E 1). The core framework for applying global hybrids in linear-response time-dependent DFT (LR-TDDFT) (Sec. IV E 2; see also Sec. II.C of Ref. 77) has been developed and will be included as a user-accessible feature in an upcoming release. Support for range-separated hybrid functionals in ONETEP is also planned, though this is currently at an early stage of development and is thus a longer term aspiration.

2. Dispersion and van der Waals interactions

Properties and processes in many scientifically interesting and technologically important systems depend on weak dispersion interactions, such as London or van der Waals (vdW) interactions.

Examples include molecular crystals, physisorption of molecules on surfaces, and inter-layer interactions in layered two-dimensional materials and heterostructures. Large-scale electronic structure methods are well-suited for studying such structurally complex systems; however, standard density functionals such as the LDA and GGA do not account for these non-local and long-range interactions. ONETEP includes two methods designed to alleviate this problem.

a. Empirical dispersion corrections. A popular approach to remedying the problem of dispersion interactions is by including a damped London energy expression as part of the total DFT energy expression with perhaps the most well-known of these methods being the Grimme D2⁷⁸ and D3⁷⁹ variants.

In addition to the original D2 empirical dispersion correction of Grimme,⁷⁸ the damping functions of Elstner⁸⁰ and Wu and Yang⁸¹ have been implemented in the ONETEP code using new parameter sets reoptimized⁸² for biomolecular interactions. This reoptimization has proven critical for the accurate simulation of dispersion interactions, with the interaction energy root mean square error computed at the PBE level improving from 5.915 kcal/mol to 0.813 kcal/mol with respect to the literature values of the 60 complexes used to fit the parameters. Dispersion corrections have also been included in the calculation of forces, enabling accurate structures to be produced through geometry optimization (see Sec. IV C 1).

b. van der Waals density functionals. Dion *et al.*⁸³ developed a general non-local contribution to the correlation energy functional for describing vdW interactions within the framework of DFT, which takes the form

$$E_c^{\text{nl}} = \frac{1}{2} \iint d\mathbf{r}_1 d\mathbf{r}_2 \rho(\mathbf{r}_1) \phi(\mathbf{r}_1, \mathbf{r}_2) \rho(\mathbf{r}_2), \quad (28)$$

where $\phi(\mathbf{r}_1, \mathbf{r}_2)$ is a kernel that depends on the separation $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ and the electronic density $\rho(\mathbf{r})$ at and in the neighborhood of \mathbf{r}_1 and \mathbf{r}_2 [see Eqs. (14)–(16) of Ref. 83 for the definitions of $\phi(\mathbf{r}_1, \mathbf{r}_2)$]. Equation (28) is a double integral over the whole simulation cell, i.e., over six spatial dimensions, and computing it directly is prohibitive for large-scale systems. Román-Pérez and Soler⁸⁴ noted that the kernel ϕ may be written as

$$\phi(\mathbf{r}_1, \mathbf{r}_2) = \phi(q_1, q_2, r_{12}) \approx \sum_{\alpha\beta} \phi(q_\alpha, q_\beta, r_{12}) p_\alpha(q_1) p_\beta(q_2), \quad (29)$$

where q_1 and q_2 are the values of a universal function $q_0[\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|]$ (see Ref. 83 for details) at the points \mathbf{r}_1 and \mathbf{r}_2 , respectively, and the kernel is then represented in terms of a set of interpolating polynomials (cubic spline functions) $p_\alpha(q)$ for a specific set of values of the interpolation points q_α for each r_{12} . In Ref. 84, it was found that ~20 interpolation points on a logarithmic mesh was sufficient to accurately describe the kernel up to a cut-off value of $q_0 \leq q_c$ at which point q_0 is smoothly saturated. The benefit of using the expression in Eq. (29) for the kernel ϕ is that Eq. (28) can be evaluated using Fourier methods as follows:

$$E_c^{\text{nl}} = \frac{1}{2} \sum_{\alpha\beta} \iint d\mathbf{r}_1 d\mathbf{r}_2 \theta_\alpha(\mathbf{r}_1) \phi_{\alpha\beta}(r_{12}) \theta_\beta(\mathbf{r}_2) \quad (30)$$

$$= \frac{1}{2} \sum_{\alpha\beta} \int d\mathbf{k} \tilde{\theta}_\alpha^*(\mathbf{k}) \tilde{\phi}_{\alpha\beta}(k) \tilde{\theta}_\beta(\mathbf{k}). \quad (31)$$

$\theta_\alpha(\mathbf{r}) \equiv \rho(\mathbf{r}) p_\alpha[q_0(\mathbf{r})]$, $\phi_{\alpha\beta}(r) \equiv \phi(q_\alpha, q_\beta, r)$, and $\tilde{\theta}_\alpha(\mathbf{k})$ and $\tilde{\phi}_{\alpha\beta}(k)$ are their respective Fourier transforms. In practice, $\phi_{\alpha\beta}(k)$ is computed and tabulated once and for all on a radial grid, and the main computational effort during a calculation is associated with the evaluation of the Fourier transforms $\tilde{\theta}_\alpha(\mathbf{k})$. Expressions for the non-local potential corresponding to the non-local energy (required for self-consistent calculations and computing Hellmann–Feynman forces) may also be evaluated (see Refs. 84 and 85 for details).

In the case of the original vdW functional of Ref. 83 (known as “vdW-DF”), the total exchange and correlation functional is given by

$$E_{\text{xc}} = E_x^{\text{revPBE}} + E_c^{\text{LDA}} + E_c^{\text{nl}}, \quad (32)$$

which is the sum of the exchange term from revPBE,⁸⁶ the correlation term from PW92 LDA,⁸⁷ and the non-local correlation discussed above. In ONETEP, we have also implemented a number of further vdW density functionals within the same implementation framework, including vdW-DF2,⁸⁸ optPBE-vdW, optB88-vdW and optPBE $_{\kappa=1}$ -vdW,⁸⁹ optB86b-vdW,⁹⁰ and C09x-vdW.⁹¹

This work has been crucial in enabling applications to van der Waals heterostructures.^{92–94} In that work, vdW-DF functionals, particularly optB88,⁸⁹ have been shown to give interlayer distances in good agreement with experimental results for 2D material bilayers and heterostructures.

3. Hubbard-corrected density functional theory: DFT+U and beyond

DFT+U,^{95–97} occasionally known as LDA+U or LSDA+U, is a technique that can be used to significantly improve on conventional DFT’s description of strongly correlated systems and also many systems not normally thought of as strongly correlated. Local and semi-local functionals often fail to qualitatively reproduce the physics associated with localized, partially filled electronic subspaces of 3d and 4f character, including the emergence of Mott–Hubbard insulating gaps and magnetic order. Challenging systems in this regard include the first-row transition metals and lanthanoids (and the oxides of both), but also certain magnetic semiconductors, organometallic molecules, and even biomolecules.⁹⁸ Other localized subspaces, including those with 2p and 4d character, or those that are not partially filled, spin-polarized, or spherically symmetric or atom-centered, are increasingly being successfully targeted using flexible DFT+U implementations.

Several varieties and generalizations of DFT+U are implemented in ONETEP,^{99–101} including the corresponding contributions to the ionic forces and phonon energies. Extensions to conduction band optimization and the excited state regime (TDDFT+U¹⁰²), as well as combinations with other functionalities such as PAW (see Sec. IV A 1), are also implemented. The DFT+U method preserves linear-scaling with respect to system size, and in its conventional fixed-projector form, it introduces only a small increase in the computational pre-factor.⁹⁹ A key strength of the ONETEP DFT+U code is its modular nature and the ability to test diverse population analysis schemes when defining the DFT+U target subspaces. It provides a convenient starting point for the implementation of constrained DFT¹⁰³ (Sec. IV B 4) and DMFT^{104,105} (Sec. IV B 5) in ONETEP. The ONETEP DFT+U code has been used for fundamental developments in strong electronic correlation effects and self-interaction error

correction,^{106–108} and it has been extended to calculate Hubbard U and Hund's J parameters from first principles¹⁰⁹ and to approximately enforce Koopman's condition.¹¹⁰

DFT+ U works by adding a corrective term to the total energy for each subspace selected.^{95–97} This is commonly interpreted as a correction for the self-interaction error exhibited by the approximate XC functional,^{111,112} applied only to the subspaces where it is thought to be most problematic. Available in ONETEP is the DFT+ U + J energy correction functional,^{97,111,113} given by

$$E = \frac{1}{2} \sum_{I\sigma} \sum_{mm'}^M \left\{ (U^I - J^I) \left(n_m^{I\sigma m'} \delta_{m'}^m - n_m^{I\sigma m'} n_{m'}^{I\sigma m} \right) + J^I \left(n_m^{I\sigma m'} n_{m'}^{I\sigma m} \right) \right\}. \quad (33)$$

Here, U^I represents the net, occupancy–occupancy type Hartree + XC self-interaction of subspace I , taking account of the spin σ and screening by the remainder of the system appropriately.¹⁰⁹ J is a counterpart to U that quantifies subspace self-interaction of magnetization–magnetization type. We simply have DFT+ U either when $J = 0$ or when the last, unlike-spin term is neglected (giving an effective U : $U_{\text{eff}} = U - J$). Another J term can be optionally added (see Ref. 113), and separate Hubbard parameters for the linear (U_1) and quadratic (U_2) energy terms and spin channels (U^σ) can be used.¹¹⁰ $n_m^{I\sigma m'}$, the density matrix of subspace I and spin σ , is defined through the set of non-orthogonal projector functions $\{|\varphi_m^I\rangle\}$ by

$$n_m^{I\sigma m'} = \sum_{m''} \langle \varphi_m^I | \hat{\rho}^\sigma | \varphi_{m''}^I \rangle O^{Im''m'}. \quad (34)$$

$\hat{\rho}^\sigma$ is the global KS density matrix, and \mathbf{O}^I is the inverse of the projector overlap matrix $O_{mm'}^I = \langle \varphi_m^I | \varphi_{m'}^I \rangle$.

Different Hubbard corrections can be applied to atoms of the same chemical species, which is helpful if their chemical environments are not equivalent. The suggested population analysis for DFT+ U is based on the pseudo-atomic orbitals used to initialize the NGWFs at the beginning of a ONETEP calculation. Hydrogenic atomic orbitals of a chosen effective nuclear charge Z may alternatively be used, as can a subset of the optimized NGWFs from a previous calculation, in a projection update procedure, which can be iterated to self-consistency.¹⁰¹ The projection of the local density of states (LDOS) (Sec. IV E 4) onto the union of the DFT+ U + J subspaces can also be calculated. With U and J set to zero, the functionality can still prove useful, e.g., for its pseudo-atomic population analysis or for breaking spin and spatial symmetries.^{107,109}

4. Constrained density functional theory

Constrained density functional theory (cDFT)^{114–118} is an established variational^{103,119} generalization of DFT that enables practical simulation of charge (spin) excitation energies and transfer, as well as exchange coupling parameters in magnetic systems. cDFT also provides a practical and computationally convenient way of correcting intrinsic deficiencies of standard DFT, such as self-interaction errors (SIEs) and static electronic correlation.^{120–122} For an extensive review of the wide range of applications of cDFT, see Ref. 118 and references therein.

cDFT finds ground state solutions of a Kohn–Sham DFT-like Hamiltonian, augmented with an additional (external) potential:

the constraining potential(s), V_c . The energy W of a DFT system (of energy $E[n]$) subject to a constraint (N_c) on the electronic occupation of a suitably defined subspace can be written as¹²³

$$W[\hat{\rho}, V_c] = E[\hat{\rho}] + V_c (\text{Tr}[\hat{\rho}\hat{\mathbb{P}}] - N_c), \quad (35)$$

where $\hat{\rho}$ is the electronic density operator and $\hat{\mathbb{P}}$ is the projection operator for the cDFT-subspace(s). Maximization of W with respect to V_c determines the V_c needed to enforce the targeted electronic occupation (N_c) in the cDFT-subspace.^{103,119} Atomic forces can then be calculated,^{100,123} allowing geometry optimization and molecular dynamics (MD) simulations.

The cDFT implementation in ONETEP rests on the use of non-orthogonal, atom-centered projectors ($|\varphi_n\rangle$) to calculate the electronic occupation of the given cDFT-subspace, needed in Eq. (35), as

$$\text{Tr}[\hat{\rho}\hat{\mathbb{P}}] = \text{Tr}[\hat{\rho}|\varphi^m\rangle\langle\varphi_m|] = \langle\varphi_m|\hat{\rho}|\varphi_n\rangle O^{nm}, \quad (36)$$

where O^{nm} is an element of the matrix \mathbf{O}^{-1} and $O_{mn} = \langle\varphi_m|\varphi_n\rangle$. The mathematical framework underlying the tensorially invariant approach to cDFT-subspace occupancies, energies, and forces is extensively discussed in Refs. 100 and 123.

In ONETEP, the cDFT problem is solved via nested conjugated gradients,¹⁴ with the W maximization loop nested between the NGWF (ϕ_α) and density kernel (\mathbf{K}) loops. Figure 9(a) illustrates this procedure for a nitrogen molecule, where an electronic occupancy difference of one electron between the two atoms is enforced, demonstrating excellent convergence within 10 NGWF iterations.

In analogy with DFT+ U ^{99–101} (Sec. IV B 3), the cDFT problem can be solved self-consistently by using the NGWFs (ϕ_α) optimized for a given cDFT constraint as cDFT-projectors (φ_n) for a successive cDFT-optimization, as shown in Figs. 9(b) and 9(c), providing rapid convergence, as shown in Fig. 9(d). In a cDFT geometry optimization, a ground state geometry can be obtained initially using cDFT-projectors obtained self-consistently for the initial geometry. New cDFT-projectors can then be found for the current geometry, and the process iterated until W is self-consistent. Figure 9(d) shows that this converges extremely rapidly—within one or two steps. Figures 9(e) and 9(f) report the rate of convergence of this fully self-consistent W and the optimized N – N bond-distance (d_{NN}), respectively, with respect to the localization radius of the cDFT projectors and NGWFs for two different cDFT-subspaces. W and d_{NN} are different for the two subspaces for small radii, but as the radii grow, increasing the variational freedom of the problem, the values of W_{SC} and d_{NN} for the subspaces converge. This indicates that the self-consistent projector implementation in ONETEP is relatively insensitive to the choice of the cDFT-subspace representation, demonstrating its robustness.

cDFT in ONETEP has been applied to photovoltaic materials,^{124,125} to interfaces in graphene-based composite materials,¹²³ and to the limitations of cDFT in constraining SIE-free solutions in standard DFT exchange–correlation functionals.¹¹⁰ Going forward, we expect applications of cDFT to large-scale systems (e.g., biomolecules, nanoparticles, nanotubes, etc.) currently not amenable to simulation by standard DFT or beyond DFT approaches.

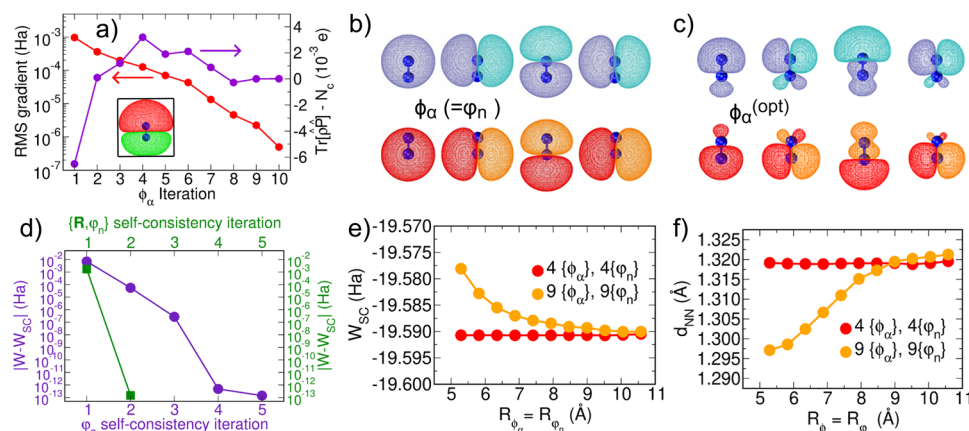


FIG. 9. (a) Convergence rate of the nested optimization of W with respect to ϕ_α and V_c for a N_2 molecule constrained to have a one electron difference between the two atomic sites. A contour plot of the final charge density difference, $\Delta\rho$, is shown in the inset (red: $\Delta\rho = +10^{-4} e \text{ \AA}^{-3}$; green: $\Delta\rho = -10^{-4} e \text{ \AA}^{-3}$). *In situ* optimization of the initial $\phi_\alpha (= \varphi_n)$ (b) in the presence of the applied constraints leads to a new set of optimized $\phi_\alpha^{(opt)}$ (c) that can then be used as cDFT-projectors (φ_n) for a successive cDFT-optimization. (d) Convergence rate of W as a function of the number of iterations toward cDFT-projector (φ_n , purple) and projector + geometry (R, φ_n , green) self-consistency (W_{SC}). Variation of the projector + geometry self-consistent optimized W_{SC} (e) and $N-N$ bond-distance (d_{NN}), (f) as a function of the localization radius of ϕ_α (R_{ϕ_α}) and φ_n (R_{φ_n}) for two different cDFT-subspace prescriptions with four ($2s, 2p$) and nine ($2s, 2p, 3d$) φ_n cDFT-projectors, respectively.

5. Dynamical mean field theory

Dynamical mean field theory (DMFT) is a Green's function method tailored for treating systems with strong local electronic correlation. Formally, correlation can be defined as the physics/chemistry due to multi-determinantal wavefunctions (that is, beyond Hartree–Fock theory). The effects of correlation encompass dispersion interactions, particle lifetimes, magnetism, satellites, collective excitations, the Kondo effect, Mott insulators, and metal–insulator transitions. Many (but not all) of these phenomena appear in systems containing transition metals or rare-earth atoms whose $3d$ - or $4f$ -electron shells are partially filled. Electrons in these shells are in especially close proximity with one another, and thus, their interaction is too pronounced to be adequately described by semi-local DFT, which can even provide qualitatively incorrect descriptions of the electronic structure.

DMFT maps the many-body electronic problem to an Anderson impurity model (AIM) Hamiltonian with a self-consistency condition.^{126,127} This model Hamiltonian includes Hubbard and Hund's-like terms, and local quantum fluctuations are fully taken into account—that is, the self-energy is dynamic but is assumed to be local. Consequently, DMFT can be expected to yield improved results where strong correlation is expected to be (a) important and (b) localized (cf. GW, which can account for non-local self-energies but only includes correlation effects perturbatively). Like DFT+ U (Sec. IV B 3), DMFT is typically used to treat localized regions where correlation is important, as defined by some atom-centered Hubbard projectors.¹²⁸ In the case of DFT+ U , each correlated subspace is subjected to an additional potential; in DFT+DMFT, these subspaces are subjected to a full Green's function treatment.

The AIM that the system is mapped on to contains both impurity and bath sites. The impurity sites typically correspond to the $3d/4f$ orbitals, while the bath sites collectively represent the rest

of the physical system. The AIM Hamiltonian is parameterized by inter-site hopping terms, which are determined via a fitting procedure, and Hubbard U and Hund's J parameters, which are external user-defined parameters that can be determined using linear response or cRPA.^{109,111,129} Alternatively, they can be treated as variational parameters.

Once the AIM Hamiltonian has been constructed, it is solved using exact diagonalization, continuous-time quantum Monte Carlo, or some other method, thereby obtaining the impurity Green's function and self-energy. This is typically the most computationally expensive step in DMFT and places a severe restriction on the total number of sites (impurity plus bath) one can use. The self-energies from each correlated subspace are then “unfolded” using these localized quantities to construct the total self-energy. During this unfolding, one must correct for the exchange and correlation already accounted for at the DFT level. The whole process is then repeated until the total self-energy is converged.

DMFT has proved capable of capturing complex electronic behavior where semi-local DFT fails. For example, DMFT corrects the lattice constants of strontium vanadate, iron oxide, elemental cerium, and delta-phase plutonium, which semi-local DFT substantially underestimates.^{130–132} Similar improvements have been observed for spectral properties such as densities of states and optical absorption.¹³³ DMFT also captures important dynamic properties that are enhanced by strong correlation, such as satellite peaks in the photoemission spectra of cerium and nickel.^{134,135} It can capture magnetic behavior both above and below the Curie temperature,¹³⁵ has been used to study high-temperature superconductivity in cuprates^{136,137} and other materials,^{138,139} and can predict mass renormalization in heavy fermion materials.^{140–144}

The implementation of DFT + DMFT using ONETEP involves an interface with the TOSCAM code toolbox.¹⁰⁵ ONETEP and TOSCAM are responsible for separate sections of the DMFT loop: ONETEP performs

much of the initialization and mappings, while *TOSCAM* is responsible for solving the AIM. As the calculation proceeds, control alternates between these two programs, with the entire procedure being driven by an overarching script. *TOSCAM* will be freely available for download in the near future. For more details, refer to Ref. 105. *ONETEP* + *TOSCAM* has already been used to explain the insulating M_1 phase of vanadium dioxide,¹⁰⁴ to demonstrate the importance of Hund's coupling in the binding energetics of myoglobin,^{145,146} and to study the superexchange mechanism in the dicopper oxo-bridge of hemocyanin and tyrosinase.¹⁴⁷

C. Dynamics and structure optimization

1. Geometry optimization

ONETEP can relax ionic positions with respect to forces either in Cartesian coordinates using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (following the original *CASTEP* implementation⁴⁹) or in the limited memory l-BFGS version;¹⁴⁸ or in delocalized internal coordinates using the approach of Andzelm *et al.*¹⁴⁹

The forces in *ONETEP* are calculated as the sum of local and non-local pseudopotential terms with Ewald and, where applicable, non-linear core correction terms.⁵³ Since *ONETEP*'s NGWF basis is atom centered and local, the forces also include Pulay terms due to the change in basis functions with respect to atomic positions.⁵⁴ The total forces in *ONETEP* are calculated as

$$\mathbf{F}_{\text{total}} = \mathbf{F}_{\text{loc}} + \mathbf{F}_{\text{NL}} + \mathbf{F}_{\text{Ewald}} + \mathbf{F}_{\text{NLCC}} + \mathbf{F}_{\text{Pulay}}. \quad (37)$$

Pre-conditioners are used to speed up the geometry optimizers in *ONETEP*. By default, (l-)BFGS uses the physics-based preconditioner suggested by Pfrommer,¹⁵⁰ which takes an estimate for the average optical phonon frequency at the Γ -point as input. The algorithm also estimates this quantity throughout the calculation in order that restarts and future calculations may be sped up. *ONETEP*'s geometry optimization performance may be improved dramatically¹⁵¹ by using the force-field based pre-conditioners of Packwood *et al.*¹⁵² These methods take simple interatomic potentials to construct initial sparse approximations to the Hessian matrix. These pre-conditioners are available with the l-BFGS geometry optimizer.

ONETEP does not perform any extrapolation of NGWFs or density kernels between geometry optimizer iterations; this helps prevent biasing the end result. For computational expediency, however, *ONETEP* instead takes the final NGWFs from the previous geometry optimizer iteration as initial NGWFs to the next iteration. The NGWFs are shifted with respect to the simulation grid to keep the ions at their center. On restarted calculations *ONETEP* can read the Hessian matrix or update vectors along with the final NGWFs and density kernel from the previous run.

ONETEP supports constraints on individual atoms in all geometry optimizer modes. Constraints may be to fix an atom in place, to keep it on a vector, or to keep it on a plane. Relaxation of the lattice parameters, however, is not possible within *ONETEP* due to the limited demand for the majority of applications of *ONETEP*. Although *ONETEP* does not calculate stresses for periodic systems, pressure can be applied to finite systems via the electronic enthalpy method, demonstrated in Sec. V C.

2. Phonons and thermodynamic properties

The calculation of phonon frequencies is implemented through the well-established finite-displacement methodology (FDM);^{153,154} in our case, the dynamical matrix is calculated using a central difference formula with a choice of either a three- or five-point stencil in each Cartesian direction (i.e., $6N/12N$ displacements in total for an N -atom system).

For crystalline materials, the FDM requires a supercell geometry to capture a sufficiently dense grid of q -points in the vibrational Brillouin zone; this is a natural choice for a linear-scaling code such as *ONETEP*. If a crystalline material is specified, only the atoms in the unit cell are displaced, and either a real-space truncation of the force constant matrix or a Slater–Koster style interpolation is used to obtain the phonon frequencies at arbitrary q -points. In order to significantly speed up the calculation, the converged self-consistent density kernel and NGWFs from the unperturbed system are taken as the starting guess for each displacement; it is also possible to limit the degrees of freedom to reduce the number of displacements required and to use a task-farming approach to run the displacements in batches as independent *ONETEP* calculations that can be run in parallel.

In addition to the phonon band structure and DOS, the code calculates a number of thermodynamic properties: the zero-point energy, and the free energy, entropy, internal energy, and specific heat in a given temperature range. These are calculated on a regular grid of q -points, which can be chosen independently of the original supercell repetitions. *ONETEP* does not calculate the effect of the long-range electric fields associated with longitudinal phonons near the Γ point (LO–TO splitting). Tests on bulk silicon have shown that well-converged quantities can be obtained using a $5 \times 5 \times 5$ cubic supercell (1000 atoms) and a $10 \times 10 \times 10$ q -point grid for the thermodynamic properties, while the phonon DOS requires a significantly denser q -point grid.¹⁵⁵

3. First principles molecular dynamics

First principles molecular dynamics (MD) simulations constitute a powerful tool to study systems that are inherently dynamical, such as in conformational analysis of (bio)molecules at finite temperatures, bond breaking in chemical reactions, adsorption, and solvation phenomena. MD simulations are also employed to search for equilibrium structures under different external conditions^{156–158} as well as for computing vibrational, IR, and Raman spectra¹⁵⁹ (see Sec. IV E 6). Two main paradigms exist to perform MD simulations within the DFT framework: Born–Oppenheimer MD (BOMD) and Car–Parrinello MD (CPMD).^{160,161} The Car–Parrinello scheme involves direct propagation of the electronic degrees of freedom along with the nuclear degrees of freedom. On the other hand, in the Born–Oppenheimer MD (BOMD) scheme, one solves the static electronic problem to find the KS energy self-consistently [$E_{\text{SCF}}(\{\mathbf{R}_j\})$] at every MD step. Hence, the electronic wavefunctions are never propagated, since electrons are assumed to instantaneously adjust to the motion of the nuclei¹⁶² (Born–Oppenheimer approximation). This makes BOMD more suitable to linear-scaling approaches since the localization constraints are more easily imposed during the KS energy minimization. The other main advantage of BOMD over CPMD is that a larger time step can be chosen, which, in turn, results in fewer MD steps and reduced simulation times. Moreover, no

extra ambiguous parameters, such as fictitious electronic masses, are required to stabilize the dynamics. On the other hand, in BOMD, the stability of the generated trajectory is quite sensitive to the accuracy of the ionic forces and more generally to the degree of convergence of the SCF loop. This is due to the well-known fact that the error in the forces is first order with respect to the error in the incompletely converged wave-function, even though the error in the Kohn–Sham energy is second order.¹⁶³

Within the ONETEP framework, the BOMD equations assume the following form:

$$M_I \frac{d^2 \mathbf{R}_I}{dt^2} = \mathbf{F}_I^{\text{SCF}} = -\nabla_I E_{\text{SCF}}(\{\mathbf{R}_I\}) \\ = -\nabla_I \left[\min_{\{\phi_\alpha\}} \left[\min_{\{\mathbf{K}\}} E[\{\phi_\alpha\}, \{\mathbf{K}\}; \{\mathbf{R}_I\}] \right] \right], \quad (38)$$

where \mathbf{R}_I is the position of the I th ion at time t and M_I is its mass. $\mathbf{F}_I^{\text{SCF}}$ is the force acting on ion I (for a detailed description of the computation of ionic forces in ONETEP, see Ref. 53).

The BOMD equations of motion are integrated using a finite difference scheme (which usually preserves the symplectic structure of phase space)—in ONETEP, we employ the Velocity-Verlet (VV) algorithm.¹⁶⁴

It is possible to couple different thermostats to the BOMD equations in Eq. (38) in order to sample ensembles other than the microcanonical (also known as *NVE*). In particular, the following thermostats are available in ONETEP: Andersen,¹⁶⁵ Langevin,¹⁶⁶ Nosé–Hoover chains,¹⁶⁷ Berendsen,¹⁶⁸ and Bussi.¹⁶⁹

For each MD step, the dominant computational effort is associated with solving the KS equations self-consistently, whereas the computation of the ionic forces and the integration of the BOMD equations amount to a small fraction of the walltime. The number of SCF cycles required to reach a given level of self-consistency can be substantially reduced by using improved initial guesses for the electronic degrees of freedom. Various approaches have been implemented, which exploit the information included in the MD history to build more efficient initial guesses and speed up the SCF loop. These methods can be grouped into two major categories: extrapolation methods and propagation methods. Various flavors of linear and polynomial extrapolation are available in ONETEP. Here, the electronic degrees of freedom are represented as polynomials with respect to time t . All these schemes assume that the NGWFs and the density kernel can be treated as independent degrees of freedom and therefore can be independently extrapolated. However, the density kernel can be considered as the representation of the density matrix in the NGWF basis. Therefore, one can also first extrapolate the NGWFs and then project the density kernel onto this new basis. Accordingly, two density kernel projection methods have been implemented: naïve projection and tensorial-compliant projection.¹⁷⁰ The latter takes into account the tensorial nature of the density kernel in the NGWF basis.

For the propagation schemes, we have implemented the approach developed by Niklasson *et al.* based on an extended Lagrangian (EL) propagation.¹⁷¹ These schemes have shown superior performances in terms of energy conservation compared to simpler schemes (by alleviating the issue of the breaking of time-reversal symmetry). In practice, three flavors of EL propagation can be chosen in ONETEP: the naïve approach (deprecated), EL with dissipation

(dEL),¹⁷¹ and inertial EL with a thermostat (iEL).¹⁷² The latter two methods show similar performance in terms of energy conservation and speed-up of the SCF cycle.¹⁷³

It should be stressed that the self-consistent solution is independent of the initial guess only at exact convergence. In practice, some breaking of time-reversal symmetry is to be expected, resulting from memory effects due to the failure of the Hellmann–Feynman theorem, as the electronic system is never exactly converged to its ground state.¹⁶³

4. Nudged elastic band calculations

Given decent initial guesses, geometry optimization (Sec. IV C 1) is a staple in locating useful minimum-energy structures. Transition states, often just as important as the endpoints in chemical reactions, diffusion problems, and solid state transitions, are less straightforward to locate. They are the maximum-energy points on the minimum-energy paths connecting structures. Several approaches exist to approximate these saddle points, with the nudged elastic band (NEB)¹⁷⁴ approach being among the most robust. NEB attempts to minimize several intermediate structures, or “beads,” onto the minimum-energy path from which the maximal point can be approximated or directly measured. The need for a strong transition state searching functionality in linear-scaling DFT packages is highlighted in a recent review of the opportunities and challenges facing this community.¹⁷⁵ To this end, a robust and scalable implementation of NEB has been designed for ONETEP.

NEB approximates the minimum energy path by connecting several intermediate structures, known as beads, typically initially guessed from linear interpolation, to their neighbors with springs of zero natural length. In order to ensure the beads lie equally distributed and on the minimum energy path, certain components of both the real and spring forces are projected out. This chain-of-states is then evolved until the “nudged” force on each bead is below some tolerance. Given a perfect approximation of the path tangent at each point on the path, the beads will lie on the minimum energy path, regardless of bead count. In practice, the path tangent at each bead’s position is approximated,¹⁷⁶ requiring a few beads to resolve minimum energy paths.

The implementation of NEB in ONETEP handles each bead in parallel with an internal driver approach. Based on a comparison of NEB minimizers,¹⁷⁷ both FIRE and the chain-global I-BFGS (GL-BFGS) minimizers were selected for the implementation in ONETEP. A climbing-image mode¹⁷⁸ is available in which the highest-energy bead moves in a modified potential energy surface (PES) in which the saddle point becomes an energy minimum, allowing the bead to climb the path to sample the transition state exactly.

NEB is not a global search—the path it finds depends on the initial guess path. For this reason, significant control is given to the user over the initial definition of the path; an optional intermediate structure can be defined in the input file or each bead’s position can be directly read from an input XYZ file.

D. Including environmental effects

1. Implicit solvent model

ONETEP’s $\mathcal{O}(N)$ formalism and scalable implementation make it an excellent tool for simulating large and complicated systems,

such as biomolecules, nanomaterials, and surfaces. The environment in which these types of systems exist—often some kind of liquid solvent—can intimately affect their behavior and properties. Some means of representing solvent environments is therefore vital for realistic and informative electronic structure calculations on these systems.

A direct approach to solvent modeling would be to simply include a number of solvent molecules in the electronic structure simulation of the system of interest. This explicit quantum mechanical approach is superficially very simple, as it does not require the system to be partitioned into solute and solvent regions and treats internal solute and solute–solvent interactions on an equal footing. However, this approach is generally prohibitively expensive, even with an $\mathcal{O}(N)$ method. In addition, even if sufficient resources are available to run an electronic structure calculation on the entire solute–solvent system, realistic results require many calculations in order to statistically average over solvent degrees of freedom.

Explicit atomistic solvent models can be made cheaper by treating the solvent environment at a lower level of theory, as in embedded mean-field theory (EMFT, Sec. IV D 3) and QM/MM, Sec. IV D 2). A further simplification is to represent the solvent as an unstructured dielectric medium. In this case, solvent–solvent interactions and solvent degrees of freedom are accounted for implicitly, and we only need to be concerned with the atomistic details of the solute.

In ONETEP's implicit solvent model, the solute is embedded in a cavity within the solvent, which is represented as a polarizable dielectric medium. The polarization of the dielectric medium by the solute charge induces a reaction potential that is incorporated into the Kohn–Sham Hamiltonian. The result is a self-consistent reaction field (SCRf) method in which the electrostatic interactions between the solute and implicit solvent are captured within the SCF procedure.

ONETEP's implicit solvation model shares the essential features of the continuum dielectric SCRf method with other widely adopted models, such as CONductor-like Screening MODEL (COSMO)¹⁷⁹ and PCM,¹⁸⁰ but differs in the specifics of its implementation. In particular, the model has the following key features:

- The generalized Poisson equation [Eq. (13) with $\alpha = -4\pi$, since ONETEP uses atomic units] is solved for the total electrostatic potential, $v(\mathbf{r})$, in 3D using a high-order defect-corrected multigrid method [using the DL_MG library (see Sec. IV A 5)].⁵⁵
- The dielectric permittivity, $\epsilon(\mathbf{r})$, is defined in terms of the electron density, $n_{\text{elec}}(\mathbf{r})$, and smoothly varies between vacuum [$\epsilon(\mathbf{r}) = 1$] and bulk solvent [$\epsilon(\mathbf{r}) = \epsilon_\infty$] values based on the form proposed by Fattebert and Gygi,¹⁸¹

$$\epsilon(\mathbf{r}) = 1 + \frac{\epsilon_\infty - 1}{2} \left(1 + \frac{1 - (n_{\text{elec}}(\mathbf{r})/n_0)^{2\beta}}{1 + (n_{\text{elec}}(\mathbf{r})/n_0)^{2\beta}} \right), \quad (39)$$

which has only two parameters: n_0 , the electron density value at the center of the solute–solvent interface, and β , which determines the width of the solute–solvent transition [between $\epsilon(\mathbf{r}) = 1$ and ϵ_∞].

- Non-electrostatic contributions to the free energy of solvation are computed by scaling the surface area of the

vacuum cavity surrounding the solute, S , by a surface tension, γ , as used by Scherlis *et al.*¹⁸² to extend Fattebert and Gygi's electrostatic model, i.e.,

$$\Delta G_{\text{non-es}} = \gamma_{\text{eff}} S [n_{\text{elec}}]. \quad (40)$$

$S[n_{\text{elec}}]$ is evaluated by integration over a thin film centered on n_0 ,^{182,183} and γ_{eff} is an effective surface tension, scaled to take account of solute–solvent dispersion–repulsion effects.¹⁸⁴

ONETEP's minimal parameter implicit solvent model (MPSM) enables the computation of free energies of solvation of small molecules with accuracy comparable to or better than other widely used implicit solvent models, but with only two empirical parameters (e.g., Fig. 10—see Ref. 184 for further details). The model has also proven to be very useful for large-scale DFT calculations, where screening due to the polarizable dielectric medium corrects for spurious electrostatic effects that arise at artificial surface–vacuum interfaces. This effect has been very convincingly demonstrated for proteins and water clusters, where unscreened dipole moments across the system *in vacuo* can artificially close the HOMO–LUMO gap (e.g., Fig. 11—see Ref. 192).

ONETEP now also includes the functionality to calculate energies in electrolyte solutions in both open boundary conditions (OBCs) and periodic boundary conditions (PBCs).¹⁹³ The source term in the

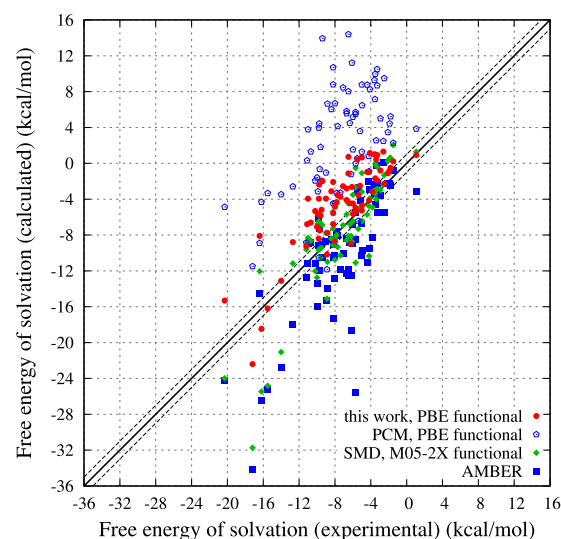


FIG. 10. Computed free energies of solvation for 71 neutral molecules^{185,186} plotted against experimental values from the Minnesota solvation database (2009).¹⁸⁷ Energies computed using the minimal parameter implicit solvent model (MPSM) in ONETEP (i.e., “this work”) are compared against those computed using other implicit solvent models (PCM¹⁸⁰ and SMD¹⁸⁸) and a force-field Poisson–Boltzmann approach in AMBER 10.¹⁸⁹ The PCM results were computed using the semi-local PBE functional,¹⁹⁰ and the SMD results were computed using the hybrid M05-2X functional.¹⁹¹ Figure reprinted with permission from Dziedzic *et al.*, *Europhys. Lett.* **95**, 43001 (2011). Copyright 2011 Europhysics Letters Association (EPLA) (see the original article for further calculation details).

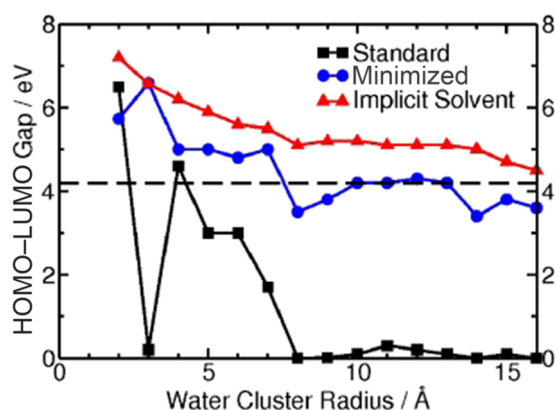


FIG. 11. HOMO–LUMO gaps for water clusters extracted from an equilibrated classical MD simulation, computed *in vacuo* (black squares), after classical relaxation (blue circles), and with an implicit solvent (red triangles). The HOMO–LUMO gap of bulk water (4.2 eV) is shown for reference (black dashed line). Figure adapted (cropped) with permission from Fig. 1 of Lever *et al.*, *J. Phys.: Condens. Matter* **25**, 152101 (2013). Copyright 2013 Author(s), licensed under a Creative Commons Attribution 3.0 Unported License. (see the original article for further calculation details).

GPE now also includes the charge density of mobile electrolyte ions, i.e.,

$$\nabla \cdot (\epsilon(\mathbf{r}) \nabla v(\mathbf{r})) = -4\pi(n_{\text{tot}}(\mathbf{r}) + n_{\text{ions}}[v](\mathbf{r})) \quad (41)$$

with $n_{\text{ions}}[v](\mathbf{r})$ dependent on the electrostatic potential, v . The charge density of mobile electrolyte ions can be written in terms of their space-dependent concentration, $c_i[v](\mathbf{r})$, as

$$n_{\text{ions}}[v](\mathbf{r}) = \sum_{i=1}^p z_i c_i[v](\mathbf{r}), \quad (42)$$

where z_i is the charge on ion i . The concentrations can be described in terms of the bulk concentrations of ions (c_i^∞) as

$$c_i^{\text{OBC}}[v](\mathbf{r}) = c_i^\infty \lambda(\mathbf{r}) \exp(-\beta z_i v(\mathbf{r})), \quad (43)$$

$$c_i^{\text{PBC}}[v](\mathbf{r}) = c_i^\infty \lambda(\mathbf{r}) \left[\frac{\int_V \lambda(\mathbf{r}) d^3 \mathbf{r}}{\int_V \lambda(\mathbf{r}) \exp(-\beta z_i v(\mathbf{r})) d^3 \mathbf{r}} \right] \exp(-\beta z_i v(\mathbf{r})), \quad (44)$$

where $\beta = \frac{1}{k_B T}$ and $\lambda(\mathbf{r})$ is the electrolyte accessibility function, which smoothly varies from zero in the solute to one in the bulk electrolyte solution and prevents electrolyte charge accumulation near the solute. The expression in OBCs is the well-known Boltzmann concentration of ions. The expression in PBCs includes a renormalization factor in order to conserve the ions. ONETEP also includes the functionality for a linear approximation of the above concentrations when the electrostatic potential is small or when $z_i v \ll k_B T$. The above set of equations is solved self consistently within DL_MG (Sec. IV A 5).

2. QM/MM and polarizable embedding

Due to the long-ranged nature of Coulombic interactions, it is often necessary to include hundreds of atoms, if not more, in DFT calculations before the relevant properties are sufficiently converged.¹⁹⁴ Suitable truncation of larger systems to computationally affordable sizes usually proves to be far from trivial.¹⁹² Hybrid (QM/MM) approaches offer one way of addressing this problem, embedding a quantum mechanical (QM) calculation within an environment described at the molecular mechanical (MM) level of theory. This is well justified in systems whose properties of interest are spatially localized, with the quality of obtained results depending on the physical soundness of the interface and the accuracy of the MM description.

ONETEP implements a fully self-consistent interface to the MM suite TINKER,¹⁹⁵ which implements the state-of-the-art polarizable force-field AMOEBA^{196,197} and standard fixed-point-charge force-fields, such as GAFF.¹⁹⁸ In the case of AMOEBA, the two subsystems (QM and MM) interact via multipolar electrostatics and are fully mutually polarizable with direct energy minimization encompassing the entire (QM + MM) system. Distributed multipole analysis (DMA,^{199–201} Sec. IV F 1) is used to represent the QM charge density in the form of point multipoles (up to a quadrupole) that are compatible with the AMOEBA formalism. In order to prevent unphysical charge spilling from the QM subsystem onto the point charges, which is an artifact of the point multipole model, damping of electrostatic interactions is used. Additionally, we employ a simple model for Pauli repulsion between QM and MM subsystems, yielding a suitable repulsive potential that ensures a physically sound localization of the QM charge density.

Our implementation^{202,203} is the only one to date enabling *in situ* minimization of localized orbitals in the presence of a polarizable MM embedding and the only one where the cost of the QM calculation is linear-scaling. This model has been demonstrated to yield remarkably good energetics across the QM/MM interface (as compared to fully QM calculations) even for minimal QM subsystems.

3. Embedded mean-field theory

In many potential applications of *ab initio* atomistic modeling, the important physics or chemistry of the system of interest is associated with a small “active region,” which would ideally be described using a high accuracy QM method. However, the rest of the system, or the “environment,” is still significant through its interaction with the active region. QM/MM (Sec. IV D 2) and implicit solvent (Sec. IV D 1) methods are two ways of approaching this problem; however, if the environment also requires QM treatment, a quantum embedding approach is required. Quantum embedding schemes enable an expensive high accuracy method to be used for the active region and a cheaper lower accuracy method for the environment.^{204,205}

The quantum embedding formalism implemented within ONETEP is the embedded mean field theory (EMFT) scheme, first proposed by Fornace *et al.*^{206,207} In general, the density matrix or kernel is partitioned into regional sub-blocks,

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{BA} & K_{BB} \end{pmatrix}, \quad (45)$$

where A labels the active region and B labels the environment. Each atom and its associated NGWFs are assigned to a region.

For DFT-in-DFT embedding, the only difference between the lower and higher accuracy methods will be in the exchange–correlation functional— E_{xc} in Eq. (6). The other terms in Eq. (6), which we will collectively call $E_0[\{\phi_\alpha\}, \mathbf{K}]$ here for brevity, are the same at both levels of theory. In EMFT, the total energy for the system is then written as

$$E^{\text{EMFT}}[\{\phi_\alpha\}, \mathbf{K}] = E_0[\{\phi_\alpha\}, \mathbf{K}] + E_{xc}^{\text{low}}[n] + (E_{xc}^{\text{high}}[n_{AA}] - E_{xc}^{\text{low}}[n_{AA}]), \quad (46)$$

where E_{xc}^{low} and E_{xc}^{high} are the exchange–correlation functionals of the lower and higher accuracy methods, respectively, and n and n_{AA} are the total and active region electron density, respectively.

The block orthogonalization procedure used in previous implementations of EMFT to avoid charge spilling²⁰⁸ interferes strongly with the ability to optimize the basis functions (the NGWFs) in ONETEP. To avoid this, all NGWFs are optimized at the lower level of theory and then fixed for the optimization of \mathbf{K} within EMFT. Results on a variety of test systems show that the error associated with this approximation is 1% or less of the difference between the full high- and low-level calculations.²⁰⁷ Embedding a hybrid functional calculation within a semi-local DFT environment enables us to extend the use of HFX to periodic structures containing a subsidiary active region. Our implementation of EMFT has been successfully applied to the calculation of excited state and encapsulation energies for a variety of systems consisting of several thousand atoms²⁰⁷ and is capable of performing hybrid-in-semi-local ground state total energy calculations. Forces are also available within the EMFT framework. Future developments will seek to make this compatible with time-dependent DFT (TDDFT) (Sec. IV E 2), enabling the use of time-dependent EMFT (TD-EMFT).²⁰⁹

4. Open boundary conditions

The psinc basis functions employed in ONETEP are unitary transformations of plane waves (cf. Sec. II). This enables ONETEP to leverage Fourier transforms for parts of the calculation that are cumbersome in real space, such as evaluating the Hartree potential, but also implies an underlying periodicity of the system. This can become problematic for systems that are not genuinely periodic, particularly when they are not charge-neutral or possess a significant dipole moment—in such cases, non-negligible artificial interactions between the periodic images arise. The traditional approach of vacuum padding and correcting²¹⁰ for the dependence of the energy on the size of the simulation cell has since been improved upon with a number of techniques for implementing open boundary conditions (OBCs) (see Ref. 211 and references therein). In ONETEP, the artificial periodicity can be abrogated (thus modeling OBCs) using three different approaches: cut-off Coulomb (CC),^{212,213} Martyna–Tuckerman (MT),²¹⁴ and direct real-space calculations using a multigrid solver (MG)⁵⁵ to solve the Poisson equation.

a. Cut-off Coulomb. The standard periodic approach evaluates the Coulomb integral over all space, assuming an infinite periodic repetition of the simulation cell. This leads to significant errors in isolated systems with net charges or large multipole moments, as

the periodic images artificially interact with one another via long-range electrostatics. The cut-off Coulomb approach counters this by only allowing Coulombic interactions within a specified region, thereby emulating the electrostatics of an isolated system with a periodic representation of the charge density.²¹² By selecting an appropriate region, \mathcal{D} , one can eliminate spurious electrostatic interactions between charges in the periodic replicas and the home cell while also retaining the correct description of the potential between charges in the isolated system,

$$v^{\text{CC}}(\mathbf{r}, \mathbf{r}') = \begin{cases} \frac{1}{|\mathbf{r}-\mathbf{r}'|}, & |\mathbf{r}-\mathbf{r}'| \in \mathcal{D} \\ 0, & |\mathbf{r}-\mathbf{r}'| \notin \mathcal{D}. \end{cases} \quad (47)$$

The modified Coulomb interaction term can then be convolved with the charge density to give the Hartree potential

$$V_{\text{H}}^{\text{CC}}(\mathbf{r}) = \int_{\mathcal{D}} v^{\text{CC}}(\mathbf{r}, \mathbf{r}') n(\mathbf{r}') d\mathbf{r}'. \quad (48)$$

Here, point \mathbf{r} experiences Coulomb potentials from all $n(\mathbf{r}')$ in region \mathcal{D} , and potentials from charge densities outside this region are set to 0. $V_{\text{H}}^{\text{CC}}(\mathbf{r})$ is evaluated in reciprocal space through a Fourier transform of (48). The form of the expression for $V_{\text{H}}^{\text{CC}}(\mathbf{G})$ is determined by the shape of the region \mathcal{D} , which also reflects the periodicity of the system. The simplest of these is the 3D Coulomb cut-off type, which truncates the Coulomb interaction to a sphere of radius R_c , thereby eliminating periodicity in the system. This can be extended to 2D (slab) and 1D (wire) geometries, maintaining periodic behavior in a desired plane or axis. The modified derivations of $V_{\text{H}}^{\text{CC}}(\mathbf{G})$ for each cut-off type are given in the work of Rozzi *et al.*²¹³ and Sohier *et al.*²¹⁵

To ensure sufficient space between the charge densities of the home cell and those of the periodic images, ONETEP places the unit cell within a larger, padded cell where $n(\mathbf{r}) = 0$ for all \mathbf{r} . Calculations of $V_{\text{H}}^{\text{CC}}(\mathbf{r})$ on a grid are performed by placing the charge density of the unit cell within the padded simulation cell. This is Fourier transformed to give $\tilde{n}(\mathbf{G})$ and multiplied with the appropriate values of $v^{\text{CC}}(\mathbf{G})$ to give $V_{\text{H}}^{\text{CC}}(\mathbf{G})$. These values are subsequently reverse Fourier transformed, and the real-space Hartree potential is extracted from the padded grid to give $V_{\text{H}}^{\text{CC}}(\mathbf{r})$. A similar approach is used for the local pseudopotential energy term, while for the core–core interaction, the usual Ewald sum is replaced with a direct Coulombic sum of interactions computed in real space. For more details about the implementation in ONETEP, refer to Ref. 211.

b. Martyna–Tuckerman approach. This approach retains Fast Fourier Transforms (FFTs) for the calculation of Hartree and local pseudopotential energy terms. The Coulomb operator is still periodic but is modified in such a way that the interactions between neighboring cells are eliminated.

Under PBCs, the real-space representation of a function $f(\mathbf{r})$ is

$$f_{\text{per}}(\mathbf{r}) = \sum_{\mathbf{G} \in \text{cell}} \tilde{f}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (49)$$

where $\tilde{f}(\mathbf{G})$ is the Fourier transform of $f(\mathbf{r})$, and the reciprocal space vectors \mathbf{G} are chosen to be commensurate with the simulation cell. The resultant function $f_{\text{per}}(\mathbf{r})$ has the periodicity of the cell and is a superposition of periodically repeated functions $f(\mathbf{r})$, one in each cell. This is illustrated on the example of a Gaussian function in Fig. 12 (top).

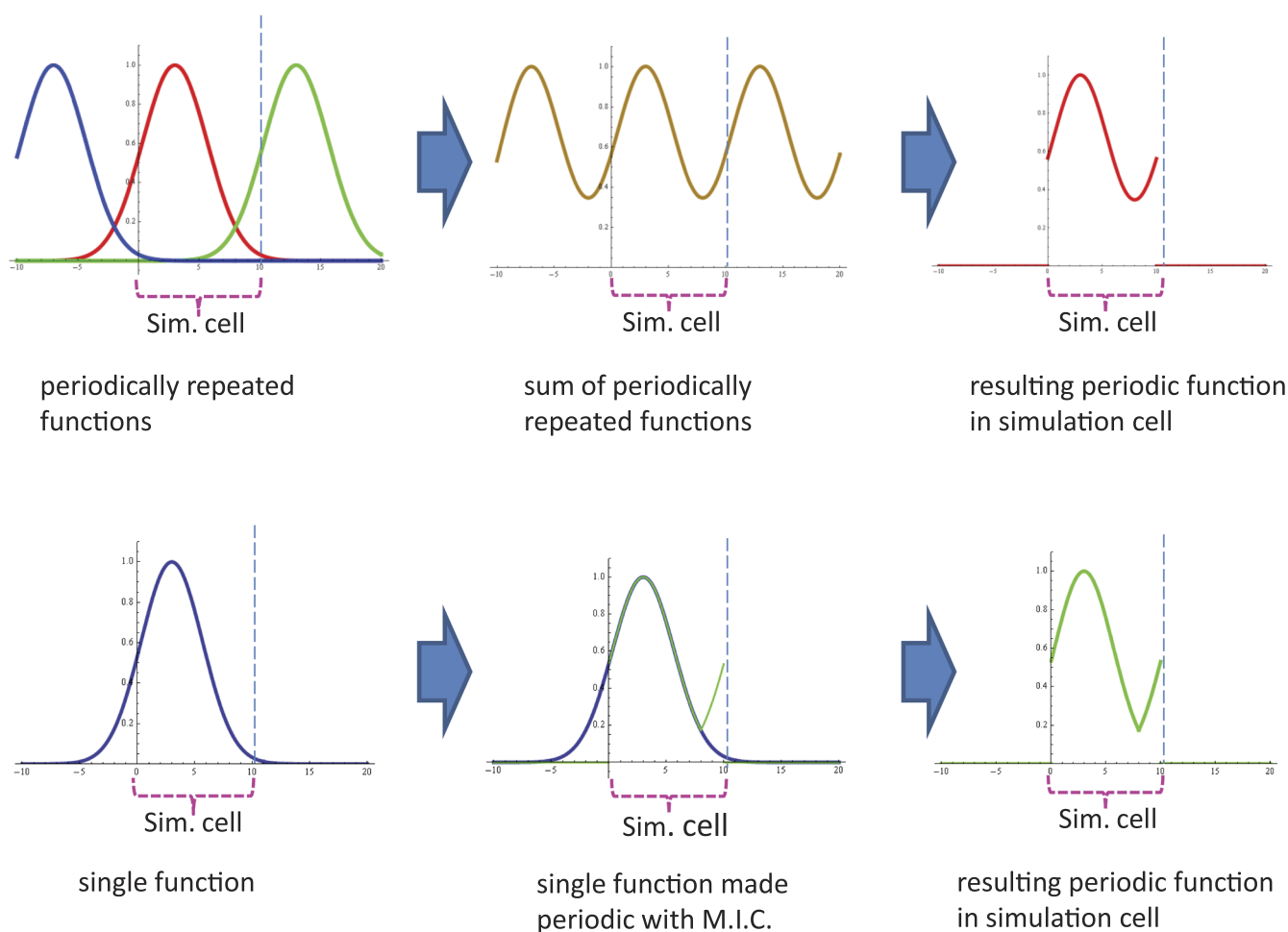


FIG. 12. Different techniques for making a function obey periodic boundary conditions inside a simulation cell, demonstrated for a Gaussian function. (Top) The Fourier transform approach. The resulting function is the same as the one that would be obtained by a superposition (sum) of periodically repeated Gaussians. (Bottom) The Minimum Image Convention (MIC) approach: the resulting function is the same as the one that would be obtained by having a single Gaussian in the simulation cell and making it periodic by applying the MIC. Figure reprinted with the permission from Hine *et al.*, *J. Chem. Phys.* **135**, 204103 (2011). Copyright 2011 AIP Publishing LLC.

The Martyna–Tuckerman approach²¹⁴ replaces $f_{\text{per}}(\mathbf{r})$ with a different function $f_{\text{MIC}}(\mathbf{r})$,

$$f_{\text{MIC}}(\mathbf{r}) = \sum_{\mathbf{G} \in \text{cell}} \bar{f}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}, \quad (50)$$

which is also periodic but is not constructed as a superposition of periodically repeated functions; rather, the minimum image convention (MIC) is used to construct it from a *single* function residing in the simulation cell while still satisfying periodicity. This is illustrated in the bottom panel of Fig. 12.

A detailed description of the form of $\bar{f}(\mathbf{G})$ required to accomplish the above, and of the treatment of the $\mathbf{G} = 0$ term, can be found in Ref. 211 and in the original paper by Martyna and Tuckerman.²¹⁴ The modified Hartree potential $\bar{V}_{\text{H}}(\mathbf{G})$ and the modified local pseudopotential $\bar{V}_{\text{locps}}(\mathbf{G})$ are obtained according to

Eqs. (22)–(25) of Ref. 211. The core–core interaction is evaluated using a direct Coulombic sum rather than via Ewald summation.

An important consequence of the MIC approach is that the length of the simulation cell must be at least twice that of the molecule in any dimension; otherwise, unphysical interactions occur between the molecule and the nearest periodic image.

c. Direct real-space calculations using a multigrid solver. In this approach, FFTs are not used in the evaluation of the Hartree, local pseudopotential, and core–core energy terms. Instead, these calculations are performed in real space.

The Hartree potential under OBCs is obtained using a multigrid solver⁵⁵ (also see Sec. IV D 1) to solve the Poisson equation,

$$\nabla^2 V_{\text{H}}(\mathbf{r}) = -4\pi n(\mathbf{r}), \quad (51)$$

on a real-space grid. Defect correction²¹⁶ is used to improve the accuracy of the second-order solution to a higher order (up to 12th).

Dirichlet boundary conditions of the form

$$V_H(\mathbf{r}) = \int_{\Omega} \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (52)$$

must be provided on the faces of the simulation cell $\partial\Omega$. To mitigate the computational cost of evaluating Eq. (52), charge coarse-graining and bilinear interpolation are used.

The local pseudopotential term is obtained directly in real space according to Eq. (B4) of Ref. 216. As in our implementation of the Martyna–Tuckerman method, the core–core interaction is evaluated using a direct Coulombic sum rather than via Ewald summation.

One important limitation of this technique is that the multigrid solver used in ONETEP only supports orthorhombic (cuboid) grids.

E. Excitations and spectroscopy

1. Conduction NGWFs and band structure calculations

As mentioned in Sec. II, while the NGWF basis is optimized to accurately represent the occupied states, the unoccupied states are typically over-estimated in energy and are in some cases missing.^{217,218} As described in Ref. 218, a second set of ‘conduction NGWFs’ may therefore be optimized to represent a select number of low energy conduction states using a projection operator-based approach. Aside from being non-self-consistent, the calculation proceeds along similar lines to a valence ONETEP calculation. The optimized conduction NGWFs are then combined with the valence NGWFs to form a joint NGWF basis. Provided that one uses sufficiently large conduction NGWF radii (typically larger than valence NGWF radii), it is possible to generate a joint NGWF basis that is capable of accurately representing both the occupied states and a given number of (bound) unoccupied states.^{218,219} One can use this basis to calculate optical absorption spectra using Fermi’s golden rule,^{218,219} which has been employed to study energy transport in the Fenna–Matthews–Olson light-harvesting pigment–protein complex,^{220,221} or TDDFT^{77,222} (Sec. IV E 2), or to calculate electron energy loss spectra²²³ (EELS) (Sec. IV E 5). Alternatively, the joint NGWF basis may be directly used to calculate band structures.

As described in Ref. 224, two possible approaches to calculating band structure are implemented in ONETEP. The first approach, denoted the tight-binding (TB) style approach, is related to both TB and Wannier interpolation, while the second approach, denoted the $\mathbf{k} \cdot \mathbf{p}$ style approach, is related to $\mathbf{k} \cdot \mathbf{p}$ perturbation theory. In the TB approach, \mathbf{k} -dependent phase factors are directly built into the construction of the Hamiltonian *matrix* at a given \mathbf{k} -point, while in the $\mathbf{k} \cdot \mathbf{p}$ approach, a \mathbf{k} -dependent Hamiltonian *operator* is derived, with the Hamiltonian matrix being the NGWF matrix elements of this operator.²²⁵ For both approaches, the Hamiltonian is constructed in the joint basis at each of the requested \mathbf{k} -points and diagonalized to get the corresponding energies. A method for unfolding supercell band structures has also been implemented.

For a well converged NGWF basis, the two methods give very similar band structures. However in cases where the NGWFs are not well converged, e.g., due to too small localization radii or too loose convergence criteria, the $\mathbf{k} \cdot \mathbf{p}$ approach tends to result in unphysical gaps at the Brillouin zone boundary with corresponding discontinuities in the unfolded band structure.²²⁴ It is therefore generally recommended to use the TB style approach. As an example, Fig. 13

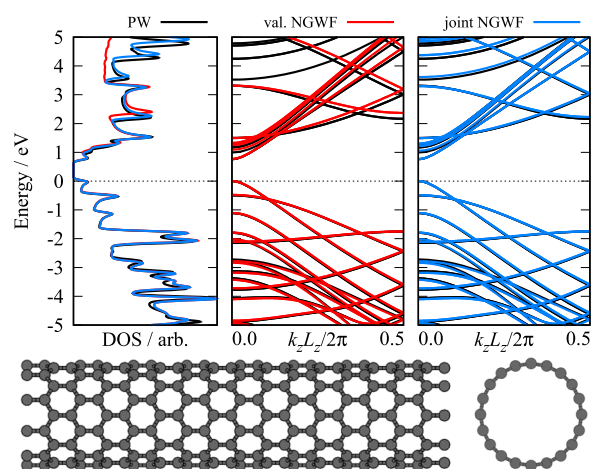


FIG. 13. Density of states (DoS) and band structure for the depicted (10,0) carbon nanotube, calculated using the CASTEP plane wave (PW) code⁴⁹ and in ONETEP using both valence NGWFs only and a joint NGWF basis. The tight-binding style approach was used for ONETEP calculations. The PW band structure was calculated in the unit cell with a Monkhorst–Pack \mathbf{k} -point sampling of $1 \times 1 \times 8$, while the ONETEP calculations were performed for an 11 unit supercell and unfolded. All calculations employed the PBE exchange–correlation functional and a kinetic energy cutoff of 916 eV. Four (five) valence (conduction) NGWFs were used per C atom, with radii of 7 bohrs (11 bohrs). A Gaussian smearing of 0.05 eV has been applied to the DoS, while the curves have been shifted so that the highest occupied molecular orbital is at zero. Figure adapted with permission from Ratcliff *et al.*, Phys. Rev. B **98**, 125123 (2018). Copyright 2018 American Physical Society.

shows the calculated band structure for a carbon nanotube (CNT) using the TB style approach. Although for this system the low energy conduction states are already relatively well represented, it is necessary to use a joint basis in order to accurately reproduce the band structure over the energy range considered.

2. Linear-response time-dependent DFT

Many potential applications of linear-scaling electronic structure methods involve the prediction of optical properties of large-scale systems, ranging from nanostructured materials to pigment–protein complexes. Often, as in the prediction of UV–vis absorption spectra, one is mainly interested in a small number of low-lying excited states. These states can be efficiently obtained in the framework of linear-response time-dependent DFT (LR-TDDFT), where the problem of calculating excitation energies is recast into solving for the lowest eigenvalues of an effective two-particle Hamiltonian.^{226,227} In ONETEP, the resulting non-Hermitian eigenvalue problem is solved by exploiting an equivalence of the problem to that of a set of classical harmonic oscillators.²²⁸

To achieve linear-scaling computations of excitation energies, in ONETEP, the transition density matrix $\rho^{(1)}(\mathbf{r}, \mathbf{r}')$ describing the lowest excitation of the system is expanded in terms of valence and conduction NGWFs optimized in a ground-state DFT calculation

$$\rho^{(1)}(\mathbf{r}, \mathbf{r}') = \chi_{\alpha}(\mathbf{r}) K^{\{X\}\alpha\beta} \phi_{\beta}(\mathbf{r}') + \chi_{\alpha}(\mathbf{r}) K^{\{Y\}\alpha\beta} \phi_{\beta}(\mathbf{r}'), \quad (53)$$

where $\mathbf{K}^{\{X\}}$ and $\mathbf{K}^{\{Y\}}$ are the excitation and de-excitation contributions to the transition density matrix, respectively. The lowest

excitation energy of the system can then be written as a functional of $\mathbf{K}^{(X)}$ and $\mathbf{K}^{(Y)}$, and higher excitations can be computed by enforcing an effective orthogonality constraint between the transition density matrices of individual excitations. Both the full LR-TDDFT method and the simplified Tamm–Dancoff approximation,²²⁹ corresponding to setting $\mathbf{K}^{(Y)}$ to zero, are implemented in the ONETEP code.^{77,222}

Linear-scaling computational effort with system size for calculating a fixed number of low-lying excited states is then achieved by exploiting sparsity in $\mathbf{K}^{(X)}$ and $\mathbf{K}^{(Y)}$. However, unlike for a ground-state DFT calculation, where the sparsity of the density kernel is defined by a simple cutoff radius, the sparsity patterns of $\mathbf{K}^{(X)}$ and $\mathbf{K}^{(Y)}$ can be chosen in a physically motivated way to restrict the space of possible solutions to the TDDFT eigenvalue problem. By dividing the system into smaller subsystems, and setting $K^{(X)\alpha\beta}$ and $K^{(Y)\alpha\beta}$ to zero when NGWFs $\chi_\alpha(\mathbf{r})$ and $\phi_\beta(\mathbf{r})$ are associated with different subsystems, it is possible to suppress all contributions to $\rho^{(1)}(\mathbf{r}, \mathbf{r}')$ arising from charge transfer between subsystems^{230,231} (see Fig. 14). Since TDDFT systematically underestimates excitation energies of charge-transfer states,²³² especially when using semi-local functionals, the subsystem approach is an efficient way to remove unphysical low-lying excited states. The approach has been used to study excitations of dyes in explicit solvent clusters^{231,233,234} and the computation of excitons in multi-chromophore pigment–protein complexes.²³⁰

a. Excited state forces. The ability to calculate excited state forces in ONETEP is not yet available in ONETEP v5.2, but it will be released in a future version. The implementation of excited state forces in ONETEP takes the approach of Furche and Ahlrichs,²³⁵ solving a Z-vector equation

$$\mathbf{F}_{\text{TDDFT}} = \frac{d\omega_{\text{min}}}{d\mathbf{R}_y} = \mathbf{F}_{\text{direct}} + \mathbf{F}_Z + \mathbf{F}_{\text{Pulay}}, \quad (54)$$

where F_{direct} is the partial derivative of the TDDFT excitation energy with respect to atomic positions and F_{Pulay} is the excited state equivalent of the ground state Pulay term (see Sec. IV C 1). F_Z is a Lagrange multiplier compensation force contribution, required because the KS orbitals of the system will change with the atomic positions but are also constrained to be orthonormal.

The Z-vector term is calculated by solving a system of linear equations.²³⁵ We map into a spherical wave basis before projecting out valence contributions because the valence and conduction NGWFs in ONETEP are not optimized to represent the Lagrange multiplier term. As with PDOS calculations (Sec. IV E 4), using the spherical wave basis introduces very little error, but using an alternative basis optimized to represent the Z-vector is the subject of active development.

3. Electronic transport

The continued miniaturization of electrical devices has led to an increasing interest in modeling charge transport at the atomic scale. Modeling electronic transport using first-principles quantum mechanics is challenging, however, as experimentally relevant systems typically require a simulation of not only the device but also substantial amounts of bulk material to model the semi-infinite leads that contact and connect with the device. Such calculations, therefore, generally require the consideration of hundreds if not thousands of atoms. ONETEP is therefore an ideal platform for calculating electronic transport. The optimized NGWFs provide a near-minimal yet highly accurate localized basis set, resulting in an efficient but unbiased calculation of electronic transport on a first-principles quantum-mechanical footing.

We have implemented within ONETEP methods to calculate zero-bias electronic transport, within the Landauer–Büttiker formalism.^{236–238} We refer the reader to Ref. 239 for the details; in short, however, the calculation involves embedding the desired device geometry within a (potentially larger) “auxiliary” ground-state DFT

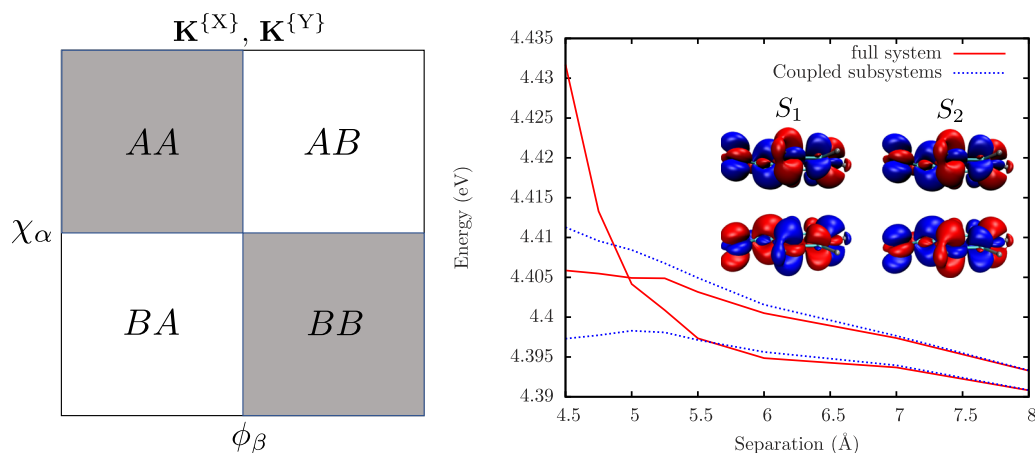


FIG. 14. Subsystem TDDFT approach in ONETEP as illustrated by the distance dependence of exciton splitting in the benzaldehyde dimer. The off-diagonal blocks of $\mathbf{K}^{(X)}$ and $\mathbf{K}^{(Y)}$, corresponding to the conduction and valence NGWFs centered on different molecules, are set to zero, suppressing all unphysical charge-transfer excitations between molecule A and molecule B. From Zuehlsdorff, *Computing the Optical Properties of Large Systems*. Copyright 2015 Springer International Publishing. Adapted with permission from Springer Nature.

calculation, then extracting the Hamiltonian and overlap matrix elements (within the NGWF basis) to calculate the electronic transport coefficients as a post-processing step. The implementation exploits the sparsity of the Hamiltonian and overlap matrices and uses an optimized block-tri-diagonal matrix inversion algorithm to compute the exact transmission spectrum.^{240,241} For quasi-one-dimensional systems, the computational complexity scales linearly with the number of atoms; for systems with higher dimensions, the algorithm performs considerably better than a naïve algorithm using dense matrix inversions. Typically, calculating the transport properties of a system corresponds to only a few percent of the total computational time associated with the electronic structure calculation.

The functionality we have implemented (for both spin-restricted and spin-polarized DFT calculations) includes electronic transmission between an arbitrary number of semi-infinite leads; bulk transmission and band structures within the leads; density of states for the device and leads; and eigenchannel decompositions, defined as the eigenvectors of the transmission probability matrix,²⁴² for transmission between arbitrary pairs of leads. Transmission eigenchannels are particularly useful for interpreting transport spectra as they provide a spatial representation of the non-mixing scattering states within the system.

4. Local and angular-momentum projected density of states

ONETEP has an extensive set of density of states (DOS) methods. Plots may be made directly from ONETEP or by outputting OPTADOS²⁴³ compatible data-files. A raw DOS may be calculated from the eigenvalues of the Hamiltonian matrix

$$\rho(\epsilon) = \sum_i \delta(\epsilon - \epsilon_i). \quad (55)$$

In practice, the delta function is replaced with a Gaussian with broadening σ , typically of the order of 0.1 eV. The local density of states (LDOS) is the projection of the DOS onto atom-local functions. The LDOS in a given region I is calculated in ONETEP by projecting each eigenstate onto the local orbitals of region I as

$$\begin{aligned} \rho_I(\epsilon) &= \sum_i \delta(\epsilon - \epsilon_i) \sum_{\beta, \gamma, \alpha \in I} (M_i^\dagger)^\gamma \langle \phi_\gamma | (|\phi^\alpha\rangle \langle \phi_\alpha|) | \phi_\beta \rangle M_i^\beta \\ &= \sum_i \delta(\epsilon - \epsilon_i) \sum_{\alpha \in I} (M_i^\dagger)^\alpha \left(\sum_\beta S_{\alpha\beta} M_i^\beta \right). \end{aligned} \quad (56)$$

This provides a series of functions $\rho_I(\epsilon)$ for each of the chosen regions I , which may include only the NGWFs of a single atom or those of a group of atom types.

ONETEP can also project the DOS onto angular momentum resolved functions to produce a projected DOS (PDOS). To calculate the PDOS, an additional resolution of the identity is inserted into Eq. (56) using a basis of angular momentum resolved functions $|\chi_{l,m}\rangle$ onto which the NGWFs are projected,

$$\rho_{I,l}(\epsilon) \approx \sum_i \delta(\epsilon - \epsilon_i) \sum_{\substack{\{\alpha,l,m\} \in I, \\ \beta,l',m'}} M_i^{\dagger\alpha} \langle \phi_\alpha | \chi'_{alm} \rangle \Lambda^{alm,\beta'l'm'} \langle \chi'_{\beta'l'm'} | \phi_\beta \rangle M_i^\beta, \quad (57)$$

where we need to include the inverse overlap matrix of angular momentum resolved functions, Λ , since this basis is also non-orthogonal.

There is a considerable amount of choice in the basis used for the angular momentum resolved functions, which are effectively a set of spherical harmonics multiplied by some radial term. In ONETEP, we have implemented two options for the radial term: spherical waves and pseudo-atomic orbitals (PAOs). The PAOs are the same as are used to initialize the NGWFs before optimization. For more details about the theory behind these options as well as tests and comparisons, refer to Ref. 244.

In using an angular momentum resolved basis, which is non-optimal for representing the orbitals, we inevitably lose something in the projection from NGWFs. This loss can be quantified through a spilling parameter,²⁴⁵ which is calculated after each PDOS run,

$$s = \frac{1}{N_{MO}} \sum_i \left\{ 1 - \sum_{alm} W_{alm,i} \right\}, \quad (58)$$

where $W_{alm,i}$ is the PDOS weight matrix,

$$W_{alm,i} = \sum_{\beta,l',m'} M_i^{\dagger\alpha} \langle \phi_\alpha | \chi'_{alm} \rangle \Lambda^{alm,\beta'l'm'} \langle \chi'_{\beta'l'm'} | \phi_\beta \rangle M_i^\beta. \quad (59)$$

We have found that PAOs are sufficient for almost all applications with less than $\approx 2\%$ spilling and qualitatively correct PDOS profiles. The PAO basis is computationally fast and memory efficient; the spherical wave basis may be used to reduce the spilling to almost zero, but at the expense of more memory usage and slower evaluation.

The PDOS implementation in ONETEP can also produce bandcenters, which are used extensively as descriptors for catalytic activity.

5. Core loss spectroscopy

Experimental characterization of nanomaterials relies on a wide range of imaging and spectroscopy techniques. Prominent among these are techniques based on core loss: the concept that incident energy in the form of electrons [in electron energy loss spectroscopy (EELS)] or x rays [in X-ray absorption spectroscopy (XAS)] can cause characteristic electronic transitions from the occupied core levels to unoccupied conduction band levels. Given sufficiently high resolution in energy and space, these methods can even allow the local structure of the conduction band to be probed and elucidate properties such as coordination, bonding, and surface reconstruction. However, such experimental results are not always easy to interpret, and simulations capable of a realistic treatment of the material are highly complementary to experiment in distinguishing between possible structures. Simulation of EELS and XAS relies on the Fermi golden rule to give the imaginary dielectric function via dipole matrix elements,

$$\epsilon_2(\omega) = \frac{1}{\Omega} \sum_c \sum_i |\mathbf{q} \cdot \langle \psi_i | \mathbf{r} | \psi_c \rangle|^2 \delta(E_i - E_c - \omega). \quad (60)$$

ONETEP implements the calculation of these expressions as a post-processing tool²²³ capability, linked to the use of the projector augmented wave method²⁰ (Sec. IV A 1). Care is taken to ensure robust evaluation of matrix elements of the form $\langle \psi_i | \mathbf{r} | \psi_c \rangle$, as detailed in

Ref. 246. Post-processing of the optical matrix elements to calculate EELS and XAS spectra can be carried out through an interface to the OPTADOS package.²⁴⁷

6. Anharmonic vibrational spectra and IR spectra

Density functional theory molecular dynamics (DFT-MD) provides an efficient framework for accurately computing several kinds of spectra, e.g., vibrational, infra-red (IR), and Raman, since it gives direct access to quantum time correlation functions of various dynamical variables from which spectra can be computed via time-dependent perturbation theory (or equivalently Green–Kubo formalism).²⁴⁸ The major advantage of DFT-MD approaches lies in its capability of inherently describing effects of polarization, temperature, and anharmonicity of the potential energy surface (PES) without having to rely on *a posteriori* corrections. In particular, DFT-MD in ONETEP (Sec. IV C 3) can provide a valuable tool for the interpretation of vibrational and IR signatures of large biological systems in both the gas-phase and the condensed-phase.

It can be shown that within linear response theory,²⁴⁸ the spectrum associated with a given perturbation, such as the spectrum associated with infrared radiation (IR), can be obtained from the Fourier transform of an autocorrelation function of the dynamical variable conjugate to the external field (in the case of IR spectra, this is the total dipole moment). This can be easily computed from a MD calculation of the unperturbed system via the fluctuation–dissipation theorem.²⁴⁹ For instance, it can be shown²⁰¹ that in the case of IR spectra, what needs to be computed is

$$I(\omega) = \frac{2\pi\beta\omega^2}{3cV} \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \tilde{\mu}(0), \dot{\mu}(t) \rangle, \quad (61)$$

where ω is the frequency of the incoming radiation, t is time, $\beta = (k_B T)^{-1}$, c is the speed of light, V is the volume of the sample, and $\dot{\mu}(t)$ is the total dipole moment operator. The integrand in Eq. (61) is the so-called Kubo autocorrelation function

$$C^{\text{kubo}}(t) = \langle \tilde{\mu}(0), \dot{\mu}(t) \rangle \equiv \left\langle \frac{1}{\beta} \int_0^\beta d\lambda \dot{\mu}(0) \dot{\mu}(t + i\hbar\lambda) \right\rangle. \quad (62)$$

$C^{\text{kubo}}(t)$ is a real function, and at high temperature ($\beta \rightarrow 0$), it approaches the classical autocorrelation function $C^{\text{cl}}(t) \equiv \langle \mu(0) \mu(t) \rangle$, where the quantum operators are replaced by their classical vectorial quantities $\dot{\mu} \rightarrow \mu$. In the case of vibrational spectra, one has a similar formula to the one in Eq. (61) with a different prefactor and using the autocorrelation function of the atomic velocities. As mentioned in Sec. IV C 2, ONETEP does not take into account the effect of LO–TO splitting.

In ONETEP, it is also possible to compute the IR spectrum of a subsystem (e.g., a solute molecule) within a larger system (e.g., a solvent). To compute the electronic dipole moment of a non-isolated system A within a larger environment B , a partitioning scheme is required. In the ONETEP framework, NGWFs are easily assigned to a given subsystem, as they are centered on atoms. However, non-orthogonality means that the density matrix ρ_A associated with subsystem A cannot be built from NGWFs centered on atoms belonging to A only. One way to overcome this obstacle, i.e., distributed multipole analysis, is described in Sec. IV F 1. A simpler method, which can be performed on-the-fly and offers results similar to DMA for IR, is given by the density kernel partitioning (DKP).²⁰¹ In essence,

the DKP method works by including all contributions from NGWFs centered on atoms in subsystem A and half of the contributions from NGWFs centered on atoms in B , which have a non-zero overlap with NGWFs centered on atoms in A .

Examples of IR spectra for simple molecules and small peptides in both the gas phase (no partitioning required) and the aqueous phase can be found in Ref. 201.

F. Partitioning the density/energy

1. Distributed multipole analysis

Distributed multipole analysis^{199,200,250} (DMA) is a technique for representing a continuous charge distribution in terms of a set of point multipoles. Atomic centers are usually, although not universally, used as the multipole sites. While DMA is typically performed in a Gaussian basis set,^{251,252} ONETEP implements it using the localized NGWF basis following a density-fitting approach.

In the first step, the electronic density is decomposed into two-center contributions from overlapping NGWFs. Each contribution is subsequently expanded in terms of auxiliary basis functions—truncated spherical waves (SWs) originating on both contributing atoms. The electrostatic (Coulomb) metric is used in the expansion, which corresponds to minimizing the self-interaction energy of the density difference between the original contribution and its SW expansion. As a result, the electronic density is decomposed into a sum of atom-centered contributions, each of which is a linear combination of SWs. Finally, spherical multipoles centered on atomic sites are obtained as linear combinations of SW expansion coefficients and simple radial integrals involving Bessel functions that can be calculated analytically.

The maximum angular momentum ℓ used in the SW basis determines the order of the expansion. ONETEP's implementation supports orders up to hexadecapole ($\ell = 4$), but values printed from the DMA procedure are restricted to monopoles, dipoles, and quadrupoles. A reference point can be specified for the multipoles, and in addition, atom-centered multipoles can be printed out, both electronic-only and total (valence-electronic + core). Additionally, ONETEP can output total multipoles in a commonly used format compatible with the GDMA²⁵¹ program. For more details, refer to Ref. 202 (Sec. II.D.2).

2. Energy decomposition analysis

Molecular interactions arise from the balance of many physically unique components such as electrostatics, polarization, and charge transfer. It is the aim of energy decomposition analysis (EDA) to partition the molecular interaction energy into these individual components.²⁵³ The ONETEP EDA scheme²⁵⁴ is a hybrid approach based on the absolutely localized molecular orbital (ALMO) method of Khaliullin *et al.*²⁵⁵ and the localized molecular orbital EDA of Su and Li.²⁵⁶ Specifically, our approach separates the interaction energy into electrostatic, exchange, correlation, Pauli repulsion, polarization, and charge transfer terms, which allows full analysis of the interaction energy. A notable benefit of our method is that the polarization and charge transfer basis set dependence problem seen in the original ALMO EDA approach is remedied through exploiting the strictly localized property of the ONETEP orbitals.²⁵⁴

The ONETEP EDA implementation features a number of modifications that make it highly suitable for the analysis of large-scale

systems. This scalability is especially important where long-range interactions can dominate, e.g., electrostatics in biomolecular assemblies. This capability has been demonstrated through performing calculations with complete basis set accuracy on an entire thrombin protein–inhibitor complex comprising 4975 atoms.²⁵⁴

The method is also capable of generating powerful electron density difference (EDD) plots²⁵⁴ that show electron density redistributions occurring on molecular polarization and charge transfer. These plots reveal key orbital features such as acceptor–donor pairs and have a high potential value for the pharmaceutical industry through suggesting regions on which to focus optimization efforts. EDA has been applied in combination with EDD plots to small, drug-like fragment interactions with the S1 pocket of the thrombin protein,²⁵⁷ highlighting a number of important interactions for thrombin–drug binding. Here, the analysis indicated the thrombin salt bridge interaction with charged amidine functional groups of drug molecules as being a key target for drug optimization efforts. This application of the ONETEP EDA to a pharmaceutical problem demonstrates the high level of insight that can be provided through visual EDD plots and quantitative EDA energies.

3. Population analysis and classical force field derivation

In order to enhance the information content of the ground state electron density, which can be particularly crucial for the large system sizes that are commonly used in ONETEP, various methods of electronic population analysis are available. Natural population analysis (NPA) and the associated natural bond orbitals (NBOs) are commonly used in quantum chemistry in conjunction with atom-centered basis sets to provide a chemical picture of bonding in terms of localized Lewis-type bond and lone pair orbitals.^{258,259} The density matrix formalism, and atom-centered NGWF basis, lends itself naturally to this scheme; the transformation of the ground state single particle density matrix to natural atomic orbitals (NAOs), and hence NPA charges, is implemented in the ONETEP code.²⁶⁰ This functionality has allowed us to quantify metal–oxygen charge transfer in very large (1000+) atom models of the oxygenated myoglobin metalloprotein, even taking dynamical mean field theory effects into account at the metal center (see Sec. IV B 5).¹⁴⁶

Through an interface with the NBO 5 analysis package,²⁵⁹ the set of atom-centered NAOs may be further transformed, via natural hybrid orbitals, into NBOs for each pair of atoms in the system.²⁶⁰ The populated NBOs form the chemically intuitive, natural Lewis structure of the molecule (bond and lone pair orbitals) and are complemented by the unoccupied antibonding and Rydberg orbitals. This analysis allows us to build up chemically familiar pictures of the nature of bonding in complex systems. For example, we were able to reconcile ligand–metal back charge transfer processes in the aforementioned myoglobin complex with experimental L-edge x-ray absorption spectroscopy by measuring metal *d* orbital occupancies.¹⁴⁶ An additional useful feature of the NBO analysis is that energetic lowering due to electronic delocalization to anti-bonding orbitals, which represents deviation from ideal Lewis description, may be estimated via second order perturbation theory. For example, we have used the relationship between this variational lowering of the energy and hydrogen bond strengths to estimate the contributions of various active site amino acids in the enzyme, chorismate mutase, to its catalytic rate enhancement.²⁶¹

A second technique commonly employed for the study of the chemical nature of molecules and materials is atomic partitioning of the electronic density among the constituent atoms in the system. In general, the partitioned atomic electron density is given by

$$n_A(\mathbf{r}) = \frac{w_A(\mathbf{r})}{\sum_B w_B(\mathbf{r})} n(\mathbf{r}), \quad (63)$$

where $n(\mathbf{r})$ is the total electron density at position \mathbf{r} and $w_A(\mathbf{r})$ is a weighting factor. A and B label atoms. There is no unique weighting scheme to perform this partitioning, but common methods include Hirshfeld, iterative Hirshfeld (IH),²⁶² and iterated stockholder atoms (ISA).²⁶³ Although Hirshfeld and ISA analyses are both available in ONETEP, our particular focus has been on density derived electrostatic and chemical (DDEC) analysis, which was first developed by Manz and Sholl,²⁶⁴ and has undergone continual development in recent years.^{265,266} We refer the reader to Refs. 264–266 for a full description, but, in brief, the DDEC method aims to optimize atomic weighting factors $w_A(\mathbf{r})$ such that the derived atomic electron densities yield net atomic charges (obtained by integrating the atomic electron densities over all space) that are chemically meaningful while reproducing the electrostatic potential (ESP) surrounding the molecule or material. Earlier versions of DDEC (up to DDEC/c3) are implemented in ONETEP,^{267,268} while the latest versions may be accessed through an interface between ONETEP and the CHARGEMOL program.²⁶⁶

DDEC charges have been computed for the 1000+ atom model of the myoglobin metalloprotein described above, and charge transfer to the oxygen molecule has been shown to agree well with the chemically intuitive value of $-1e$.²⁶⁸ Furthermore, analysis of DDEC charges of several small proteins reveals a number of desirable properties. Not only do the charges reproduce the ESP of the underlying ONETEP calculation well, but they are still accurate when transferred to conformations close to the native state of the protein, even for buried atoms (a feature that traditional charge derivation schemes struggle with).²⁶⁷ These features make DDEC electron density partitioning well suited to molecular mechanics force field design, uniquely providing a route between static large-scale electronic structure calculations and long time scale dynamics of biological molecules. To facilitate this, we have additionally implemented methods to derive off-site charges in ONETEP to effectively model electron anisotropy in force fields (such as lone pairs and σ -holes)²⁶⁹ as well as dispersion (or van der Waals) C_6 coefficients using both the widely used Tkatchenko–Scheffler method^{270,271} and a new method to rigorously account for environmental screening.²⁷²

Through an interface between ONETEP and the QUBEKIT software,²⁶⁹ molecular mechanics force fields (named QUBE) for both small molecules and entire proteins may be readily computed using the electron density partitioning methods described above. For example, QUBE force fields have been derived for the alanine pentapeptide and two proteins, ubiquitin and GB3. Microsecond molecular dynamics simulations using the derived force fields are in good agreement with experiment [as measured by comparisons with nuclear magnetic resonance (NMR) J coupling data].²⁷³ Furthermore, the QUBE force field outperforms the established Optimized Potentials for Liquid Simulations (OPLS) force field in the prediction of binding free energies of six benzene analogs to the L99A mutant of T4 lysozyme (mean unsigned error of 0.85 kcal/mol).²⁷⁴ With several avenues for further improvement, there is hence

significant potential for force fields derived from large-scale electronic structure calculations in the simulation of biomolecular dynamics and the design of future medicines.^{275,276}

4. Electron localization descriptors

Electron localization descriptors provide a quantum valence shell electron pair repulsion (VSEPR) representation of molecules, crystalline solids, surfaces and coordination compounds. They are a useful visualization tool for indicating regions of lone pairs and distinguishing between bond orders. The electron localization function (ELF) was first introduced by Edgecombe and Becke in 1990 for Hartree–Fock theory and later extended to DFT by Becke *et al.*^{277–279} It is based on the conditional probability of an electron existing in the proximity of a like-spin reference electron, derived from the Hartree–Fock probability of finding two same-spin particles at two positions. As this quantity does not have an upper bound, it is inverted for more convenient graphical interpretation such that the localization of the reference electron is represented at unity. This visualization tool is applicable to large complex systems using the linear-scaling capabilities of ONETEP, particularly for biomolecular simulations, catalysis, and the design of nanostructured materials.

The exact kinetic energy density as defined in ONETEP, $\tau_\sigma^{\text{exact}}$, is extended into a *non-negative* probability density, the Pauli kinetic energy density, to describe electron localization,

$$\tau_\sigma^P = \tau_\sigma^{\text{exact}} - \frac{1}{4} \frac{(\nabla n_\sigma(\mathbf{r}))^2}{n_\sigma(\mathbf{r})}, \quad (64)$$

where

$$\tau_\sigma^{\text{exact}}(\mathbf{r}) = \sum_\alpha \tau(\mathbf{r}; \alpha), \text{ and } \tau(\mathbf{r}; \alpha) = (\nabla \phi_\alpha(\mathbf{r})) \cdot \left(\nabla \sum_\beta K^{\alpha\beta} \phi_\beta(\mathbf{r}) \right), \quad (65)$$

where β runs over NGWFs that overlap with ϕ_α . $n_\sigma(\mathbf{r})$ is the charge density for each spin σ , and $\nabla n_\sigma(\mathbf{r})$ is its gradient. To make τ_σ^P dimensionless, it is compared with the kinetic energy density of the uniform electron gas as a reference, i.e., the local spin density approximation (LSDA), $\tau_\sigma^{\text{LSDA}}(\mathbf{r}) = \frac{3}{5} (6\pi^2)^{\frac{2}{3}} n_\sigma^{\frac{5}{3}}(\mathbf{r})$. This is reformulated to avoid open bounds, limiting the ELF to a more desirable finite range of values from zero to one,

$$\text{ELF}_\sigma = \frac{1}{1 + \chi_\sigma^2} \text{ where } \chi_\sigma = \frac{\tau_\sigma^P(\mathbf{r})}{\tau_\sigma^{\text{LSDA}}(\mathbf{r})}. \quad (66)$$

The code has also been extended to include the simpler localized orbital locator (LOL). Here, only the exact kinetic energy is required instead of τ_σ^P , again with a range of zero to one,

$$\text{LOL}_\sigma = \frac{1}{1 + \frac{\tau_\sigma^{\text{exact}}(\mathbf{r})}{\tau_\sigma^{\text{LSDA}}(\mathbf{r})}}. \quad (67)$$

V. APPLICATIONS OF ONETEP: EXAMPLES

Primarily due to its favorable scaling with system size, ONETEP has found widespread application to studies of nanomaterials,

particularly nanocrystals,^{280,281} nanorods,^{282,283} and 2D materials.^{92–94} In all these cases, the combination of scaling to models of hundreds or thousands of atoms with plane-wave accuracy has allowed a quantitative prediction of optical and electronic properties. In this section, we present several applications of ONETEP to such systems as well as other problems of interest, spanning a wide range of topics.

A. Conductance between carbon nanotubes

The electronic transport functionality of ONETEP (Sec. IV E 3) has been used to carry out calculations of conductance between two semi-infinite carbon nanotubes (CNTs): a fullerene-capped (5, 5) CNT and a zig-zag terminated (9, 0) CNT with two hydrogen atoms passivating dangling bonds on the terminal carbon atoms. Figure 15 shows the geometry of the auxiliary system for which the ground-state DFT calculation is performed. The device, indicated by the atoms within the box, consists of a left and right lead and a central region where electrons can tunnel between the leads. Outside the device, a number of buffer atoms are included in the ground-state electronic structure calculation only to ensure that the matrix elements at the extreme left/right of the device region are bulk-like, i.e., they are not directly influenced by the presence of the terminated ends of the CNTs.

Figure 16 shows the conductance spectra between the two CNTs, which are generally found to be two orders of magnitude lower than the maximum possible conductance for CNTs ($4e^2/h$). This is because the two CNTs have different chirality, suppressing inter-tube conductance.²⁸⁴ Three sharp peaks are observed at -0.5 eV, -0.2 eV, and 0.16 eV and can be explained by considering the eigenchannel decompositions. Inset (a) shows the eigenchannel with largest conductance close to a resonance. It can be seen that a significant eigenchannel weight has transferred across the junction to the second CNT. The sharp peaks suggest the presence of localized states. Indeed, by considering the (9, 0) CNT from the original auxiliary geometry in isolation and examining the eigenstates, a localized end-state is discovered at approximately the same energy as the conductance resonance. Inset (b) shows this end state, demonstrating that it has the same form as the eigenchannel in inset (a) and is responsible for the resonance in the conductance. The asymmetry in the line shape of the resonances is typical of Fano resonant scattering between bulk states and localized defect states.²⁸⁵ In contrast, inset (c) shows the same eigenchannel as shown in inset (a), but away from a resonance where the conductance is low, and

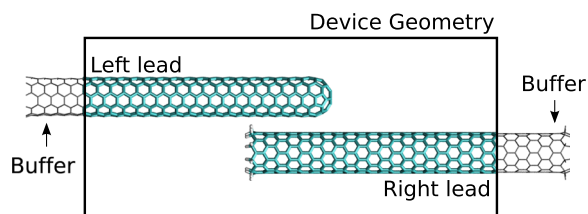


FIG. 15. The geometry used to calculate conductance between two semi-infinite carbon nanotubes. From Bell, *First-Principles Conductance between Carbon Nanotubes*. Copyright 2015 Springer International Publishing. Figure adapted with permission from Springer Nature.²⁸⁶

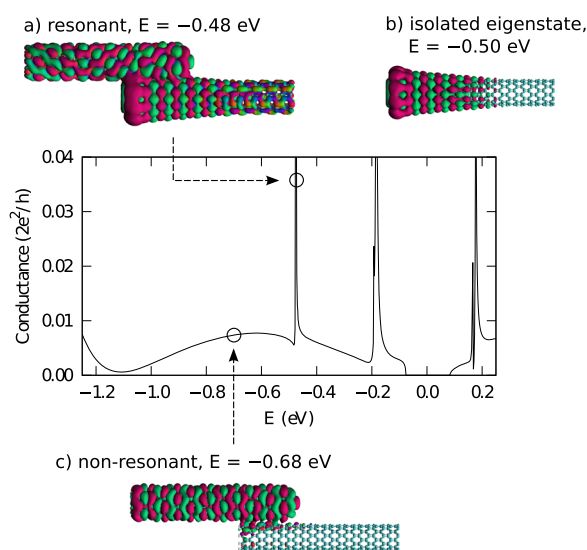


FIG. 16. (Main figure) The conductance spectrum between the semi-infinite CNTs in the geometry shown in Fig. 15. Insets: (a) an eigenchannel at a resonance; (b) a localized end state of the isolated (9, 0) fragment; and (c) an eigenchannel away from resonance. Each eigenchannel/eigenstate is colored by the complex phase, and eigenchannels originate from deep within the left CNT. From Bell, *First-Principles Conductance between Carbon Nanotubes*. Copyright 2015 Springer International Publishing. Figure adapted with permission from Springer Nature.²⁸⁶

the eigenchannel is seen to be almost entirely localized within the source CNT.

B. Supported nanoparticles in heterogeneous catalysis

Metal oxide supports often play an active part in heterogeneous catalysis by moderating both the structure and the electronic properties of the metallic catalyst particle. In order to provide some fundamental understanding on these effects, ONETEP has been used in collaboration with Johnson Matthey for an investigation of the binding of O and CO on Pt nanoparticles supported on titania (anatase) surfaces²⁸⁷ (see Fig. 17). They have found that there is strong interaction of the Pt nanoparticles with the support so it is important to include both in the simulations: simpler models where either isolated nanoparticles are used or slab surfaces are used instead of nanoparticles cannot capture these crucial effects. Furthermore, they were able to separate geometric and electronic effects of the support and found that a larger contribution to ligand binding energy arose from the former. They also made use of electronic descriptors, including the d-band center and an electronic density based descriptor in order to further rationalize their results in these complex systems. This pioneering work provided an approach for creating realistic models for the study of supported metal catalysts, which are ubiquitous in many industrial applications ranging from batteries and fuel cells to the bulk synthesis of industrial chemicals.

C. Piezochromic properties of nanocrystals

ONETEP has been used to study the effect of pressure-induced deformations on the optoelectronic properties of CdS nanocrystals

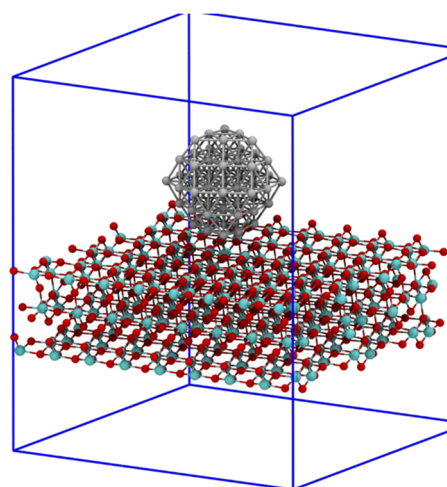


FIG. 17. An example of a structure used to investigate the support effects on a Pt nanoparticle supported by a titania surface.

coated with organic ligands.²⁸¹ The application of pressure to these finite systems employs the electronic enthalpy method,^{183,288} and optical absorption spectra were calculated using the linear-response implementation of time-dependent DFT in ONETEP (Sec. IV E 2). Whereas in the bulk, the application of pressure increases the bandgap, for these nanocrystals, the sign of the shift depends upon the choice of ligand, as shown in Fig. 18. Hole-accepting ligands such as phenyl groups hybridize strongly with the HOMO causing it to be localized at the surface, in contrast to the LUMO that is confined within the core, leading to suppression of the oscillator strength for transitions between these states. Increasing pressure causes the HOMO to spread into the core and increasing overlap with the LUMO and the associated oscillator strength, resulting in an overall red shift of the absorption. This dependence upon the choice of ligand, which derives from the environment of the nanocrystal, suggests new possibilities for controlling the optoelectronic properties of nanomaterials for applications such as pressure sensors.

D. Band structure projection in 2D material heterostructures

Heterostructures are ubiquitous in real applications of 2D materials. Any 2D material that is incorporated into any kind of device automatically becomes, in effect, part of a heterostructure, since the atomic-scale widths of these materials mean that their “bulk” regions are never really independent of their surroundings. Furthermore, novel physics often emerges from the hybridization of the band structure that results when materials are combined. Controlling the properties of layered material heterostructures is therefore crucial to the success of devices based on the novel capabilities of 2D materials.

However, while monolayer 2D materials are very well-studied, the theoretical insight into heterostructures available from traditional DFT is limited by the large system sizes required to study the interfaces between pairs of materials, which, in general, will be

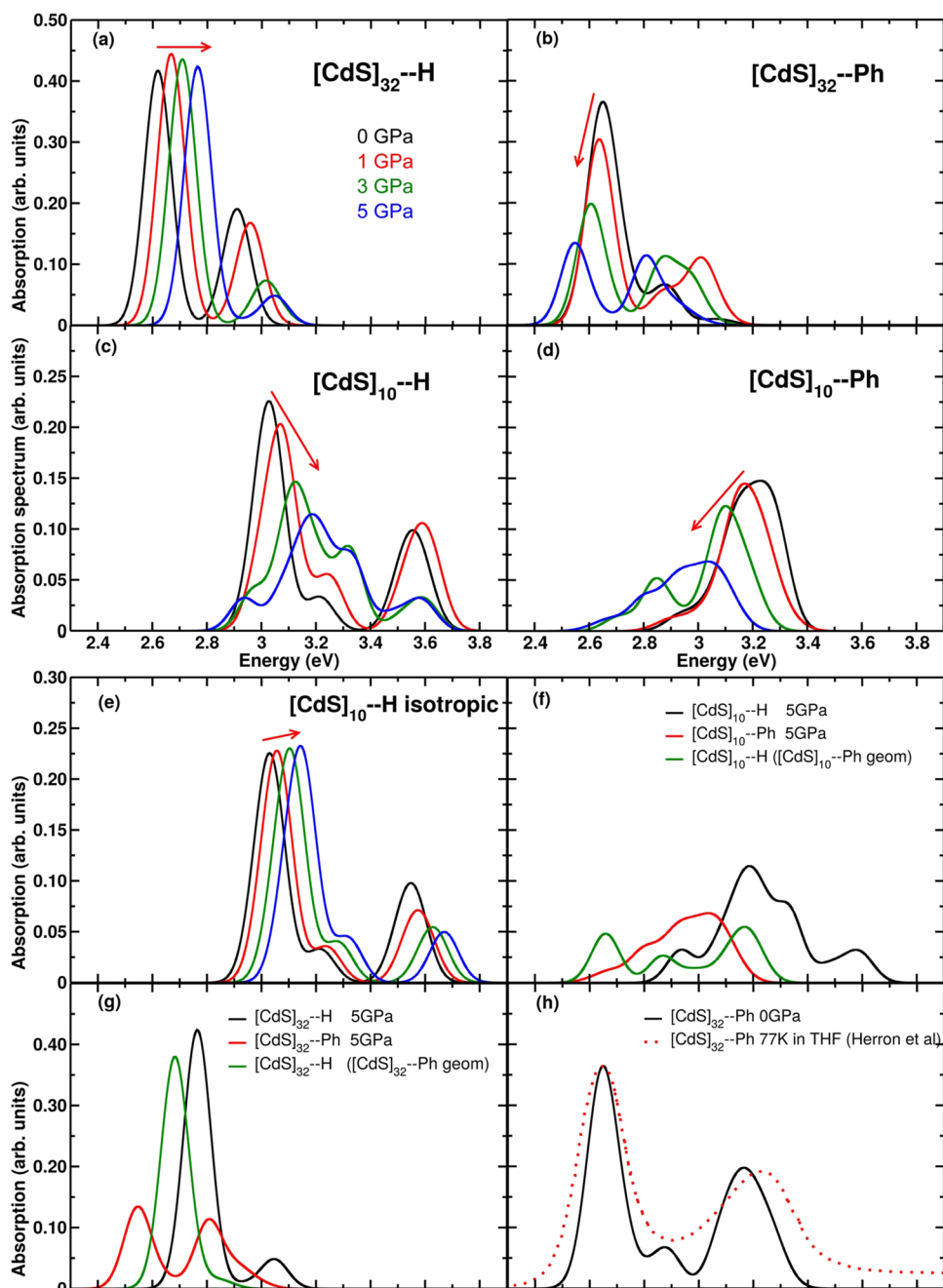


FIG. 18. Calculated optical absorption spectra at several different pressures for two different sizes of the CdS nanocrystal bonded to two different ligands. The arrows show the shift in the absorption spectrum with increasing pressure. Figure adapted (cropped) with permission from Fig. 5 of Corsini *et al.*, Nano Lett. 17, 1042 (2017). Copyright 2017 Author(s), licensed under a Creative Commons Attribution 3.0 Unported License. (see the original article for further calculation details).

incommensurate and may be rotated with respect to one another. Furthermore, while the effect of interlayer interactions is weak, given that the layers are bonded by only van der Waals interactions, there can be subtle, yet crucial, effects from hybridization. The combination available in ONETEP of non-local vdW functionals and band structure unfolding techniques with large-scale models of heterostructures is thus of great interest for applications in this field.

ONETEP has been applied to calculations on a range of heterostructure combinations including MoS₂/MoSe₂,⁹² MoSe₂/WSe₂,⁹⁴ and hBN/phosphorene.⁹³ Band structure changes caused by stacking and rotation of the layers have been observed and characterized in all these cases. Changes in spectral weight and band structure between the monolayers and heterostructured interfaces show how lattice mismatch (MoS₂/MoSe₂) or spacer layers (phosphorene/hBN/phosphorene) can allow the component monolayers to

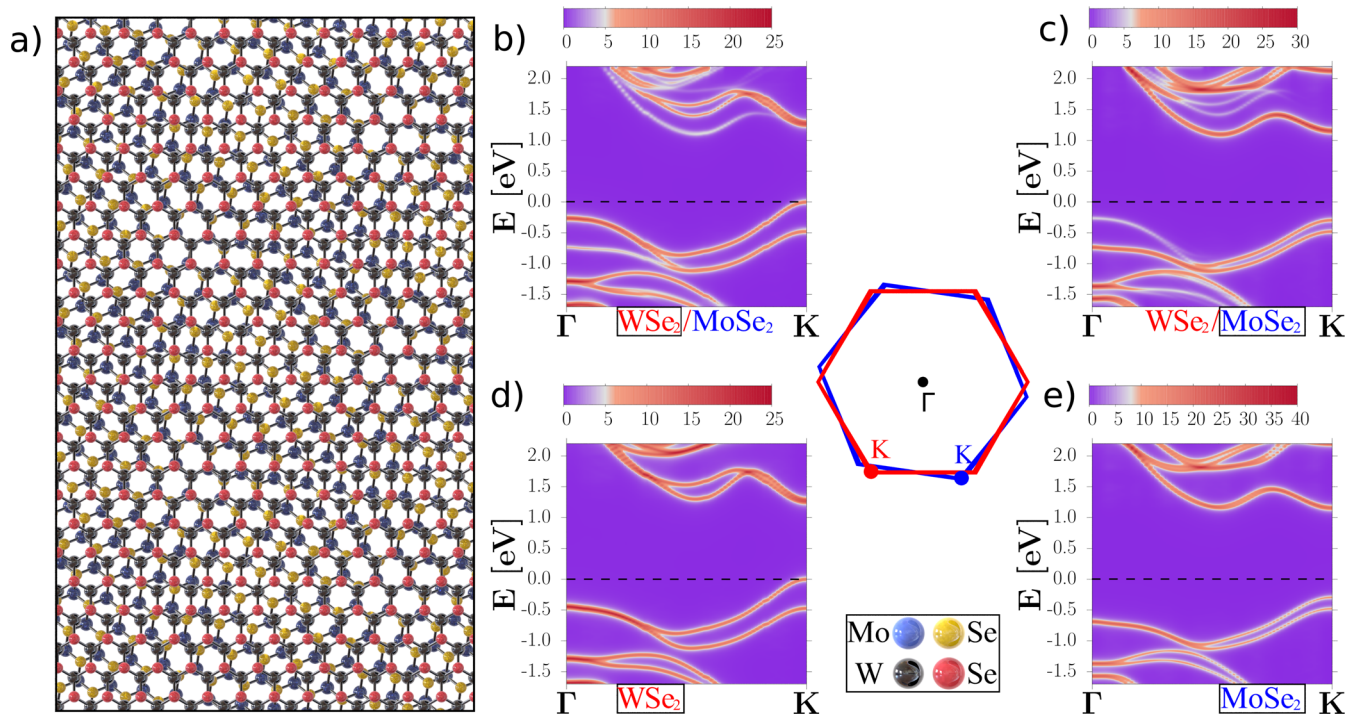


FIG. 19. Unfolded spectral function of the $\text{WSe}_2/\text{MoSe}_2$ heterostructure depicted in (a), projected on the WSe_2 monolayer (b) and MoSe_2 layer (c). The Γ -K directions are rotated by 8.21° , as shown in the central inset depicting the 1st Brillouin zones. The spectral-function representation of the independent monolayers is also shown for WSe_2 (c) and MoSe_2 (d). The energy reference is the energy of the valence band maximum (VBM) at K in the heterostructure. Based on Fig. S5 from the supplementary material of Ref. 94, which is published under a CC-BY license, with panel (a) reproduced from that work.

retain more independence in heterostructures than in homo-stacks. Significantly, comparison of band structures with ARPES results for aligned and non-aligned stack models (see Fig. 19) led to the confirmation that there can be co-existence of commensurate and incommensurate domains within the interface between a single pair of monolayer transition metal dichalcogenide flakes.⁹⁴

VI. COMMUNITY ACTIVITIES

ONETEP was developed from the beginning to be an academic platform to enable cutting edge developments in large-scale electronic structure theory. At the same time, the philosophy behind the actual software has been to make it as efficient and simple (user friendly) to use as possible, even if this requires significantly more software engineering effort than a traditional academic code that is often targeted toward a small number of expert, highly specialized users. The reasoning behind this choice is that the new electronic structure capabilities that are developed in ONETEP are meant to be available long term for use in the future for application in challenging problems in industry and academia. This has led to the early adoption of the code by industry via commercialization by Accelrys (since acquired by Dassault Systèmes BIOVIA) in late 2004. This relationship has grown over the years, and ONETEP is currently seen as a powerful computational simulation tool by industry, which is marketed currently by Dassault Systèmes BIOVIA.²⁸⁹

The popularity and usage of the code has also increased in academia with a large number of groups that are not associated with the developers using it for a multitude of applications. In 2017, ONETEP became the Flagship Project of CCP9, the EPSRC-funded UK Collaborative Computational Project for the Study of the Electronic Structure of Condensed Matter.²⁹⁰ An academic license for the code covers an entire research group and provides access to the full source code and no limitations on the size of calculations that can be run, the use of HPC resources, or the number of calculations that can be run simultaneously. An academic license is obtained from the ONETEP website²⁹¹ and is currently free to UK academics members of CCP9 (all UK academics can join). For non-UK academics, there is a modest administration fee for the license.

The community of ONETEP users is growing. The leading events are the ONETEP Masterclasses, which provide training to both academic and industrial users of the code. A distinct feature of these events is that, unlike usual code training schools where the participants train on set exercises, each participant is allocated two personal tutors who help them work directly on the problems of their interest so that they are ready to continue work with the code after the Masterclass. Apart from this hands-on training, the Masterclass consists of lectures on the general theory of ONETEP and specialized lectures of specific topics voted for by the participants. Each Masterclass has about 20 participants who are a mixture of academic and industrial users of the code.

Another more specialized community activity are the ONETEP coding retreats. These last for about a week and bring together people who have considerable experience of developing the code so that they can work together intensively to develop specific functionalities, achieve inter-operability between existing capabilities, eliminate known bugs, clean up or refactor parts of the code, and write documentation.

ONETEP is ultimately a community code and as such it is one of the codes of the CCP9 and UKCP²⁹² consortia for the electronic structure community. Through these consortia, the community can learn about the most recent developments and applications of the code and also influence the developments according to its needs. On the international side, ONETEP is one of the codes of the Psi-k European network for electronic structure;²⁹³ it is part of activities of this network and its large community, which extends worldwide and provides a forum for interaction between the users and developers of electronic structure codes via conferences, workshops, mailing list, and other resources. ONETEP is also one of the community code partners of The Molecular Sciences Software Institute (MolSSI)²⁹⁴ in the US, a nexus for science, education, and cooperation serving the worldwide community of computational molecular scientists.

VII. CONCLUSIONS AND FUTURE DIRECTIONS

We have presented an overview of the ONETEP program for linear-scaling density functional theory (DFT) calculations with large basis set (plane-wave) accuracy on parallel computers. The DFT energy is computed from the density matrix, which is constructed from spatially localized orbitals we call non-orthogonal generalized Wannier functions (NGWFs). During the calculation, the density matrix and the NGWFs are optimized with localization constraints. A basis set of periodic sinc (psinc) functions, which is equivalent to plane waves, is used to expand the NGWFs. By taking advantage of density matrix localization, ONETEP is able to perform calculations with thousands of atoms with computational effort, which scales linearly with the number of atoms. We have described the large and diverse range of capabilities of the code, which include different boundary conditions, many types of DFT exchange–correlation functionals, and finite electronic temperature DFT methods for metallic systems. The code also has a range of methods that go beyond DFT such as non-local exchange and strong correlation. There are advanced capabilities for geometry relaxation, molecular dynamics, phonons, and transition states. Excited state calculations are available through time-dependent DFT as well as relevant experimental observables such as electronic transport, UV/VIS spectra, core loss spectroscopy, and anharmonic vibrational spectra. The code is able to include the effect of the chemical environment in calculations via a solvation model, QM/MM embedding, and QM-in-QM embedding approaches. An extensive variety of electronic properties are also available ranging from density of states to distributed multipole analysis and methods for partitioning charges and interactions between fragments. Calculations with ONETEP provide unique insights into large and complex systems that require an accurate description at the atomic level ranging from biomolecular to chemical, to materials and physical problems, as we showed with a small selection of illustrative examples.

The first version of the ONETEP code was completed in summer 2004.⁷ While able to run calculations with thousands of atoms

and demonstrate linear-scaling behavior with plane-wave basis set accuracy, the only type of calculation it was able to do was single-point energy. Since then there has been a tremendous increase in the capabilities of the code that now contains an enormous and varied range of functionalities, as we have described in this paper, and the code is being used for applications that cover all possible areas of applicability of electronic structure theory.

A community code and project with unique capabilities such as ONETEP needs to remain at the cutting edge of physical theory method developments as well as software developments. The future directions for the developments of the code will depend on the needs of the users and the applications that they wish to do with the code. It is likely that the range of capabilities will continue to grow. It is also expected that more developments toward multiscale methods will be done in order to include environment (e.g., solvent, support, etc.) effects in the calculations. This will likely involve implementations of computationally cheaper semi-empirical variants of DFT that will take advantage of the linear-scaling framework of ONETEP to perform extensive configurational sampling, as in molecular dynamics simulations. Developments are likely to take place also at the other end of the spectrum of accuracy with the implementation of more sophisticated exchange–correlation functionals and the implementation of quantum chemistry wavefunction-based methods, which will be based on the two electron repulsion integral engine for NGWFs that we have developed for hybrid functionals. Another interesting direction would be to develop machine-learning methods to estimate certain quantities and hence reduce the calculational cost.

Finally, a code such as ONETEP needs to keep up with progress in supercomputing resources. In particular, it is imperative to be able to take advantage of the new disruptive technologies that are emerging and will make exascale computing a reality. Due to its linear-scaling computational cost, ONETEP has unique potential for simulations of complex materials with high industrial relevance. Often, a common feature of these challenging problems is that they require very large scales of simulation with many thousands of atoms and also a large number of these simulations (configuration sampling). Thus, to make such simulations realistic enough to be able to have predictive power exascale computing is absolutely necessary. While we have made excellent progress in developing ONETEP to run efficiently on “conventional” large multicore HPC resources, we have a lot of work to do to meet the challenge of exascale computing. Radically different approaches are needed for the redesign of our core algorithms to run efficiently on these resources that will require significant software engineering work over the coming years. This will involve also taking advantage of GPUs, an area where ONETEP already has some capability²⁹⁵ and can be significantly improved in terms of performance. Overall, a code and project such as ONETEP needs to be always under development in order to remain useful for its target applications.

AUTHOR'S CONTRIBUTIONS

J.C.A.P, J.A., and J.C.W. contributed equally to this work.

ACKNOWLEDGMENTS

Computing resources were provided by a number of facilities including the Scientific Computing Research Technology

Platform of the University of Warwick, the IRIDIS-5 supercomputer of the University of Southampton, the Imperial College High Performance Computing Service, the University of Cambridge High Performance Computing Service, and the TASK Academic Computer Centre (Gdańsk, Poland). The use of the ARCHER supercomputer was supported by both the UKCP consortium, funded by EPSRC Grant No. EP/P022030/1, and the Materials Chemistry Consortium, funded by EPSRC Grant Nos. EP/L000202/1 and EP/R029431/1. We acknowledge the use of Athena at HPC Midlands+, which was funded by the EPSRC through Grant No. EP/P020232/1, as part of the HPC Midlands+ consortium, and the UK Materials and Molecular Modeling Hub for computational resources, which was partially funded by EPSRC (Grant No. EP/P020194/1). We acknowledge funding under the embedded CSE program (Project Nos. eCSE01-004, eCSE07-006, and eCSE08-15) of the ARCHER UK National Supercomputing Service, and the distributed CSE program of HEC-ToR, the former UK National Supercomputing Service, which supported the development of DL_MG. J.A., J.C.A.P., J.C.W., J.D., C.-K.S., P.D.H., A.A.M., and N.D.M.H. acknowledge the support of EPSRC (Grant Nos. EP/P02209X/1, and previously EP/F010974/1, EP/G05567X/1, and EP/J015059/1), which supported earlier development of ONETEP. J.C.A.P., L. Andrinopoulos, R.A.B., R. J. Charlton, F.C., A.G., D.D.O., L.E.R., V.V., T.J.Z., P.D.H., N.D.M.H., and A.A.M. all acknowledge support from the Thomas Young Centre for Theory and Simulation of Materials through Grant No. TYC-101. J.C.A.P. also acknowledges support from St Edmund Hall, University of Oxford. L. Andrinopoulos was supported by a studentship from the EPSRC Grant No. EP/G05567X/1. R. J. Charlton, A.G., and T.J.Z. were supported through studentships in the Centre for Doctoral Training in Theory and Simulation of Materials at Imperial College London, funded by the EPSRC through Grant Nos. EP/G036888/1 and EP/L015579/1. T.J.Z. acknowledges funding under the embedded CSE program of the ARCHER UK National Supercomputing Service (eCSE02-15). A.E.A.A., K.K.B.D., and E.L. were supported through studentships in the Centre for Doctoral Training in Computational Methods for Materials Science at the University of Cambridge, funded by the EPSRC through Grant No. EP/L015552/1. E.W.T. was supported through a studentship at the Cambridge Centre for Doctoral Training in Nanoscience and Nanotechnology, funded by the EPSRC through Grant No. EP/G037221/1. V.V. and A.A.M. were supported by the EPSRC through Grant No. EP/S025324/1. N.D.M.H. was supported by the EPSRC through Grant No. EP/P01139X/1. V.V. was also supported through a studentship at the Doctoral Training Centre of the Institute for Complex System Simulations, funded by the EPSRC through Grant No. EP/G03690X/1. G.A.B. acknowledges the EPSRC and Pacific Northwest National Laboratory (PNNL) for Ph.D. funding. Q.H. also acknowledges the EPSRC for research studentship funding. M.J.S.P. would like to thank the BBSRC and Boehringer Ingelheim for an industrial CASE studentship. R. J. Clements was supported through a studentship in the Centre for Doctoral Training in Next Generation of Computational Modelling, funded by the EPSRC through Grant No. EP/L015382/1. A.B., J.D., and J.C.W. acknowledge funding from the Faraday Institution (faraday.ac.uk); EPSRC Grant No. EP/S003053/1) through Grant No. FIRG003. A.R.S. acknowledges the support of the EPSRC for a high end computing studentship, funded by Grant No. EP/F038038/1, through the UKCP consortium. M.C.P., J.M.E., and T.J.Z. acknowledge

support from the EPSRC through Grant No. EP/J017639/1. M.C.P. and J.M.E. acknowledge support from the EPSRC through Grant Nos. EP/J015059/1 and EP/P034616/1. M.C.P., N.D.M.H., and S.M.-M.D. acknowledge support from the EPSRC through Grant No. EP/G055904/1. J.M.E. also acknowledges support from Ministerio de Ciencia e Innovación of Spain through the María de Maeztu grant MDM-2017-0767. D.J.C. acknowledges support from the EPSRC through Grant No. EP/R010153/1. G.T. acknowledges support from the EPSRC through Grant Nos. EP/I004483/1, EP/P022189/1, and EP/P022189/2. D.D.O.R. acknowledges the support of Science Foundation Ireland (SFI) through AMBER (The Advanced Materials and Bioengineering Research Centre, Grant Nos. SFI/12/RC/2278 and SFI/12/RC/2278_P2), the European Regional Development Fund (ERDF), and the National University of Ireland. M.C.P. and D.D.O.R. acknowledge support from the EPSRC through Grant No. EP/F032773/1. D.D.O. and G.T. also acknowledge support by the Royal Irish Academy-Royal Society International Exchange Cost Share Programme (No. IE131505). Structures of the PBTZT-stat-BDIT-8 polymer (Figs. 7 and 8) were kindly provided by Gabriele Boschetto.

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- ¹X.-P. Li, R. W. Nunes, and D. Vanderbilt, "Density-matrix electronic-structure method with linear system-size scaling," *Phys. Rev. B* **47**, 10891 (1993).
- ²E. Prodan and W. Kohn, "Nearsightedness of electronic matter," *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11635 (2005).
- ³C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, "Exponential localization of Wannier functions in insulators," *Phys. Rev. Lett.* **98**, 046402 (2007).
- ⁴P. W. Anderson, "Self-consistent pseudopotentials and ultralocalized functions for energy bands," *Phys. Rev. Lett.* **21**, 13 (1968).
- ⁵G. Galli and M. Parrinello, "Large scale electronic structure calculations," *Phys. Rev. Lett.* **69**, 3547 (1992).
- ⁶E. Hernández and M. J. Gillan, "Self-consistent first-principles technique with linear scaling," *Phys. Rev. B* **51**, 10157 (1995).
- ⁷C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, "Introducing ONETEP: Linear-scaling density functional simulations on parallel computers," *J. Chem. Phys.* **122**, 084119 (2005).
- ⁸M. J. Gillan, D. R. Bowler, A. S. Torralba, and T. Miyazaki, "Order-N first-principles calculations with the CONQUEST code," *Comput. Phys. Commun.* **177**, 14 (2007).
- ⁹S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, and T. Deutsch, "Accurate and efficient linear scaling DFT calculations with universal applicability," *Phys. Chem. Chem. Phys.* **17**, 31360 (2015).
- ¹⁰C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Diéguez, and M. C. Payne, "Nonorthogonal generalized Wannier function pseudopotential plane-wave method," *Phys. Rev. B* **66**, 035119 (2002).
- ¹¹A. A. Mostofi, P. D. Haynes, C.-K. Skylaris, and M. C. Payne, "Preconditioned iterative minimization for linear-scaling electronic structure calculations," *J. Chem. Phys.* **119**, 8842–8848 (2003).
- ¹²C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, C. J. Pickard, and M. C. Payne, "Accurate kinetic energy evaluation in electronic structure calculations with localized functions on real space grids," *Comput. Phys. Commun.* **140**, 315–322 (2001).
- ¹³A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, and M. C. Payne, "Total-energy calculations on a real space grid with localized functions and a plane-wave basis," *Comput. Phys. Commun.* **147**, 788–802 (2002).

- ¹⁴C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, "Implementation of linear-scaling plane wave density functional theory on parallel computers," *Phys. Status Solidi B* **243**, 973–988 (2006).
- ¹⁵N. D. M. Hine, P. D. Haynes, A. A. Mostofi, C.-K. Skylaris, and M. C. Payne, "Linear-scaling density-functional theory with tens of thousands of atoms: Expanding the scope and scale of calculations with ONETEP," *Comput. Phys. Commun.* **180**, 1041–1053 (2009).
- ¹⁶K. A. Wilkinson, N. D. M. Hine, and C.-K. Skylaris, "Hybrid MPI-OpenMP parallelism in the ONETEP linear-scaling electronic structure code: Application to the delamination of cellulose nanofibrils," *J. Chem. Theory Comput.* **10**, 4782–4794 (2014).
- ¹⁷M. Challacombe, "A general parallel sparse-blocked matrix multiply for linear scaling SCF theory," *Comput. Phys. Commun.* **128**, 93–107 (2000).
- ¹⁸R. L. Graham, "Bounds for certain multiprocessing anomalies," *Bell Syst. Tech. J.* **45**, 1563–1581 (1966).
- ¹⁹J. C. Womack, N. Mardirossian, M. Head-Gordon, and C.-K. Skylaris, "Self-consistent implementation of meta-GGA functionals for the ONETEP linear-scaling electronic structure package," *J. Chem. Phys.* **145**, 204114 (2016).
- ²⁰N. D. M. Hine, "Linear-scaling density functional theory using the projector augmented wave method," *J. Phys.: Condens. Matter* **29**, 024001 (2016).
- ²¹P. E. Blöchl, "Projector augmented-wave method," *Phys. Rev. B* **50**, 17953 (1994).
- ²²M. Torrent, F. Jollet, F. Bottin, G. Zerah, and X. Gonze, "Implementation of the projector augmented-wave method in the ABINIT code: Application to the study of iron under pressure," *Comput. Mater. Sci.* **42**, 337–351 (2008).
- ²³N. D. Mermin, "Thermal properties of the inhomogeneous electron gas," *Phys. Rev.* **137**, A1441 (1965).
- ²⁴N. Marzari, D. Vanderbilt, and M. C. Payne, "Ensemble density-functional theory for *ab initio* molecular dynamics of metals and finite-temperature insulators," *Phys. Rev. Lett.* **79**, 1337 (1997).
- ²⁵C. Freysoldt, S. Boeck, and J. Neugebauer, "Direct minimization technique for metals in density functional theory," *Phys. Rev. B* **79**, 241103 (2009).
- ²⁶M. Methfessel and A. T. Paxton, "High-precision sampling for Brillouin-zone integration in metals," *Phys. Rev. B* **40**, 3616 (1989).
- ²⁷C.-L. Fu and K.-M. Ho, "First-principles calculation of the equilibrium ground-state properties of transition metals: Applications to Nb and Mo," *Phys. Rev. B* **28**, 5480 (1983).
- ²⁸R. J. Needs, R. M. Martin, and O. H. Nielsen, "Total-energy calculations of the structural properties of the group-V element arsenic," *Phys. Rev. B* **33**, 3778 (1989).
- ²⁹Á. Ruiz-Serrano and C.-K. Skylaris, "A variational method for density functional theory calculations on metallic systems with thousands of atoms," *J. Chem. Phys.* **139**, 054107 (2013).
- ³⁰S. Goedecker and L. Colombo, "Efficient linear scaling algorithm for tight-binding molecular dynamics," *Phys. Rev. Lett.* **73**, 122 (1994).
- ³¹S. Goedecker and M. Teter, "Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals," *Phys. Rev. B* **51**, 9455 (1995).
- ³²S. Goedecker, "Low complexity algorithms for electronic structure calculations," *J. Comput. Phys.* **118**, 261–268 (1995).
- ³³W. Liang, C. Saravanan, Y. Shao, R. Baer, A. T. Bell, and M. Head-Gordon, "Improved Fermi operator expansion methods for fast electronic structure calculations," *J. Chem. Phys.* **119**, 4117–4125 (2003).
- ³⁴S. Goedecker, "Integral representation of the Fermi distribution and its applications in electronic-structure calculations," *Phys. Rev. B* **48**, 17573 (1993).
- ³⁵L. Lin, J. Lu, L. Ying, and E. Weinan, "Pole-based approximation of the Fermi-Dirac function," *Chin. Ann. Math. Ser. B* **30**, 729–742 (2009).
- ³⁶M. Ceriotti, T. D. Kühne, and M. Parrinello, "An efficient and accurate decomposition of the Fermi operator," *J. Chem. Phys.* **129**, 024707 (2008).
- ³⁷P. Suryanarayana, "On spectral quadrature for linear-scaling density functional theory," *Chem. Phys. Lett.* **584**, 182–187 (2013).
- ³⁸A. M. N. Niklasson, M. J. Cawkwell, E. H. Rubensson, and E. Rudberg, "Canonical density matrix perturbation theory," *Phys. Rev. E* **92**, 063301 (2015).
- ³⁹J. Aarons and C.-K. Skylaris, "Electronic annealing Fermi operator expansion for DFT calculations on metallic systems," *J. Chem. Phys.* **148**, 074107 (2018).
- ⁴⁰C. Sgiarovello, M. Peressi, and R. Resta, "Electron localization in the insulating state: Application to crystalline semiconductors," *Phys. Rev. B* **64**, 115202 (2001).
- ⁴¹N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, "Maximally localized Wannier functions: Theory and applications," *Rev. Mod. Phys.* **84**, 1419 (2012).
- ⁴²F. Giustino and A. Pasquarello, "Theory of atomic-scale dielectric permittivity at insulator interfaces," *Phys. Rev. B* **71**, 144104 (2005).
- ⁴³M. Stengel and N. A. Spaldin, "Accurate polarization within a unified Wannier function formalism," *Phys. Rev. B* **73**, 075121 (2006).
- ⁴⁴S. Coh and D. Vanderbilt, "Electric polarization in a Chern insulator," *Phys. Rev. Lett.* **102**, 107603 (2009).
- ⁴⁵H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, "Topological insulators in Bi_2Se_3 , Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface," *Nat. Phys.* **5**, 438 (2009).
- ⁴⁶N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," *Phys. Rev. B* **56**, 12847 (1997).
- ⁴⁷F. Gygi, J.-L. Fattebert, and E. Schwegler, "Computation of maximally localized Wannier functions using a simultaneous diagonalization algorithm," *Comput. Phys. Commun.* **155**, 1 (2003).
- ⁴⁸A. Greco, "Development and application of first-principles methods for complex oxide surfaces and interfaces," Ph.D. thesis, Imperial College London, 2017, <http://hdl.handle.net/10044/1/56213>.
- ⁴⁹S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. Payne, "First principles methods using CASTEP," *Z. Kristall.* **220**, 567–570 (2005).
- ⁵⁰J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, "The SIESTA method for *ab initio* order- N materials simulation," *J. Phys.: Condens. Matter* **14**, 2745 (2002).
- ⁵¹M. Tañipolsky and R. Schmid, "A general and efficient pseudopotential Fourier filtering scheme for real space methods using mask functions," *J. Chem. Phys.* **124**, 174102 (2006).
- ⁵²E. Artacho, E. Anglada, O. Diéguez, J. D. Gale, A. García, J. Junquera, R. M. Martin, P. Ordejón, J. M. Pruneda, D. Sánchez-Portal, and J. M. Soler, "The SIESTA method: Developments and applicability," *J. Phys.: Condens. Matter* **20**, 064208 (2008).
- ⁵³N. D. M. Hine, M. Robinson, P. D. Haynes, C.-K. Skylaris, M. C. Payne, and A. A. Mostofi, "Accurate ionic forces and geometry optimization in linear-scaling density-functional theory with local orbitals," *Phys. Rev. B* **83**, 195102 (2011).
- ⁵⁴Á. Ruiz-Serrano, N. D. M. Hine, and C.-K. Skylaris, "Pulay forces from localized orbitals optimized *in situ* using a psinc basis set," *J. Chem. Phys.* **136**, 234101 (2012).
- ⁵⁵J. C. Womack, L. Anton, J. Dziejczak, P. J. Hasnip, M. I. J. Probert, and C.-K. Skylaris, "DL_MG: A parallel multigrid Poisson and Poisson-Boltzmann solver for electronic structure calculations in vacuum and solution," *J. Chem. Theory Comput.* **14**, 1412–1432 (2018).
- ⁵⁶J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy," *AIP Conf. Proc.* **577**, 1–20 (2001).
- ⁵⁷J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, "Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits," *J. Chem. Phys.* **123**, 062201 (2005).
- ⁵⁸R. Neumann, R. H. Nobes, and N. C. Handy, "Exchange functionals and potentials," *Mol. Phys.* **87**, 1–36 (1996).
- ⁵⁹A. V. Arbuznikov and M. Kaupp, "The self-consistent implementation of exchange-correlation functionals depending on the local kinetic energy density," *Chem. Phys. Lett.* **381**, 495–504 (2003).
- ⁶⁰A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, "Generalized Kohn-Sham schemes and the band-gap problem," *Phys. Rev. B* **53**, 3764–3774 (1996).

- ⁶¹J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, "Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation," *Phys. Rev. Lett.* **82**, 2544–2547 (1999).
- ⁶²J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, "Erratum: Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation [Phys. Rev. Lett. 82, 2544 (1999)]," *Phys. Rev. Lett.* **82**, 5179 (1999).
- ⁶³N. Mardirossian and M. Head-Gordon, "Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V," *J. Chem. Phys.* **142**, 074111 (2015).
- ⁶⁴N. Mardirossian, L. Ruiz Pestana, J. C. Womack, C.-K. Skylaris, T. Head-Gordon, and M. Head-Gordon, "Use of the rVV10 nonlocal correlation functional in the B97M-V density functional: Defining B97m-rV and related functionals," *J. Phys. Chem. Lett.* **8**, 35–40 (2017).
- ⁶⁵N. Mardirossian and M. Head-Gordon, "Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals," *Mol. Phys.* **115**, 2315–2372 (2017).
- ⁶⁶O. A. Vydrov and T. Van Voorhis, "Nonlocal van der Waals density functional: The simpler the better," *J. Chem. Phys.* **133**, 244103 (2010).
- ⁶⁷R. Sabatini, T. Gorni, and S. de Gironcoli, "Nonlocal van der Waals density functional made simple and efficient," *Phys. Rev. B* **87**, 041108 (2013).
- ⁶⁸A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *J. Chem. Phys.* **98**, 5648–5652 (1993).
- ⁶⁹A. V. Arbuznikov, "Hybrid exchange correlation functionals and potentials: Concept elaboration," *J. Struct. Chem.* **48**, S1–S31 (2007).
- ⁷⁰P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," *J. Phys. Chem.* **98**, 11623–11627 (1994).
- ⁷¹J. Dziedzic, Q. Hill, and C.-K. Skylaris, "Linear-scaling calculation of Hartree-Fock exchange energy with non-orthogonal generalised Wannier functions," *J. Chem. Phys.* **139**, 214103 (2013).
- ⁷²Q. Hill, "Development of more accurate computational methods within linear-scaling density functional theory," Ph.D. thesis, University of Southampton, Southampton, 2010, <https://eprints.soton.ac.uk/id/eprint/193151>.
- ⁷³P. D. Haynes and M. C. Payne, "Localised spherical-wave basis set for O(N) total-energy pseudopotential calculations," *Comput. Phys. Commun.* **102**, 17–27 (1997).
- ⁷⁴G. Boschetto, H.-T. Xue, J. Dziedzic, M. Krompiec, and C.-K. Skylaris, "Effect of polymerization statistics on the electronic properties of copolymers for organic photovoltaics," *J. Phys. Chem. C* **121**, 2529–2538 (2017).
- ⁷⁵J. C. Womack, J. Dziedzic, and C.-K. Skylaris, "Fast and accurate Coulomb metric matrix evaluation in a truncated spherical wave basis via mixed numerical/analytic integration" (unpublished) (2020).
- ⁷⁶J. Dziedzic, J. C. Womack, R. Ali, and C.-K. Skylaris, "Massively parallel linear-scaling Hartree-Fock exchange and hybrid exchange-correlation functionals with plane-wave basis set accuracy" (unpublished) (2020).
- ⁷⁷T. J. Zuehlsdorff, N. D. M. Hine, J. S. Spencer, N. M. Harrison, D. J. Riley, and P. D. Haynes, "Linear-scaling time-dependent density-functional theory in the linear response formalism," *J. Chem. Phys.* **139**, 064104 (2013).
- ⁷⁸S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range dispersion correction," *J. Comput. Chem.* **27**, 1787–1799 (2006).
- ⁷⁹S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "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 (2010).
- ⁸⁰M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, "Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment," *J. Chem. Phys.* **114**, 5149–5155 (2001).
- ⁸¹Q. Wu and W. Yang, "Empirical correction to density functional theory for van der Waals interactions," *J. Chem. Phys.* **116**, 515–524 (2002).
- ⁸²Q. Hill and C.-K. Skylaris, "Including dispersion interactions in the ONETEP program for linear-scaling density functional theory calculations," *Proc. R. Soc. A* **465**, 669–683 (2009).
- ⁸³M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der Waals density functional for general geometries," *Phys. Rev. Lett.* **92**, 246401 (2004).
- ⁸⁴G. Román-Pérez and J. M. Soler, "Efficient implementation of a van der Waals density functional: Application to double-wall carbon nanotubes," *Phys. Rev. Lett.* **103**, 096102 (2009).
- ⁸⁵L. Andrinopoulos, "Including van der Waals interactions in first-principles electronic structure calculations," Ph.D. thesis, Imperial College London, 2013, <http://hdl.handle.net/10044/1/22152>.
- ⁸⁶Y. Zhang and W. Yang, "Comment on "Generalized gradient approximation made simple"," *Phys. Rev. Lett.* **80**, 890 (1998).
- ⁸⁷J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," *Phys. Rev. B* **45**, 13244 (1992).
- ⁸⁸K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, "Higher-accuracy van der Waals density functional," *Phys. Rev. B* **82**, 081101 (2010).
- ⁸⁹J. Klimeš, D. R. Bowler, and A. Michaelides, "Chemical accuracy for the van der Waals density functional," *J. Phys.: Condens. Matter* **22**, 022201 (2009).
- ⁹⁰J. Klimeš, D. R. Bowler, and A. Michaelides, "Van der Waals density functionals applied to solids," *Phys. Rev. B* **83**, 195131 (2011).
- ⁹¹V. R. Cooper, "Van der Waals density functional: An appropriate exchange functional," *Phys. Rev. B* **81**, 161104 (2010).
- ⁹²G. C. Constantinescu and N. D. M. Hine, "Energy landscape and band-structure tuning in realistic MoS₂/MoSe₂ heterostructures," *Phys. Rev. B* **91**, 195416 (2015).
- ⁹³G. C. Constantinescu and N. D. M. Hine, "Multipurpose black-phosphorus/hBN heterostructures," *Nano Lett.* **16**, 2586–2594 (2016).
- ⁹⁴N. R. Wilson, P. V. Nguyen, K. Seyler, P. Rivera, A. J. Marsden, Z. P. L. Laker, G. C. Constantinescu, V. Kandyba, A. Barinov, N. D. M. Hine, X. Xu, and D. H. Cobden, "Determination of band offsets, hybridization, and exciton binding in 2D semiconductor heterostructures," *Sci. Adv.* **3**, e1601832 (2017), <https://advances.sciencemag.org/content/3/2/e1601832.full.pdf>.
- ⁹⁵V. I. Anisimov, J. Zaanen, and O. K. Andersen, "Band theory and Mott insulators: Hubbard *U* instead of Stoner *I*," *Phys. Rev. B* **44**, 943–954 (1991).
- ⁹⁶V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, "First-principles calculations of the electronic structure and spectra of strongly correlated systems: The LDA + *U* method," *J. Phys.: Condens. Matter* **9**, 767 (1997).
- ⁹⁷S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+*U* study," *Phys. Rev. B* **57**, 1505–1509 (1998).
- ⁹⁸D. J. Cole, D. D. O'Regan, and M. C. Payne, "Ligand discrimination in myoglobin from linear-scaling DFT+*U*," *J. Phys. Chem. Lett.* **3**, 1448–1452 (2012).
- ⁹⁹D. D. O'Regan, N. D. M. Hine, M. C. Payne, and A. A. Mostofi, "Linear-scaling DFT+*U* with full local orbital optimization," *Phys. Rev. B* **85**, 085107 (2012).
- ¹⁰⁰D. D. O'Regan, M. C. Payne, and A. A. Mostofi, "Subspace representations in *ab initio* methods for strongly correlated systems," *Phys. Rev. B* **83**, 245124 (2011).
- ¹⁰¹D. D. O'Regan, N. D. M. Hine, M. C. Payne, and A. A. Mostofi, "Projector self-consistent DFT + *U* using nonorthogonal generalized Wannier functions," *Phys. Rev. B* **82**, 081102 (2010).
- ¹⁰²O. K. Orhan and D. D. O'Regan, "TDDFT + *U*: A critical assessment of the Hubbard *U* correction to exchange-correlation kernels and potentials," *Phys. Rev. B* **99**, 165120 (2019).
- ¹⁰³D. D. O'Regan and G. Teobaldi, "Optimization of constrained density functional theory," *Phys. Rev. B* **94**, 035159 (2016).
- ¹⁰⁴C. Weber, D. D. O'Regan, N. D. M. Hine, M. C. Payne, G. Kotliar, and P. B. Littlewood, "Vanadium dioxide: A Peierls-Mott insulator stable against disorder," *Phys. Rev. Lett.* **108**, 256402 (2012).
- ¹⁰⁵E. B. Linscott, D. J. Cole, N. D. M. Hine, M. C. Payne, and C. Weber, "ONETEP + TOSCAM: Uniting dynamical mean field theory and linear-scaling density functional theory," [arXiv:1911.07752](https://arxiv.org/abs/1911.07752) (2019).
- ¹⁰⁶D. D. O'Regan, *Optimised Projections for the ab initio Simulation of Large and Strongly Correlated Systems*, Springer Theses XVI (Springer, Berlin, Heidelberg, 2012).
- ¹⁰⁷G. Moynihan, "A self-contained ground-state approach for the correction of self-interaction error in approximate density-functional theory," Ph.D. thesis, Trinity College Dublin, 2018, <http://www.tara.tcd.ie/handle/2262/82220>.

- ¹⁰⁸O. K. Orhan, "Corrective first-principles approaches for the theoretical spectroscopy of transition-metal systems," Ph.D. thesis, Trinity College Dublin, 2018, <http://www.tara.tcd.ie/handle/2262/84975>.
- ¹⁰⁹E. B. Linscott, D. J. Cole, M. C. Payne, and D. D. O'Regan, "Role of spin in the calculation of Hubbard U and Hund's J parameters from first principles," *Phys. Rev. B* **98**, 235157 (2018).
- ¹¹⁰G. Moynihan, G. Teobaldi, and D. D. O'Regan, "Inapplicability of exact constraints and a minimal two-parameter generalization to the DFT+ U based correction of self-interaction error," *Phys. Rev. B* **94**, 220104 (2016).
- ¹¹¹M. Cococcioni and S. de Gironcoli, "Linear response approach to the calculation of the effective interaction parameters in the LDA + U method," *Phys. Rev. B* **71**, 035105 (2005).
- ¹¹²H. J. Kulik, M. Cococcioni, D. A. Scherlis, and N. Marzari, "Density functional theory in transition-metal chemistry: A self-consistent Hubbard U approach," *Phys. Rev. Lett.* **97**, 103001 (2006).
- ¹¹³B. Himmetoglu, R. M. Wentzcovitch, and M. Cococcioni, "First-principles study of electronic and structural properties of CuO," *Phys. Rev. B* **84**, 115108 (2011).
- ¹¹⁴P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, "Ground states of constrained systems: Application to cerium impurities," *Phys. Rev. Lett.* **53**, 2512–2515 (1984).
- ¹¹⁵Q. Wu and T. Van Voorhis, "Constrained density functional theory and its application in long-range electron transfer," *J. Chem. Theory Comput.* **2**, 765–774 (2006).
- ¹¹⁶Q. Wu and T. Van Voorhis, "Direct calculation of electron transfer parameters through constrained density functional theory," *J. Phys. Chem. A* **110**, 9212–9218 (2006).
- ¹¹⁷Q. Wu and T. Van Voorhis, "Extracting electron transfer coupling elements from constrained density functional theory," *J. Chem. Phys.* **125**, 164105 (2006).
- ¹¹⁸B. Kaduk, T. Kowalczyk, and T. Van Voorhis, "Constrained density functional theory," *Chem. Rev.* **112**, 321 (2012).
- ¹¹⁹Q. Wu and T. Van Voorhis, "Direct optimization method to study constrained systems within density-functional theory," *Phys. Rev. A* **72**, 024502 (2005).
- ¹²⁰Q. Wu, C.-L. Cheng, and T. Van Voorhis, "Configuration interaction based on constrained density functional theory: A multireference method," *J. Chem. Phys.* **127**, 164119 (2007).
- ¹²¹Q. Wu, B. Kaduk, and T. Van Voorhis, "Constrained density functional theory based configuration interaction improves the prediction of reaction barrier heights," *J. Chem. Phys.* **130**, 034109 (2009).
- ¹²²B. Kaduk and T. Van Voorhis, "Communication: Conical intersections using constrained density functional theory-configuration interaction," *J. Chem. Phys.* **133**, 061102 (2010).
- ¹²³S. Roychoudhury, D. D. O'Regan, and S. Sanvito, "Wannier-function-based constrained DFT with nonorthogonality-correcting Pulay forces in application to the reorganization effects in graphene-adsorbed pentacene," *Phys. Rev. B* **97**, 205120 (2018).
- ¹²⁴D. H. P. Turban, G. Teobaldi, D. D. O'Regan, and N. D. M. Hine, "Super-cell convergence of charge-transfer energies in pentacene molecular crystals from constrained DFT," *Phys. Rev. B* **93**, 165102 (2016).
- ¹²⁵S. Lukman, K. Chen, J. M. Hodgkiss, D. H. P. Turban, N. D. M. Hine, S. Dong, J. Wu, N. C. Greenham, and A. J. Musser, "Tuning the role of charge-transfer states in intramolecular singlet exciton fission through side-group engineering," *Nat. Commun.* **7**, 13622 (2016).
- ¹²⁶W. Metzner and D. Vollhardt, "Correlated lattice fermions in $d = \infty$ dimensions," *Phys. Rev. Lett.* **62**, 324–327 (1989).
- ¹²⁷A. Georges and G. Kotliar, "Hubbard model in infinite dimensions," *Phys. Rev. B* **45**, 6479–6483 (1992).
- ¹²⁸V. I. Anisimov, A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar, "First-principles calculations of the electronic structure and spectra of strongly correlated systems: Dynamical mean-field theory," *J. Phys.: Condens. Matter* **9**, 7359 (1997).
- ¹²⁹F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, "Calculations of Hubbard U from first-principles," *Phys. Rev. B* **74**, 125106 (2006).
- ¹³⁰K. Haule and T. Birol, "Free energy from stationary implementation of the DFT + DMFT functional," *Phys. Rev. Lett.* **115**, 256402 (2015).
- ¹³¹J. D. Becker, J. M. Wills, L. Cox, and B. R. Cooper, "Electronic structure of Pu compounds with group-IIIB metals: Two regimes of behavior," *Phys. Rev. B* **54**, R17265–R17268 (1996).
- ¹³²X. Dai, S. Y. Savrasov, G. Kotliar, A. Miglieri, H. Ledbetter, and E. Abrahams, "Calculated phonon spectra of plutonium at high temperatures," *Science* **300**, 953–955 (2003).
- ¹³³S. Mandal, K. Haule, K. M. Rabe, and D. Vanderbilt, "Systematic beyond-DFT study of binary transition metal oxides," *npj Comput. Mater.* **5**, 115 (2019); [arXiv:1907.10498](https://arxiv.org/abs/1907.10498).
- ¹³⁴A. K. McMahan, K. Held, and R. T. Scalettar, "Thermodynamic and spectral properties of compressed Ce calculated using a combined local-density approximation and dynamical mean-field theory," *Phys. Rev. B* **67**, 075108 (2003).
- ¹³⁵A. I. Lichtenstein, M. I. Katsnelson, and G. Kotliar, "Finite-temperature magnetism of transition metals: An *ab initio* dynamical mean-field theory," *Phys. Rev. Lett.* **87**, 067205 (2001).
- ¹³⁶C. Weber, C.-H. Yee, K. Haule, and G. Kotliar, "Scaling of the transition temperature of hole-doped cuprate superconductors with the charge-transfer energy," *Europhys. Lett.* **100**, 37001 (2011); [arXiv:1108.3028](https://arxiv.org/abs/1108.3028).
- ¹³⁷C.-H. Yee, T. Birol, and G. Kotliar, "Guided design of copper oxysulfide superconductors," *Europhys. Lett.* **111**, 17002 (2015).
- ¹³⁸Z. P. Yin and G. Kotliar, "Rational material design of mixed-valent high- T_c superconductors," *Europhys. Lett.* **101**, 27002 (2013).
- ¹³⁹C.-J. Kang, T. Birol, and G. Kotliar, "Phase stability and large in-plane resistivity anisotropy in the 112-type iron-based superconductor $\text{Ca}_{1-x}\text{La}_x\text{FeAs}_2$," *Phys. Rev. B* **95**, 014511 (2017).
- ¹⁴⁰F. Lechermann, S. Biermann, and A. Georges, "Interorbital charge transfers and Fermi-surface deformations in strongly correlated metals: Models, BaVS_3 and Na_xCoO_2 ," *Prog. Theor. Phys. Suppl.* **160**, 233–252 (2005).
- ¹⁴¹J. H. Shim, K. Haule, and G. Kotliar, "Modeling the localized-to-itinerant electronic transition in the heavy fermion system CeIrIn_5 ," *Science* **318**, 1615–1617 (2007).
- ¹⁴²R. Arita, K. Held, A. V. Lukoyanov, and V. I. Anisimov, "Doped Mott insulator as the origin of heavy-fermion behavior in LiV_2O_4 ," *Phys. Rev. Lett.* **98**, 166402 (2007).
- ¹⁴³K. Haule, C.-H. Yee, and K. Kim, "Dynamical mean-field theory within the full-potential methods: Electronic structure of CeIrIn_5 , CeCoIn_5 , and CeRhIn_5 ," *Phys. Rev. B* **81**, 195107 (2010).
- ¹⁴⁴H. C. Choi, B. I. Min, J. H. Shim, K. Haule, and G. Kotliar, "Temperature-dependent Fermi surface evolution in heavy fermion CeIrIn_5 ," *Phys. Rev. Lett.* **108**, 016402 (2012).
- ¹⁴⁵C. Weber, D. D. O'Regan, N. D. M. Hine, P. B. Littlewood, G. Kotliar, and M. C. Payne, "Importance of many-body effects in the kernel of hemoglobin for ligand binding," *Phys. Rev. Lett.* **110**, 106402 (2013).
- ¹⁴⁶C. Weber, D. J. Cole, D. D. O'Regan, and M. C. Payne, "Renormalization of myoglobin-ligand binding energetics by quantum many-body effects," *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5790–5795 (2014).
- ¹⁴⁷M. Ali al-Badri, E. Linscott, A. Georges, D. J. Cole, and C. Weber, "Superexchange mechanism and quantum many body excitations in the archetypal di-Cu oxo-bridge," *Commun. Phys.* **3**, 4 (2020).
- ¹⁴⁸J. Aarons, "A new CASTEP and ONETEP geometry optimiser," 2011, www.hector.ac.uk/cse/distributedcse/reports/castep-geom; this work was funded under the embedded CSE programme of the HECTOR UK National Supercomputing Service, <http://www.hector.ac.uk>.
- ¹⁴⁹J. Andzelm, R. D. King-Smith, and G. Fitzgerald, "Geometry optimization of solids using delocalized internal coordinates," *Chem. Phys. Lett.* **335**, 321–326 (2001).
- ¹⁵⁰B. G. Pfommer, M. Côté, S. G. Louie, and M. L. Cohen, "Relaxation of crystals with the quasi-Newton method," *J. Comput. Phys.* **131**, 233–240 (1997).
- ¹⁵¹L. Mones, J. Aarons, M. I. J. Probert, N. D. M. Hine, C. Ortner, and J. R. Kermode, "Preconditioned geometry optimisers for the CASTEP and ONETEP codes," 2018, <https://www.archer.ac.uk/community/eCSE/eCSE11-07/eCSE11-07.php>; this work was funded under the embedded CSE programme of the ARCHER UK National Supercomputing Service, <http://www.archer.ac.uk>.

- ¹⁵²D. Packwood, J. Kermode, L. Mones, N. Bernstein, J. Woolley, N. Gould, C. Ortner, and G. Csányi, "A universal preconditioner for simulating condensed phase materials," *J. Chem. Phys.* **144**, 164109 (2016).
- ¹⁵³D. Alfè, "PHON: A program to calculate phonons using the small displacement method," *Comput. Phys. Commun.* **180**, 2622–2633 (2009).
- ¹⁵⁴A. Togo and I. Tanaka, "First principles phonon calculations in materials science," *Scr. Mater.* **108**, 1–5 (2015).
- ¹⁵⁵F. Corsetti, "On the properties of point defects in silicon nanostructures from *ab initio* calculations," Ph.D. thesis, Imperial College London, 2012, <http://hdl.handle.net/10044/1/9754>.
- ¹⁵⁶G. Kresse and J. Hafner, "*Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium," *Phys. Rev. B* **49**, 14251–14269 (1994).
- ¹⁵⁷M. E. Tuckerman, "*Ab initio* molecular dynamics: Basic concepts, current trends and novel applications," *J. Phys.: Condens. Matter* **14**, R1297–R1355 (2002).
- ¹⁵⁸P. Carloni, U. Rothlisberger, and M. Parrinello, "The role and perspective of *ab initio* molecular dynamics in the study of biological systems," *Acc. Chem. Res.* **35**, 455–464 (2002).
- ¹⁵⁹M.-P. Gaigeot, "Theoretical spectroscopy of floppy peptides at room temperature. A DFTMD perspective: Gas and aqueous phase," *Phys. Chem. Chem. Phys.* **12**, 3336–3359 (2010).
- ¹⁶⁰R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
- ¹⁶¹D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, 2014).
- ¹⁶²M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln," *Ann. Phys.* **389**, 457–484 (1927).
- ¹⁶³M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, "Iterative minimization techniques for *ab initio* total-energy calculations: Molecular dynamics and conjugate gradients," *Rev. Mod. Phys.* **64**, 1045–1097 (1992).
- ¹⁶⁴L. Verlet, "Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules," *Phys. Rev.* **159**, 98–103 (1967).
- ¹⁶⁵H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," *J. Chem. Phys.* **72**, 2384–2393 (1980).
- ¹⁶⁶M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, England, 1996).
- ¹⁶⁷G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics," *J. Chem. Phys.* **97**, 2635–2643 (1992).
- ¹⁶⁸H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *J. Chem. Phys.* **81**, 3684–3690 (1984).
- ¹⁶⁹G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *J. Chem. Phys.* **126**, 014101 (2007).
- ¹⁷⁰E. Artacho and D. D. O'Regan, "Quantum mechanics in an evolving Hilbert space," *Phys. Rev. B* **95**, 115155 (2017).
- ¹⁷¹A. M. N. Niklasson, P. Steneteg, A. Odell, N. Bock, M. Challacombe, C. J. Tymczak, E. Holmström, G. Zheng, and V. Weber, "Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation," *J. Chem. Phys.* **130**, 214109 (2009).
- ¹⁷²A. Albaugh, O. Demerdash, and T. Head-Gordon, "An efficient and stable hybrid extended Lagrangian/self-consistent field scheme for solving classical mutual induction," *J. Chem. Phys.* **143**, 174104 (2015).
- ¹⁷³V. Vitale, J. Dziedzic, A. Albaugh, A. M. N. Niklasson, T. Head-Gordon, and C.-K. Skylaris, "Performance of extended Lagrangian schemes for molecular dynamics simulations with classical polarizable force fields and density functional theory," *J. Chem. Phys.* **146**, 124115 (2017).
- ¹⁷⁴H. Jónsson, G. Mills, and K. W. Jacobsen, "Nudged elastic band method for finding minimum energy paths of transitions," in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, 1997), pp. 385–404.
- ¹⁷⁵L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, "Challenges in large scale quantum mechanical calculations," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **7**(1), e1290 (2016).
- ¹⁷⁶G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," *J. Chem. Phys.* **113**, 9978 (2000).
- ¹⁷⁷D. Sheppard, R. Terrell, and G. Henkelman, "Optimization methods for finding minimum energy paths," *J. Chem. Phys.* **128**, 134106 (2008).
- ¹⁷⁸G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," *J. Chem. Phys.* **113**, 9901 (2000).
- ¹⁷⁹A. Klamt and G. Schüürmann, "COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient," *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
- ¹⁸⁰S. Miertuš, E. Scrocco, and J. Tomasi, "Electrostatic interaction of a solute with a continuum. A direct utilization of *ab initio* molecular potentials for the prevision of solvent effects," *Chem. Phys.* **55**, 117–129 (1981).
- ¹⁸¹J.-L. Fattebert and F. Gygi, "Density functional theory for efficient *ab initio* molecular dynamics simulations in solution," *J. Comput. Chem.* **23**, 662–666 (2002).
- ¹⁸²D. A. Scherlis, J.-L. Fattebert, F. Gygi, M. Cococcioni, and N. Marzari, "A unified electrostatic and cavitation model for first-principles molecular dynamics in solution," *J. Chem. Phys.* **124**, 074103 (2006).
- ¹⁸³M. Cococcioni, F. Mauri, G. Ceder, and N. Marzari, "Electronic-enthalpy functional for finite systems under pressure," *Phys. Rev. Lett.* **94**, 145501 (2005).
- ¹⁸⁴J. Dziedzic, H. H. Helal, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, "Minimal parameter implicit solvent model for *ab initio* electronic-structure calculations," *Europhys. Lett.* **95**, 43001 (2011).
- ¹⁸⁵A. Nicholls, D. L. Mobley, J. P. Guthrie, J. D. Chodera, C. I. Bayly, M. D. Cooper, and V. S. Pande, "Predicting small-molecule solvation free energies: An informal blind test for computational chemistry," *J. Med. Chem.* **51**, 769–779 (2008).
- ¹⁸⁶J. P. Guthrie, "A blind challenge for computational solvation free energies: Introduction and overview," *J. Phys. Chem. B* **113**, 4501–4507 (2009).
- ¹⁸⁷A. V. Marenich, C. P. Kelly, J. D. Thompson, G. D. Hawkins, C. C. Chambers, D. J. Giesen, P. Winget, C. J. Cramer, and D. G. Truhlar, Minnesota Solvation Database, version 2009, 2009, <https://comp.chem.umn.edu/mnsol/>.
- ¹⁸⁸A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," *J. Phys. Chem. B* **113**, 6378–6396 (2009).
- ¹⁸⁹R. Salomon-Ferrer, D. A. Case, and R. C. Walker, "An overview of the Amber biomolecular simulation package," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **3**, 198–210 (2013).
- ¹⁹⁰J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865 (1996).
- ¹⁹¹Y. Zhao, N. E. Schultz, and D. G. Truhlar, "Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions," *J. Chem. Theory Comput.* **2**, 364 (2006).
- ¹⁹²G. Lever, D. J. Cole, N. D. M. Hine, P. D. Haynes, and M. C. Payne, "Electrostatic considerations affecting the calculated HOMO-LUMO gap in protein molecules," *J. Phys.: Condens. Matter* **25**, 152101 (2013).
- ¹⁹³J. Dziedzic, A. Bhandari, L. Anton, C. Peng, J. Womack, M. Famili, D. Kramer, and C.-K. Skylaris, "A practical approach to large scale electronic structure calculations in electrolyte solutions via continuum-embedded linear-scaling DFT," *J. Phys. Chem. C* **124**(14), 7860–7872 (2020).
- ¹⁹⁴D. J. Cole and N. D. M. Hine, "Applications of large-scale density functional theory in biology," *J. Phys.: Condens. Matter* **28**, 393001 (2016).
- ¹⁹⁵P. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," *J. Phys. Chem. B* **107**, 5933–5947 (2003).

- ¹⁹⁶P. Ren, C. Wu, and J. W. Ponder, "Polarizable atomic multipole-based molecular mechanics for organic molecules," *J. Chem. Theory Comput.* **7**, 3143–3161 (2011).
- ¹⁹⁷Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable atomic multipole-based AMOEBA force field for proteins," *J. Chem. Theory Comput.* **9**, 4046–4063 (2013).
- ¹⁹⁸J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general Amber force field," *J. Comput. Chem.* **25**, 1157–1174 (2004).
- ¹⁹⁹A. J. Stone, "Distributed multipole analysis, or how to describe a molecular charge distribution," *Chem. Phys. Lett.* **83**, 233–239 (1981).
- ²⁰⁰A. J. Stone and M. Alderton, "Distributed multipole analysis," *Mol. Phys.* **56**, 1047–1064 (1985).
- ²⁰¹V. Vitale, J. Dziedzic, S. M.-M. Dubois, H. Fangohr, and C.-K. Skylaris, "Anharmonic infrared spectroscopy through the Fourier transform of time correlation function formalism in ONETEP," *J. Chem. Theory Comput.* **11**, 3321–3332 (2015).
- ²⁰²J. Dziedzic, Y. Mao, Y. Shao, J. Ponder, T. Head-Gordon, M. Head-Gordon, and C.-K. Skylaris, "TINKTEP: A fully self-consistent, mutually polarizable QM/MM approach based on the AMOEBA force field," *J. Chem. Phys.* **145**, 124106 (2016).
- ²⁰³J. Dziedzic, T. Head-Gordon, M. Head-Gordon, and C.-K. Skylaris, "Mutually polarizable QM/MM model with *in situ* optimized localized basis functions," *J. Chem. Phys.* **150**, 074103 (2019).
- ²⁰⁴P. Huang and E. A. Carter, "Advances in correlated electronic structure methods for solids, surfaces, and nanostructures," *Annu. Rev. Phys. Chem.* **59**, 261–290 (2008).
- ²⁰⁵A. S. P. Gomes and C. R. Jacob, "Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems," *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **108**, 222–277 (2012).
- ²⁰⁶M. E. Fornace, J. Lee, K. Miyamoto, F. R. Manby, and T. F. Miller, "Embedded mean-field theory," *J. Chem. Theory Comput.* **11**, 568–580 (2015).
- ²⁰⁷J. C. A. Prentice, R. J. Charlton, A. A. Mostofi, and P. D. Haynes, "Combining embedded mean-field theory with linear-scaling density-functional theory," *J. Chem. Theory Comput.* **16**, 354 (2020).
- ²⁰⁸F. Ding, F. R. Manby, and T. F. Miller, "Embedded mean-field theory with block-orthogonalized partitioning," *J. Chem. Theory Comput.* **13**, 1605 (2017).
- ²⁰⁹F. Ding, T. Tsuchiya, F. R. Manby, and T. F. Miller, "Linear-response time-dependent embedded mean-field theory," *J. Chem. Theory Comput.* **13**, 4216 (2017).
- ²¹⁰G. Makov and M. C. Payne, "Periodic boundary conditions in *ab initio* calculations," *Phys. Rev. B* **51**, 4014 (1995).
- ²¹¹N. D. M. Hine, J. Dziedzic, P. D. Haynes, and C.-K. Skylaris, "Electrostatic interactions in finite systems treated with periodic boundary conditions: Application to linear-scaling density functional theory," *J. Chem. Phys.* **135**, 204103 (2011).
- ²¹²M. R. Jarvis, I. D. White, R. W. Godby, and M. C. Payne, "Supercell technique for total-energy calculations of finite charged and polar systems," *Phys. Rev. B* **56**, 14972–14978 (1997).
- ²¹³C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, "Exact Coulomb cutoff technique for supercell calculations," *Phys. Rev. B* **73**, 205119 (2006).
- ²¹⁴G. J. Martyna and M. E. Tuckerman, "A reciprocal space based method for treating long range interactions in *ab initio* and force-field-based calculations in clusters," *J. Chem. Phys.* **110**, 2810 (1999).
- ²¹⁵T. Sohler, M. Calandra, and F. Mauri, "Density functional perturbation theory for gated two-dimensional heterostructures: Theoretical developments and application to flexural phonons in graphene," *Phys. Rev. B* **96**, 075448 (2017).
- ²¹⁶J. Dziedzic, S. J. Fox, T. Fox, C. S. Tautermann, and C.-K. Skylaris, "Large-scale DFT calculations in implicit solvent—A case study on the T4 lysozyme L99A/M102Q protein," *Int. J. Quantum Chem.* **113**, 771–785 (2013).
- ²¹⁷C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, "Using ONETEP for accurate and efficient density functional calculations," *J. Phys.: Condens. Matter* **17**, 5757–5769 (2005).
- ²¹⁸L. E. Ratcliff, N. D. M. Hine, and P. D. Haynes, "Calculating optical absorption spectra for large systems using linear-scaling density functional theory," *Phys. Rev. B* **84**, 165131 (2011).
- ²¹⁹L. E. Ratcliff and P. D. Haynes, "Ab initio calculations of the optical absorption spectra of C₆₀-conjugated polymer hybrids," *Phys. Chem. Chem. Phys.* **15**, 13024–13031 (2013).
- ²²⁰A. S. Fokas, D. J. Cole, N. D. M. Hine, S. A. Wells, M. C. Payne, and A. W. Chin, "Evidence of correlated static disorder in the Fenna-Matthews-Olson complex," *J. Phys. Chem. Lett.* **8**, 2350 (2017).
- ²²¹D. J. Cole, A. W. Chin, N. D. M. Hine, P. D. Haynes, and M. C. Payne, "Toward ab initio optical spectroscopy of the Fenna-Matthews-Olson complex," *J. Phys. Chem. Lett.* **4**, 4206 (2013).
- ²²²T. J. Zuehlsdorff, N. D. M. Hine, M. C. Payne, and P. D. Haynes, "Linear-scaling time-dependent density-functional theory beyond the Tamm-Dancoff approximation: Obtaining efficiency and accuracy with *in situ* optimised local orbitals," *J. Chem. Phys.* **143**, 204107 (2015).
- ²²³E. W. Tait, L. E. Ratcliff, M. C. Payne, P. D. Haynes, and N. D. M. Hine, "Simulation of electron energy loss spectra of nanomaterials with linear-scaling density functional theory," *J. Phys.: Condens. Matter* **28**, 195202 (2016).
- ²²⁴L. E. Ratcliff, G. J. Conduit, N. D. M. Hine, and P. D. Haynes, "Band structure interpolation using optimized local orbitals from linear-scaling density functional theory," *Phys. Rev. B* **98**, 125123 (2018).
- ²²⁵C. J. Pickard and M. C. Payne, "Second-order $\mathbf{k} \cdot \mathbf{p}$ perturbation theory with Vanderbilt pseudopotentials and plane waves," *Phys. Rev. B* **62**, 4383 (2000).
- ²²⁶M. E. Casida, in *Recent Advances in Density Functional Methods, Part I*, edited by D. P. Chong (World Scientific, Singapore, 1994), p. 155.
- ²²⁷G. Onida, L. Reining, and A. Rubio, "Electronic excitations: Density-functional versus many-body Green's-function approaches," *Rev. Mod. Phys.* **74**, 601 (2002).
- ²²⁸E. V. Tsiper, "A classical mechanics technique for quantum linear response," *J. Phys. B: At. Mol., Opt. Phys.* **34**, L401 (2019).
- ²²⁹S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the Tamm-Dancoff approximation," *Chem. Phys. Lett.* **314**, 291–299 (1999).
- ²³⁰T. J. Zuehlsdorff, *Computing the Optical Properties of Large Systems* (Springer International Publishing, Switzerland, 2015), <http://hdl.handle.net/10044/1/29208>.
- ²³¹T. J. Zuehlsdorff, P. D. Haynes, F. Hanke, M. C. Payne, and N. D. M. Hine, "Solvent effects on electronic excitations of an organic chromophore," *J. Chem. Theory Comput.* **12**, 1853–1861 (2016).
- ²³²A. Dreuw and M. Head-Gordon, "Failure of time-dependent density functional theory for long-range charge-transfer excited states: The zincbacteriochlorin-bacteriochlorophyll and bacteriochlorophyll-spheroidene complexes," *J. Am. Chem. Soc.* **126**, 4007–4016 (2004).
- ²³³T. J. Zuehlsdorff, P. D. Haynes, M. C. Payne, and N. D. M. Hine, "Predicting solvatochromic shifts and colours of a solvated organic dye: The example of Nile red," *J. Chem. Phys.* **146**, 124504 (2017).
- ²³⁴M. A. P. Turner, M. D. Horbury, V. G. Stavros, and N. D. M. Hine, "Determination of secondary species in solution through pump-selective transient absorption spectroscopy and explicit-solvent TDDFT," *J. Phys. Chem. A* **123**, 873–880 (2019).
- ²³⁵F. Furche and R. Ahlrichs, "Adiabatic time-dependent density functional methods for excited state properties," *J. Chem. Phys.* **117**, 7433–7447 (2002).
- ²³⁶R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Dev.* **1**, 223–231 (1957).
- ²³⁷R. Landauer, "Electrical resistance of disordered one-dimensional lattices," *Philos. Mag.* **21**, 863–867 (1970).
- ²³⁸M. Büttiker, "Four-terminal phase-coherent conductance," *Phys. Rev. Lett.* **57**, 1761–1764 (1986).
- ²³⁹R. A. Bell, S. M.-M. Dubois, M. C. Payne, and A. A. Mostofi, "Electronic transport calculations in the ONETEP code: Implementation and applications," *Comput. Phys. Commun.* **193**, 78–88 (2015).

- ²⁴⁰E. M. Godfrin, "A method to compute the inverse of an n-block tridiagonal quasi-Hermitian matrix," *J. Phys.: Condens. Matter* **3**, 7843 (1991).
- ²⁴¹D. E. Petersen, H. H. B. Sørensen, P. C. Hansen, S. Skelboe, and K. Stokbro, "Block tridiagonal matrix inversion and fast transmission calculations," *J. Comput. Phys.* **227**, 3174–3190 (2008).
- ²⁴²M. Paulsson and M. Brandbyge, "Transmission eigenchannels from nonequilibrium Green's functions," *Phys. Rev. B* **76**, 115117 (2007).
- ²⁴³A. J. Morris, R. J. Nicholls, C. J. Pickard, and J. R. Yates, "OptaDOS: A tool for obtaining density of states, core-level and optical spectra from electronic structure codes," *Comput. Phys. Commun.* **185**, 1477–1485 (2014).
- ²⁴⁴J. Aarons, L. G. Verga, N. D. M. Hine, and C.-K. Skylaris, "Atom-projected and angular momentum resolved density of states in the ONETEP code," *Electron. Struct.* **1**, 035002 (2019).
- ²⁴⁵D. Sanchez-Portal, E. Artacho, and J. M. Soler, "Projection of plane-wave calculations into atomic orbitals," *Solid State Commun.* **95**, 685 (1995).
- ²⁴⁶E. W. Tait, "Linear-scaling density functional theory and theoretical electron energy loss spectroscopy investigations of surfaces and defects in nanomaterials," Ph.D. thesis, University of Cambridge, United Kingdom, 2019, <https://www.repository.cam.ac.uk/handle/1810/292108>.
- ²⁴⁷R. J. Nicholls, A. J. Morris, C. J. Pickard, and J. R. Yates, "OptaDOS: A new tool for EELS calculations," *J. Phys.: Conf. Ser.* **371**, 012062 (2012).
- ²⁴⁸R. Zwanzig, "Time-correlation functions and transport coefficients in statistical mechanics," *Annu. Rev. Phys. Chem.* **16**, 67–102 (1965).
- ²⁴⁹R. Kubo, "The fluctuation-dissipation theorem," *Rep. Prog. Phys.* **29**, 255–284 (1966).
- ²⁵⁰R. Rein, "On physical properties and interactions of polyatomic molecules: With application to molecular recognition in biology," in *Advances in Quantum Chemistry*, edited by P.-O. Löwdin (Academic Press, 1973), Vol. 7, pp. 335–396.
- ²⁵¹A. J. Stone, "GDMA: Distributed multipoles from Gaussian98 wavefunctions," Technical Report, University of Cambridge, 1998, available upon request at <https://www.cl.cam.ac.uk/techreports/>.
- ²⁵²A. J. Stone, "Distributed multipole analysis: Stability for large basis sets," *J. Chem. Theory Comput.* **1**, 1128–1132 (2005).
- ²⁵³M. J. S. Phipps, T. Fox, C. S. Tautermann, and C.-K. Skylaris, "Energy decomposition analysis approaches and their evaluation on prototypical protein-drug interaction patterns," *Chem. Soc. Rev.* **44**, 3177–3211 (2015).
- ²⁵⁴M. J. S. Phipps, T. Fox, C. S. Tautermann, and C.-K. Skylaris, "Energy decomposition analysis based on absolutely localised molecular orbitals for large-scale density functional theory calculations in drug design," *J. Chem. Theory Comput.* **12**, 3135–3148 (2016).
- ²⁵⁵R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, and M. Head-Gordon, "Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals," *J. Phys. Chem. A* **111**, 8753–8765 (2007).
- ²⁵⁶P. Su and H. Li, "Energy decomposition analysis of covalent bonds and intermolecular interactions," *J. Chem. Phys.* **131**, 014102 (2009).
- ²⁵⁷M. J. S. Phipps, T. Fox, C. S. Tautermann, and C.-K. Skylaris, "Intuitive density functional theory-based energy decomposition analysis for protein-ligand interactions," *J. Chem. Theory Comput.* **13**, 1837–1850 (2017).
- ²⁵⁸A. E. Reed, L. A. Curtiss, and F. Weinhold, "Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint," *Chem. Rev.* **88**, 899–926 (1988).
- ²⁵⁹E. D. Glendening, J. K. Badenhop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, and F. Weinhold, NBO 5, Theoretical Chemistry Institute, University of Wisconsin, Madison, 2001, <http://nbo6.chem.wisc.edu/>.
- ²⁶⁰L. P. Lee, D. J. Cole, M. C. Payne, and C.-K. Skylaris, "Natural bond orbital analysis in the ONETEP code: Applications to large protein systems," *J. Comput. Chem.* **34**, 429–444 (2013).
- ²⁶¹G. Lever, D. J. Cole, R. Lonsdale, K. E. Ranaghan, D. J. Wales, A. J. Mulholland, C.-K. Skylaris, and M. C. Payne, "Large-scale density functional theory transition state searching in enzymes," *J. Phys. Chem. Lett.* **5**, 3614–3619 (2014).
- ²⁶²P. Bultinck, C. Van Alsenoy, P. W. Ayers, and R. Carbó-Dorca, "Critical analysis and extension of the Hirshfeld atoms in molecules," *J. Chem. Phys.* **126**, 144111 (2007).
- ²⁶³T. C. Lillestolen and R. J. Wheatley, "Atomic charge densities generated using an iterative stockholder procedure," *J. Chem. Phys.* **131**, 144101 (2009).
- ²⁶⁴T. A. Manz and D. S. Sholl, "Chemically meaningful atomic charges that reproduce the electrostatic potential in periodic and nonperiodic materials," *J. Chem. Theory Comput.* **6**, 2455–2468 (2010).
- ²⁶⁵T. A. Manz and D. S. Sholl, "Improved atoms-in-molecule charge partitioning functional for simultaneously reproducing the electrostatic potential and chemical states in periodic and nonperiodic materials," *J. Chem. Theory Comput.* **8**, 2844–2867 (2012).
- ²⁶⁶T. A. Manz and N. G. Limas, "Introducing DDEC6 atomic population analysis: Part 1. Charge partitioning theory and methodology," *RSC Adv.* **6**, 47771–47801 (2016).
- ²⁶⁷L. P. Lee, D. J. Cole, C.-K. Skylaris, W. L. Jorgensen, and M. C. Payne, "Polarized protein-specific charges from atoms-in-molecule electron density partitioning," *J. Chem. Theory Comput.* **9**, 2981–2991 (2013).
- ²⁶⁸L. P. Lee, N. G. Limas, D. J. Cole, M. C. Payne, C.-K. Skylaris, and T. A. Manz, "Expanding the scope of density derived electrostatic and chemical charge partitioning to thousands of atoms," *J. Chem. Theory Comput.* **10**, 5377–5390 (2014).
- ²⁶⁹J. T. Horton, A. E. A. Allen, L. S. Dodda, and D. J. Cole, "QUBEKit: Automating the derivation of force field parameters from quantum mechanics," *J. Chem. Inf. Model.* **59**, 1366–1381 (2019).
- ²⁷⁰D. J. Cole, J. Z. Vilseck, J. Tirado-Rives, M. C. Payne, and W. L. Jorgensen, "Biomolecular force field parameterization via atoms-in-molecule electron density partitioning," *J. Chem. Theory Comput.* **12**, 2312–2323 (2016).
- ²⁷¹A. Tkatchenko and M. Scheffler, "Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data," *Phys. Rev. Lett.* **102**, 073005–073009 (2009).
- ²⁷²T. A. Manz, T. Chen, D. J. Cole, N. G. Limas, and B. Fiszbein, "New scaling relations to compute atom-in-material polarizabilities and dispersion coefficients: Part 1. Theory and accuracy," *RSC Adv.* **9**, 19297–19324 (2019).
- ²⁷³A. E. A. Allen, M. J. Robertson, M. C. Payne, and D. J. Cole, "Development and validation of the quantum mechanical bespoke protein force field," *ACS Omega* **4**, 14537–14550 (2019).
- ²⁷⁴D. J. Cole, I. Cabeza de Vaca, and W. L. Jorgensen, "Computation of protein ligand binding free energies using quantum mechanical bespoke force fields," *Med. Chem. Commun.* **10**, 1116–1120 (2019).
- ²⁷⁵D. J. Cole, J. T. Horton, L. Nelson, and V. Kurdekar, "The future of force fields in computer aided drug design," *Future Med. Chem.* **11**, 2359–2363 (2019).
- ²⁷⁶J. T. Horton, A. E. A. Allen, and D. J. Cole, "Modelling flexible protein-ligand binding in p38 α MAP kinase using the QUBE force field," *Chem. Commun.* **56**, 932 (2020).
- ²⁷⁷A. D. Becke and K. E. Edgecombe, "A simple measure of electron localization in atomic and molecular systems," *J. Chem. Phys.* **92**, 5397–5403 (1990).
- ²⁷⁸A. D. Becke and H. L. Schmider, "Chemical content of the kinetic energy density," *J. Mol. Struct. Theochem* **527**, 51–61 (2000).
- ²⁷⁹A. Savin, O. Jepsen, J. Flad, O. K. Andersen, H. Preuss, and H. G. von Schnering, "Electron localization in solid-state structures of the elements: The diamond structure," *Angew. Chem., Int. Ed. Engl.* **31**, 187–188 (1992).
- ²⁸⁰N. R. C. Corsini, Y. Zhang, W. R. Little, A. Karatutlu, O. Ersoy, P. D. Haynes, C. Molteni, N. D. M. Hine, I. Hernandez, J. Gonzalez, F. Rodriguez, V. V. Brazhkin, and A. Sapelkin, "Pressure-induced amorphization and a new high density amorphous metallic phase in matrix-free Ge nanoparticles," *Nano Lett.* **15**, 7334–7340 (2015).
- ²⁸¹N. R. C. Corsini, N. D. M. Hine, P. D. Haynes, and C. Molteni, "Unravelling the roles of sizes, ligands, and pressure in the piezochromic properties of CdS nanocrystals," *Nano Lett.* **17**, 1042 (2017).
- ²⁸²P. W. Avraam, N. D. M. Hine, P. Tangney, and P. D. Haynes, "Factors influencing the distribution of charge in polar nanocrystals," *Phys. Rev. B* **83**, 241402 (2011).
- ²⁸³P. W. Avraam, N. D. M. Hine, P. Tangney, and P. D. Haynes, "Fermi-level pinning can determine polarity in semiconductor nanorods," *Phys. Rev. B* **85**, 115404 (2012).

- ²⁸⁴R. A. Bell, M. C. Payne, and A. A. Mostofi, "Improving the conductance of carbon nanotube networks through resonant momentum exchange," *Phys. Rev. B* **89**, 245426 (2014).
- ²⁸⁵U. Fano, "Effects of configuration interaction on intensities and phase shifts," *Phys. Rev.* **124**, 1866–1878 (1961).
- ²⁸⁶R. A. Bell, *First-Principles Conductance between Carbon Nanotubes* (Springer International Publishing, Switzerland, 2015), <https://www.repository.cam.ac.uk/handle/1810/290758>.
- ²⁸⁷T. Ellaby, L. Briquet, M. Sarwar, D. Thompsett, and C.-K. Skylaris, "Modification of O and CO binding on Pt nanoparticles due to electronic and structural effects of titania supports," *J. Chem. Phys.* **151**, 114702 (2019).
- ²⁸⁸N. R. C. Corsini, A. Greco, N. D. M. Hine, C. Molteni, and P. D. Haynes, "Simulations of nanocrystals under pressure: Combining electronic enthalpy and linear-scaling density-functional theory," *J. Chem. Phys.* **139**, 084117 (2013).
- ²⁸⁹See <https://www.3dsbiovia.com> for Dassault systèmes biovia; accessed 05 February 2020.
- ²⁹⁰See <https://www.ccp9.ac.uk> for Collaboration computation project for the study of the electronic structure of condensed matter (ccp9); accessed 05 February 2020.
- ²⁹¹See <https://www.onetep.org> for Onetep website; accessed 05 February 2020.
- ²⁹²See <http://ukcp.ac.uk> for the United Kingdom Car-Parrinello consortium (UKCP); accessed 05 February 2020.
- ²⁹³See <http://psi-k.net> for Psi-k network; accessed 05 February 2020.
- ²⁹⁴See <https://molssi.org> for the Molecular Sciences Software Institute (MolSSI); accessed 05 February 2020.
- ²⁹⁵K. Wilkinson and C.-K. Skylaris, "Porting ONETEP to graphical processing unit-based coprocessors. 1. FFT box operations," *J. Comput. Chem.* **34**, 2446 (2013).