**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

Local approaches for electric dipole moments in periodic systems and their application to real-time time-dependent density functional theory

Schreder, Lukas ; Lubner, Sandra

Abstract: Within periodic boundary conditions, the traditional quantum mechanical position operator is ill-defined, necessitating the use of alternative methods, most commonly the Berry phase formulation in the modern theory of polarization. Since any information about local properties is lost in this change of framework, the Berry phase formulation can only determine the total electric polarization of a system. Previous approaches toward recovering local electric dipole moments have been based on applying the conventional dipole moment operator to the centers of maximally localized Wannier functions (MLWFs). Recently, another approach to local electric dipole moments has been demonstrated in the field of subsystem density functional theory (DFT) embedding. We demonstrate in this work that this approach, aside from its use in ground state DFT-based molecular dynamics, can also be applied to obtain electric dipole moments during real-time propagated time-dependent DFT (RT-TDDFT). Moreover, we present an analogous approach to obtain local electric dipole moments from MLWFs, which enables subsystem analysis in cases where DFT embedding is not applicable. The techniques were implemented in the quantum chemistry software CP2K for the mixed Gaussian and plane wave method and applied to cis-diimide and water in the gas phase, cis-diimide in aqueous solution, and a liquid mixture of dimethyl carbonate and ethylene carbonate to obtain absorption and infrared spectra decomposed into localized subsystem contributions.

DOI: <https://doi.org/10.1063/5.0058793>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-215171>

Journal Article

Accepted Version

Originally published at:

Schreder, Lukas; Lubner, Sandra (2021). Local approaches for electric dipole moments in periodic systems and their application to real-time time-dependent density functional theory. *Journal of Chemical Physics*, 155(13):134116.

DOI: <https://doi.org/10.1063/5.0058793>

Local approaches for electric dipole moments in periodic systems and their application to real-time time-dependent density functional theory

Lukas Schreder and Sandra Luber^{a)}

*University of Zürich, Winterthurerstrasse 190, 8057 Zürich,
Switzerland*

(Dated: September 15, 2021)

Within periodic boundary conditions, the traditional quantum mechanical position operator is ill-defined, necessitating the use of alternative methods, most commonly the Berry phase formulation in the modern theory of polarization. Since any information about local properties is lost in this change of framework, the Berry phase formulation can only determine the total electric polarization of a system. Previous approaches towards recovering local electric dipole moments have been based on applying the conventional dipole moment operator to the centers of maximally localized Wannier functions (MLWFs). Recently, another approach to local electric dipole moments has been demonstrated in the field of subsystem density functional theory (DFT) embedding. We demonstrate in this work that this approach, aside from its use in ground state DFT-based molecular dynamics, can also be applied to obtain electric dipole moments during real-time propagated time-dependent DFT (RT-TDDFT). Moreover, we present an analogous approach to obtain local electric dipole moments from MLWFs, which enables subsystem analysis in cases where DFT embedding is not applicable. The techniques were implemented in the quantum chemistry software CP2K for the mixed Gaussian and plane wave method and applied to *cis*-diimide and water in gas phase, *cis*-diimide in aqueous solution and a liquid mixture of dimethyl carbonate and ethylene carbonate to obtain absorption and infrared spectra decomposed into localized subsystem contributions.

^{a)}Email: sandra.luber@chem.uzh.ch

I. INTRODUCTION

One of the most important physical observables to be computed by quantum chemical methods is the electric dipole moment. The electric dipole response of a system to a given perturbation is used to compute spectra which, in a dynamic approach, depend on the time-dependent evolution of the electric polarization, such as absorption and infrared (IR) spectra.^{1,2} Both spectroscopic techniques are also widely used in experimental analysis of chemical systems ranging from isolated gas phase molecules to bulk materials. The expectation values of the electronic contribution to the electric dipole moment in nonperiodic systems are calculated trivially by adapting the conventional electric dipole definition $\mathbf{d} = q\mathbf{r}$, where q is the charge and $\mathbf{r}$ is the position vector, to the quantum mechanical position operator applied to the set of orbitals. As this is a purely one-electron operator, the total electric dipole is strictly the sum of local contributions. Accordingly, local contributions can be separated out exactly, down to single orbitals. Unfortunately, this conventional approach cannot be applied to periodic systems, since the classical position operator is ill-defined. This problem has been resolved by the modern theory of polarization for stationary as well as time-dependent electric fields.³⁻⁷ By reformulating the electric polarization in terms of the Berry phase, the observable is made gauge-independent. It is, however, no longer trivially separable into subsystem contributions. Moreover, it is notoriously difficult to generalize this approach to higher order multipole moments or different observables altogether.⁸

In calculations for periodic systems, the derivation of localized properties requires special care. One widespread approach is localization, which can produce maximally localized Wannier functions (MLWFs) that resemble localized molecular orbitals in the nonperiodic case.⁹⁻¹¹ Another approach has been suggested recently¹² which relies on periodic subsystem density functional theory (DFT) embedding methods to obtain subsystem densities.¹³⁻¹⁸ DFT embedding is first and foremost employed as a cost-saving measure, but can achieve a complete separation of the electronic densities by design using local atomic basis functions. Thus, it lends itself well to subsystem analysis when the error introduced by embedding is reasonably low. Further, the MLWF approach usually requires one localization procedure at each sample point, increasing the computational cost. To its advantage, this variational method does not introduce any additional approximations into the chosen method of electronic structure theory aside from the error made during the localization procedure (which generally relies on approximate diagonalization). Re-

Regardless of the chosen approach, an accurate calculation of a subsystem’s electric dipole moment in periodic boundary conditions (PBC) requires modifications to the Berry phase formalism. One separable approximation which is based on a power series expansion of the many-electron operator has been successfully applied to DFT embedding calculations.¹²

In addition to robust formulations of single point properties, obtaining spectra requires probing nuclear and electron dynamics. Since absorption spectra mostly describe the response to an external electric field in the ultraviolet and visible range of the electromagnetic spectrum, it is dominated by relatively short-lived excited states. Thus, the key to computationally affordable prediction of absorption spectra is to find efficient solutions to the time-dependent Schrödinger equation (TDSE) at fixed molecular geometries. As traditional DFT is constrained to ground state calculations by the underlying Hohenberg-Kohn theorem,¹⁹ it is necessary to employ time-dependent density functional theory (TDDFT) which describes the time evolution of electronic densities by means of the Runge-Gross theorem.²⁰ One active field of TDDFT research is real-time propagated TDDFT (RT-TDDFT)^{21–27}, where the time-evolution propagator is determined numerically between timesteps. The efficient scaling of this approach has led to a significant increase in system sizes that are accessible to TDDFT-based investigations. Moreover, the RT-TDDFT approach has recently been successfully applied to localized orbitals.²⁸ As the absorption spectrum is connected to the electric dipole evolution by what amounts to a simple Fourier transform, it follows naturally from RT-TDDFT simulations. IR spectra on the other hand can usually be simulated to high accuracy by considering molecular dynamics of systems in the electronic ground state. In this case, the IR spectrum can be recovered by Fourier transforming the autocorrelation function of the dipole strength function obtained from e.g. *ab initio* molecular dynamics (AIMD)^{29,30} simulations at the DFT level of theory.

In this work, we demonstrate that the approximate gauge-invariant formulation presented by Lubert¹² is not only applicable to DFT embedding in AIMD, but is equally valid for excited states simulated by RT-TDDFT. Additionally, we show that an analogous approach can be made for MLWFs, which enables subsystem analysis in cases where DFT embedding is expected to perform poorly, e.g. RT-TDDFT of certain condensed systems. To enable a head to head comparison, we implement both methods of subsystem electric dipole analysis into CP2K³¹, and extend their localization routines to enable Time-dependent MLWFs. To this end, we adapt the localization approaches presented by Gygi et al.³² and Yost et al.²⁸ for plane waves used in the Qb@ll fork of

the Qbox^{33,34} code to the mixed Gaussian and plane wave approach used in CP2K. The methods are then applied to *cis*-diimide and water in gas phase and solution as well as a liquid 1:1 mixture of dimethyl carbonate (DMC) and ethylene carbonate (EC) which have seen frequent use as electrolytes in lithium-ion batteries. While they are a popular model system for MLWF-based methods, we did not include crystalline systems into the test set as subsystems will always be connected by covalent bonds. Since the embedding scheme used in this work is not expected to be well suited for subsystems connected via covalent bonds, it would render any comparisons between localization- and embedding-based moments methods naught. Electronic absorption and IR spectra are then collected by means of RT-TDDFT and AIMD respectively.

II. THEORY

In this work, we make use of the Berry phase formulation of polarization introduced by the modern theory of polarization as the reference method for the total electronic part of the electric dipole moment. To separate the contributions according to the chosen subsystems, either localization or embedding methods were employed. Subsystem contributions to the electric dipole were calculated with the approximate gauge-invariant dipole approach.¹² In the equations given, bold lower case letters denote vectors, while bold upper case letters denote matrices. Nuclear contributions to the subsystem moments are treated as conventional point dipoles in all approaches and added to the electronic contribution. In phase-based approaches, this summation is done within the exponent. For RT-TDDFT, this contribution is constant between steps as the nuclei remain stationary. With readability in mind, the following formulae are therefore given for the electronic contributions.

A. Electric polarization in periodic systems

When transitioning from finite to periodic systems, the conventional position operator $\hat{\mathbf{r}} = [\hat{x} \hat{y} \hat{z}]^T$ is ill-defined and thus needs replacement in order to obtain a gauge-invariant definition of the electronic moments. The modern theory of polarization achieves this by describing the electronic polarization $\mathbf{d}$ (electric dipole per unit length) along Cartesian direction α (x-, y-, or z-direction) in terms of a Berry phase, given at the Γ -point of the Brillouin zone of a large supercell and

insulator in a vanishing macroscopic electric field:^{3-5,35-37}

$$d_{\alpha}^{\text{ref}} = \frac{e}{2\pi} \sum_{\beta} h_{\alpha\beta} \text{Im} \ln \det \mathbf{S}_{\alpha} \quad (1)$$

with the elementary charge e and the elements of matrix $\mathbf{S}_{\alpha}$ being

$$S_{\alpha\mu\nu} = \langle \phi_{\mu} | \exp[-i2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}] | \phi_{\nu} \rangle \quad (2)$$

where μ and ν enumerate the occupied orbitals and $\mathbf{h}_{\alpha} = [a_{\alpha} \ b_{\alpha} \ c_{\alpha}]$ is the row of the simulation cell matrix $\mathbf{h} = [\mathbf{a} \ \mathbf{b} \ \mathbf{c}]$ with the primitive lattice vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, along the α direction. Within a Kohn-Sham (KS) framework, ϕ refers to a set of non-interacting KS orbitals. As is the case for nonperiodic systems, this formalism assumes electrically neutral subsystems to uphold translational invariance. While this formulation provides us with a rigorous theory of bulk polarization, it is inherently non-separable into local contributions. Another apparent downside is that there has been no successful attempt to generalize it to higher electric moments so far. Ono et al. recently presented shortcomings of two attempts at generalizing Resta's and Vanderbilt's approach to electric quadrupole moments,^{38,39} casting doubt on whether it can be a feasible approach to localize multipole moments in periodic systems.⁸ Specifically, the proposed methods were found to be in violation of periodicity and to be origin-dependent in a simple quadrupole insulator model.

B. Maximally localized Wannier Functions

In PBC, MO functions $\{\phi\}$ are usually represented in the Bloch form, which is given by a plane wave $e^{i\mathbf{k}\mathbf{r}}$ modulated by $u_{n\mathbf{k}}$, which has the same periodicity as the Hamiltonian, and where n is the MO index, $\mathbf{k}$ is the wave vector, $\mathbf{r}$ the 3-dimensional position vector and t is time:⁴⁰

$$\phi_{n\mathbf{k}}(\mathbf{r}, t) = u_{n\mathbf{k}}(\mathbf{r}, t) e^{i\mathbf{k}\mathbf{r}}. \quad (3)$$

For brevity, we omit the position- and time-dependence of wave functions and electronic densities from this point onward. An alternative, equally valid representation of the electronic structure is given in the Wannier form, transforming each Bloch orbital at $\mathbf{k}$ into a Wannier function w at the real-space lattice vector $\mathbf{R}$. One qualitative advantage of the Wannier form is that

this representation produces a set of functions which, when a suitable spread functional is chosen, resemble molecular orbitals and thus align much closer with the intuition of a trained chemist than Bloch functions. The set of Wannier functions span the same Hilbert space as the Bloch functions while being entirely localized in one lattice cell at lattice vector $\mathbf{R}$. As the transformation from Bloch into Wannier function form is unitary, it does not change any resulting physical observables, and the unitary transformation $\mathbf{U}^L$ is the localization matrix¹¹:

$$w_{n\mathbf{R}} = \sum_m \mathbf{U}_{mn}^L \phi_{m\mathbf{k}}. \quad (4)$$

While there are infinitely many choices for $\mathbf{U}^L$, it can be made unique by imposing that it must be the transformation for which maximal localization is achieved which we call $\mathbf{U}^{ML}$, commonly defined by the minimum of a given spread functional of the N Wannier functions:

$$\mathbf{U}^{ML} \in \left\{ \mathbf{U} : \min_{\forall \mathbf{U} \in \mathbb{C}^{N \times N}} \Omega(\mathbf{U}) \right\}. \quad (5)$$

In finite systems with localized molecular orbitals, one can use the Foster-Boys criterion, which takes the second moment of the position $\langle \mathbf{r}^2 \rangle - \langle \mathbf{r} \rangle^2$ as the spread functional.⁴¹ To be consistent with PBCs and MLWFs however, the functional must be constructed in terms of a Berry phase, as given in its generalized form by Berghold et al.:^{37,42}

$$\Omega = \sum_{\alpha} \sum_{\mu} \omega_{\alpha} \ln \langle w_{\mu} | \exp [i2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}] | w_{\mu} \rangle^2, \quad (6)$$

with a weight factor ω_{α} derived from the simulation cell dimensions. The periodic operator can be conveniently split into sines and cosines as well as the spatial dimensions, as it is commonly done in computational implementations:

$$e^{i2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}} = \cos[2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}] + i \sin[2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}]. \quad (7)$$

This representation results in 6 $\mathbb{R}^{N \times N}$ matrices (or 6 $\mathbb{C}^{N \times N}$ for time-dependent orbitals, as is the case during RT-TDDFT) which is generally a more convenient format for distributed programming compared to the combined monolithic $\mathbb{C}^{3 \times N \times N}$ matrix. It also happens to be the standard data layout used in joint diagonalization schemes, allowing for easy application of established numerical methods.

C. Subsystem DFT-based embedding

In the theoretical framework of subsystem DFT-based embedding (in the latter referred to as embedding for brevity), the electronic density ρ of the total system under study is split into a sum of subsystem electronic densities ρ_m which can be calculated from the KS orbitals using atomic basis functions as

$$\rho = \sum_m \rho_m = \sum_{m\mu} f_{m\mu}^{\text{occ}} |\phi_{m\mu}|^2, \quad (8)$$

where m enumerates the subsystems, μ the orbitals of a subsystem, and $f_{m\mu}^{\text{occ}}$ denotes the occupation of the μ -th KS orbital of the m -th subsystem.^{12–17,43–46} Embedding is most commonly used to save computational cost. The partitioning allows for energy evaluation by calculating the total energy of the system at a computationally more affordable level of theory which is then corrected by more expensive terms for each subsystem. The simulations presented here employed the total functional represented by the subtraction scheme^{12,47}

$$E^{\text{tot}}[\rho] = E^{\text{HK}}[\rho] - \sum_m E^{\text{HK}}[\rho_m] + \sum_m E^{\text{KS}}[\rho_m] \quad (9)$$

with E^{HK} being the orbital-free Hohenberg-Kohn functional

$$E^{\text{HK}}[\rho] = T^{\text{S}}[\rho] + E^{\text{ext}}[\rho] + E^{\text{c}}[\rho] + E^{\text{xc}}[\rho] \quad (10)$$

built from the kinetic energy density functional T^{S} , the external potential interaction E^{ext} , the Coulomb energy E^{c} and the exchange-correlation functional E^{xc} . The electronic energy of the subsystems is calculated within the Kohn-Sham framework, where the kinetic energy density functional $T^{\text{S}}[\rho]$ is replaced by the kinetic energy orbital functional $T^{\text{S}}[\{\phi\}]$

$$E^{\text{KS}}[\rho[\{\phi\}]] = T^{\text{S}}[\{\phi\}] + E^{\text{ext}}[\rho] + E^{\text{c}}[\rho] + E^{\text{xc}}[\rho]. \quad (11)$$

In its current implementation in CP2K, it is assumed that nonlocal pseudopotential contributions are limited to the respective subsystem, and that the MO set of any given subsystem can be adequately represented by combinations of atomic basis functions centered on atoms contained in that subsystem. This embedding keeps the predefined subdensities separate by design and therefore does not produce approximate divisions between subsystems. This is in contrast to localization

methods, which generally rely on joint approximate diagonalization of matrices, resulting in approximate subsystems. In CP2K, this exact separation is achieved by storing the KS matrix in a block-diagonal form where each block corresponds to a subsystem, resulting in linear scaling with respect to the number of subsystems and thus significant savings in computational cost for large systems. Use of DFT embedding for RT-TDDFT has not been explored extensively. However, recent advances have shown that there is interest in accelerating RT-TDDFT calculations and Ehrenfest molecular dynamics (EMD) via DFT embedding approaches.^{48–54}

D. Local electric dipole moments in periodic systems

Local electric dipole moments are usually obtained for periodic systems by using the Wannier functions belonging to a subsystem and taking point charges located at their respective Wannier centers,

$$d_{\alpha}^{\text{WC}} = \sum_m d_{\alpha m}^{\text{WC}} = -e \sum_{m\mu} f_{m\mu}^{\text{occ}} r_{\alpha m\mu}^{\text{WC}}, \quad (12)$$

where m enumerates the subsystems and μ the respective subsystem's Wannier functions, and the α -component of the μ -th Wannier center is given by

$$r_{\alpha m\mu}^{\text{WC}} = -\frac{1}{2\pi} \sum_{\beta} h_{\alpha\beta} \text{Im} \ln \langle w_{m\mu} | \exp[-i2\pi \mathbf{h}_{\alpha}^{-1} \hat{\mathbf{r}}] | w_{m\mu} \rangle. \quad (13)$$

This approach can be valid for periodic systems only if the spread functional used is valid within PBC, which makes Eq. 6 the spread functional of choice. The major disadvantage of localization-based methods is that the localization process is fundamentally numerical as it relies on approximate simultaneous diagonalization. Thus, the separation into subsystems will always carry an error and computational cost as opposed to approaches which use intrinsically separated subsystems such as DFT embedding. Based on DFT embedding, another approach has been presented.¹² By exploiting the identity $\ln \det \mathbf{S}_{\alpha} = \text{Tr} \ln \mathbf{S}_{\alpha}$ and expanding the matrix elements to first order in $\mathbf{h}_{\alpha}^{-1}$ one can approximate the matrix elements of $\mathbf{S}_{\alpha}$ as

$$S_{\alpha\mu\nu} \approx 1 - i2\pi \mathbf{h}_{\alpha}^{-1} \langle \phi_{\mu} | \hat{\mathbf{r}} | \phi_{\nu} \rangle. \quad (14)$$

For this approximation to be valid, it is important that the simulation cell is chosen sufficiently large so that the density matrix and the electronic density of the subsystems do not change with respect to $\mathbf{h}$. As usual, the subsystems must be charge neutral so that translational invariance is preserved. The embedding is expected to introduce significant errors when subsystems are strongly interacting, e.g. via covalent bonds. Thus, subsystems connected by covalent bonds should be analyzed in the localization approach.

Since the approximation made in Eq. 14 is dependent only on the single-electron operators $\hat{\mathbf{r}}_\alpha$, it is separable into subsystem contributions. Using these approximate matrix elements in the Berry phase dipole definition (Eq. 1) and applying it to subsets of Wannier functions and DFT embedding subdensities respectively gives us the subsystem electric dipole expressions used in the simulations:

$$d_\alpha^{\text{emb}} = -\frac{e}{2\pi} \sum_m \sum_\beta h_{\alpha\beta} \text{Im} \ln \exp \left[i2\pi \mathbf{h}_\alpha^{-1} \sum_{\mu, \nu \in m} P_{\mu\nu} \langle \chi_\mu | \hat{\mathbf{r}} | \chi_\nu \rangle \right] \quad (15a)$$

$$d_\alpha^{\text{loc}} = -\frac{e}{2\pi} \sum_m \sum_\beta h_{\alpha\beta} \text{Im} \ln \exp \left[i2\pi \mathbf{h}_\alpha^{-1} \sum_{\mu \in m} \langle w_\mu | \hat{\mathbf{r}} | w_\mu \rangle \right] \quad (15b)$$

where χ_μ and χ_ν are μ -th and ν -th atomic basis functions of subsystem m , $P_{\mu\nu}$ is the corresponding element of the density matrix, and w_μ the μ -th Wannier function of subsystem m . A derivation of Eq. 15b can be obtained from Eq. 1 analogous to Luber¹² by reformulating $\mathbf{S}_\alpha$ in terms of the set of Wannier functions, which also gives Eq. 12 as an intermediate result after the determinant identity. Since this formulation is ambiguous towards the branch of the complex logarithm, it is only defined modulo to its period $e\mathbf{h}_\alpha$, similar to Eq. 1.

III. COMPUTATIONAL METHODOLOGY

All simulations were conducted using a development version of CP2K 7.1.³¹ The double-zeta Goedecker-Teter-Hutter (DVZP-GTH) basis set was used with the corresponding GTH pseudopotentials and the PBE exchange-correlation functional.^{55–57} For calculations on isolated systems, a sufficiently large cubic simulation cell with a length of 10 Å was chosen while the boxes of the liquid phase systems were equilibrated individually to 8.987 Å for solvated *cis*-diimide and 12.03 Å for the DMC/EC mixture. RT-TDDFT calculations were run for at least $500 \hbar E_h^{-1}$ with

a $0.1 \hbar E_h^{-1}$ time step using the enforced time reversible symmetry propagator for time-dependent KS equations⁵⁸. An initial small delta pulse perturbation with a scale parameter of 0.001 was applied before propagation. In CP2K, the effective field strength is dependent on the cell size via $\|E = k \cdot \mathbf{v} 2\pi \hbar^{-1}\|$, with k being the scale parameter and $\mathbf{v}$ being the direction, resulting in 3.32×10^{-4} for the gas phase systems, 3.70×10^{-4} for *cis*-diimide in water and 2.76×10^{-4} for the DMC/EC mixture (all field strengths in atomic units). While the perturbations vary slightly between the test systems, they are chosen small enough to be in the linear response regime, meeting the assumptions made for spectral analysis.

AIMD calculations were run for at least 30 ps with a 0.5 fs time step, with electric dipole moments being sampled every 2.5 fs. All AIMD runs made use of the same reference trajectory obtained at the same level of theory without embedding. This choice of using non-embedding MD trajectories to test our embedding-based electric dipole methods was made to exclude any error in the calculation of MD forces which is to be expected when running a DFT embedding MD simulation, so any observed error can be entirely attributed to the approach used for the calculation of electric dipole moments. The spectral post-processing of the embedding (Eq. 15a) and Wannier (Eq. 15b) methods was conducted with the internal "RTana" analysis tools using a damping factor of $0.1 E_h^{-1}$ for all systems,^{59–62} supplemented with Scipy⁶³ and the Matplotlib plotting library for visualisation.⁶⁴

Finding an optimal set of parameters for DFT embedding calculations often involves some trial and error, especially when extrapolating the initial approximate electron density as in RT-TDDFT, and small errors can add up over the course of the simulation. Still, the embedding method performed on par with the localization method in the AIMD case when the same exchange-correlation functional was used. Since we were only interested in a consistent setup for comparability with the localization approach, we chose to perform the embedding RT-TDDFT calculations with the same sets of parameters and subsystems as their localization-based counterparts. Thus, the functional for the kinetic energy used for embedding was also PBE. We refrained from applying advanced post-processing or smoothing in order to not accidentally hiding any errors of the subsystem methods, with the only exceptions being zero-filling and apodization before discrete Fourier transforms.

A. Spectral analysis

We calculate the frequency (ω)-dependent IR intensities $A(\omega)$ by Fourier transformation of the autocorrelation function with τ as the lag with respect to the time t of the first derivative of the electric dipole moment $\mathbf{d}$ over time. $\mathbf{d}$ can refer to either $\mathbf{d}^{\text{ref}}$, $\mathbf{d}^{\text{loc}}$ or $\mathbf{d}^{\text{emb}}$ depending on the method chosen.⁶⁵

$$A(\omega) \propto \int \left\langle \dot{\mathbf{d}}(\tau) \dot{\mathbf{d}}(t + \tau) \right\rangle_{\tau} e^{-i\omega t} dt \quad (16)$$

Time-derivatives of the electric dipole moments can be easily obtained by finite differences between sampling steps. To calculate absorption spectra from RT-TDDFT simulations, the 3×3 polarizability tensor $\mathbf{P}$ in the dipole approximation is constructed first by running three simulations where the initial delta pulse is applied parallel to each cartesian direction. The elements of $\mathbf{P}$ are then given by

$$P_{\alpha\beta}(\omega) = \frac{\langle d_{\alpha}(\omega) \rangle}{E_{\beta}(\omega)} \quad (17)$$

where α enumerates the components of the electric dipole moment $\mathbf{d}$ and β the components of the respective Fourier-transformed electric field $E(\omega)$. $\langle d_{\alpha}(\omega) \rangle$ is obtained from Fourier-transforming the time evolution of the electric dipole moment in direction α as

$$\langle d_{\alpha}(\omega) \rangle = \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} dt (\langle d_{\alpha}(t) \rangle - d_{0,\alpha}) e^{i\omega t} e^{-\gamma t}, \quad (18)$$

where $d_{0,\alpha}$ denotes the electric dipole moment of the unperturbed system and γ is a qualitatively determined dampening factor. The absorption spectrum is then computed from the imaginary part of the trace of $\mathbf{P}$ as⁶⁶

$$S(\omega) = \frac{4\pi\omega}{3c} \text{Tr}(\text{Im}(\mathbf{P}(\omega))). \quad (19)$$

Eckart conditions⁶⁷ were taken into account to decouple rotational and vibrational modes for cleaner IR spectra in the gas phase samples.

Algorithm 1 Parallel cyclic Jacobi method with Cardoso-Souloumiac angles for joint approximate diagonalization of eigenmatrices (JADE): This procedure overwrites the set of matrices $\{\mathbf{A}\}$ with its diagonalized form, i.e. its spread functional $\Omega(\{\mathbf{A}\})$ is minimized to tolerance ϵ by a unitary transformation $\mathbf{U}$. The given algorithm is executed on each one of the n_{proc} processors which are 0-indexed by i_{proc} and connected by a message passing protocol. n_{MO} is the number of Wannier functions.

procedure JADEPARALLEL($\{\mathbf{A}\}, \mathbf{U}, \epsilon$)

$n_{\text{pairs}} \leftarrow n_{\text{MO}}/n_{\text{proc}}/2$

$\{top\} \leftarrow 2 \cdot i_{\text{proc}} \cdot n_{\text{pairs}} + 1 : 2 \cdot (i_{\text{proc}} + 1) \cdot n_{\text{pairs}} - 1 : 2$ $\triangleright$ in NumPy slicing notation

$\{bot\} \leftarrow 2 \cdot i_{\text{proc}} \cdot n_{\text{pairs}} + 2 : 2 \cdot (i_{\text{proc}} + 1) \cdot n_{\text{pairs}} : 2$

$\{\{\mathbf{a}^{\text{top}}\}\} \leftarrow \{\mathbf{A}_{\{top\}}\}$

$\{\{\mathbf{a}^{\text{bot}}\}\} \leftarrow \{\mathbf{A}_{\{bot\}}\}$

$\{\mathbf{u}^{\text{top}}\} \leftarrow \mathbf{I}_{\{top\}}$

$\{\mathbf{u}^{\text{bot}}\} \leftarrow \mathbf{I}_{\{bot\}}$

while $\Omega(\{\mathbf{A}\}) > \epsilon$ **do**

for $i \in 1 : n_{\text{MO}} - 1$ **do**

for $j \in 1 : n_{\text{MO}}$ **do**

$p \leftarrow \text{MIN}(top_j, bot_j)$

$q \leftarrow \text{MAX}(top_j, bot_j)$

$(c, s) \leftarrow \text{CARDOSO SOULOUMIAC}(\{\mathbf{a}_p^p\}, \{\mathbf{a}_q^p\}, \{\mathbf{a}_p^q\}, \{\mathbf{a}_q^q\})$

 UPDATECOLUMNS($\{\mathbf{a}_j^p\}, \{\mathbf{a}_j^q\}, \mathbf{u}_j^p, \mathbf{u}_j^q, c, s$)

 BROADCAST(c, s, p, q)

 UPDATEROWS($\{\{\mathbf{a}^p\}\}, \{\{\mathbf{a}^q\}\}, \{\mathbf{u}^p\}, \{\mathbf{u}^q\}, \{c\}, \{s\}, \{p\}, \{q\}$)

 UPDATECOLUMNS($\{\mathbf{a}_j^p\}, \{\mathbf{a}_j^q\}, c, \bar{s}$)

 UPDATEROWS($\{\{\mathbf{a}^p\}\}, \{\{\mathbf{a}^q\}\}, \{c\}, \{s\}$)

end for

 ROTATE($\{top\}, \{bot\}, \{\{\mathbf{a}^{\text{top}}\}\}, \{\{\mathbf{a}^{\text{bot}}\}\}, \{\mathbf{u}^{\text{top}}\}, \{\mathbf{u}^{\text{bot}}\}$)

end for

end while

end procedure

B. Localization procedure during real-time propagation

CP2K offers a variety of localization procedures for time-independent systems where the MO coefficients are purely real which we used for AIMD calculations. However, localization of time-dependent orbitals, where the coefficients are complex, required a newly implemented localization algorithm.

Since the maximum localization criterion (Eq. 5) presents us with a minimization problem, it can be solved numerically with a variational approach. Moreover, since the spread functional (Eq. 6) can be interpreted as a measure of off-diagonality, this problem reduces to a joint matrix diagonalization on 6 $\mathbb{R}^{N \times N}$ matrices in the electronic ground states (AIMD) and 6 $\mathbb{C}^{N \times N}$ matrices for time-dependent orbital sets (RT-TDDFT), denoted as $\{\mathbf{A}\}$: 3 cartesian directions with a cosine and sine component each.³² An example of Wannierization during RT-TDDFT was previously implemented in Qb@ll as a proof of concept.^{28,34} As a robust first implementation we designed the code to apply a full localization procedure at each time step of an RT-TDDFT simulation. Like the approaches by Gygi et al.³² and Yost et al.,²⁸ we base the implementation on the cyclic Jacobi algorithm⁶⁸ using the Cardoso-Souloumiac formulation of optimal Jacobi rotation angles in complex Hermitian matrices.⁶⁹ The implementation was integrated into CP2K, adapted to its parallel architecture, but can be adapted to any other parallelism in a straightforward way if required.

The localization matrix is initialized to the identity matrix $\mathbf{U} \leftarrow \mathbf{I}$. Each processor then receives an equal amount of the MO indices, which are stored as pairs in an array of top (i.e. odd) and bottom (i.e. even) indices (Fig. 1). According to this distribution, each processor then loads the corresponding columns of $\mathbf{U}$ and each of the matrices in $\{\mathbf{A}\}$ into its local memory. If necessary, the indices and columns can be padded with zeroes until the total number of columns is divisible by two and the number of processors. These dummy columns can then be skipped during iteration.

Until convergence of Ω to a given tolerance ϵ , the program then performs Jacobi sweeps: For each local index pairing (p, q) (vertical neighbors in Fig. 1, ordered first to avoid redundant operations on the Hermitian matrices in $\{\mathbf{A}\}$), the elements of the local columns of $\{\mathbf{A}\}$ corresponding to these indices are then used to create the $\mathbb{R}^{N \times N}$ eigenvalue problem $\mathbf{G}$ of Cardoso-Souloumiac rotation angles⁶⁹

Figure 1: Schematic depiction of the column distribution scheme of 12 molecular orbitals between three separate processors with no shared memory. Arrows indicate the rotation of columns and global indices between the steps of a Jacobi sweep.

$$\mathbf{G} = \text{Re} \sum_k \mathbf{g}_k^H \mathbf{g}_k \quad (20)$$

where k enumerates the matrices in $\{\mathbf{A}\}$ and $\mathbf{g}$ is built as follows, with a referring to the indexed element of the k -th column of the local set and i being the imaginary unit:

$$\mathbf{g}_k = [a_{kp}^p - a_{kq}^q, a_{kq}^p + a_{kp}^q, i(a_{kp}^q - a_{kq}^p)]. \quad (21)$$

The eigenvector corresponding to the largest eigenvalue, $[x \ y \ z]^T$, is then chosen so that $x > 0$, which avoids overly large rotation angles for the sake of numerical stability. The optimal rotation matrix $\mathbf{R}$ is then defined as⁶⁹

$$\mathbf{R} = \begin{pmatrix} c & \bar{s} \\ -s & \bar{c} \end{pmatrix} \quad (22a)$$

$$c = \sqrt{\frac{x+r}{2r}} \quad (22b)$$

$$s = \frac{y - iz}{\sqrt{2r(x+r)}} \quad (22c)$$

$$r = \sqrt{x^2 + y^2 + z^2}. \quad (22d)$$

Since c is always entirely real (within numerical error), the corresponding Jacobi rotation $\mathbf{J}(c, s, p, q)$ which applies this rotation only to columns and rows p and q can be easily applied to a matrix in local memory by two calls of the ZROT LAPACK routine^{70–72} or, when

adhering to a ScaLAPACK-compatible distribution scheme, in the corresponding parallel PZROT variants⁷³. However, since the design of the cyclic Jacobi algorithm requires this specific cyclic distribution scheme, the rotations must be done manually, rotating the local columns of $\{\mathbf{A}\}$ and the local columns of $\mathbf{U}$, for each column $\mathbf{v}$:

$$\mathbf{v}^p \leftarrow c\mathbf{v}^p + s\mathbf{v}^q \quad (23a)$$

$$\mathbf{v}^q \leftarrow -\bar{s}\mathbf{v}^p + c\mathbf{v}^q. \quad (23b)$$

c , s , p and q are then broadcast to all remaining processors so that all processors can execute the row update on their local portion of four matrix elements:

$$\begin{pmatrix} v_p^p & v_q^p \\ v_p^q & v_q^q \end{pmatrix} \leftarrow \mathbf{R} \begin{pmatrix} v_p^p & v_q^p \\ v_p^q & v_q^q \end{pmatrix}. \quad (24)$$

Since this is the most communication-heavy part of the algorithm, it is likely to be the bottleneck of the procedure. To complete the left-hand side of the similarity transform $\{\mathbf{A}\} \leftarrow \mathbf{J}^\dagger \{\mathbf{A}\} \mathbf{J}$, all local columns and rows of $\{\mathbf{A}\}$ are updated a second time with s replaced by its complex conjugate. The local columns and corresponding indices are then rotated between processors as shown in Fig. 1 and the process is repeated from Eq. 20. One sweep is complete after $n_{\text{MO}} - 1$ steps, when the columns have completed one cycle of the distribution scheme. $\text{off}(\{\mathbf{A}\})$ is checked for convergence and another sweep is performed if needed. The algorithm is given in pseudocode in Alg. 1.

IV. RESULTS

In the following spectra, black represents the Berry phase dipole approach (Eq. 1) and thus the reference spectrum, blue the Wannier centre approach (Eq. 12), green the embedding approach (Eq. 15a), and red the localized approach (Eq. 15b). Renders of all investigated systems are given in Fig. 2. In addition to the spectra included in this manuscript, the single molecule spectra we discuss are given in the supplementary material.

Table I: Root mean square error of the spectra obtained from the tested subsystem methods against the Berry phase approach, given in arbitrary units of the normalized spectra. "wc" refers to the Wannier centers method (Eq. 12), "emb" to the DFT embedding approach (Eq. 15a), and "loc" to the localization approach (Eq. 15b).

system	absorption			IR		
	wc	emb	loc	wc	emb	loc
H ₂ O	4.30×10^{-3}	—	8.33×10^{-3}	2.70×10^{-3}	—	3.27×10^{-3}
N ₂ H ₂	3.36×10^{-2}	—	2.46×10^{-2}	9.42×10^{-2}	—	1.22×10^{-2}
2 H ₂ O	2.38×10^{-2}	1.31×10^{-1}	2.38×10^{-2}	2.78×10^{-2}	1.03×10^{-1}	1.10×10^{-1}
N ₂ H ₂ + H ₂ O	1.49×10^{-2}	3.72×10^{-2}	1.50×10^{-2}	1.90×10^{-1}	6.96×10^{-2}	9.20×10^{-2}
N ₂ H ₂ (aq)	4.54×10^{-2}	4.03×10^{-2}	7.98×10^{-2}	1.11×10^{-2}	2.77×10^{-1}	1.66×10^{-2}
DMC/EC	9.21×10^{-2}	9.06×10^{-2}	9.71×10^{-2}	3.04×10^{-2}	3.18×10^{-2}	3.42×10^{-2}

Figure 2: Molecular structures of the water dimer (a), the *cis*-diimide-water dimer (b), *cis*-diimide in water (c), and DMC/EC mixture (d).

A. RT-TDDFT

The test case of single molecules *in vacuo* without any defined subsystems resulted in spectra which are indistinguishable from the reference (see Figs. 1-4 in the SI). This comes as no surprise since the approximations made in Eqs. 15a and 15b are negligible when the simulation box is much bigger than the system under investigation and no error from embedding or approximate

Figure 3: Absorption spectra simulated via RT-TDDFT with the total spectrum given in solid lines and a selected subspectrum from the respective method given in black dashed lines. For each simulation, every single molecule was treated as an own subsystem. (a) the water dimer (subsystem: in-plane water), (b) the *cis*-diimide-water dimer (subsystem: diimide), (c) one *cis*-diimide molecule solvated in water (subsystem: diimide), and (d) a 1:1 liquid mixture of ethylene carbonate and dimethyl carbonate (subsystem: EC). "ref" refers to the Berry phase dipole formulation used as reference, "wc" refers to the Wannier centers method (Eq. 12), "emb" to the DFT embedding approach (Eq. 15a), and "loc" to the localization approach (Eq. 15b).

diagonalization is introduced.

Moving on to the first test involving subsystems, investigation of the dimer spectrum provides a cleanly separated water subspectrum with the expected features (Fig. 3(a)), just as the subsystem

spectrum of the *cis*-diimide-water dimer (Fig. 3(b)). The water dimer subspectrum also reveals that peaks above the ionization energy of water (experiment: 12.62 eV⁷⁴, PBE tends to underestimate this value by ≈ 0.5 eV in the GW approach⁷⁵ even with large basis sets) are assigned to one of the subsystems only. Intuitively, one would expect these excitations to be located on both systems, either by Rydberg or ionized states. This "overlocalization" is caused by the design of the algorithm implementing Eq. 15b, which assigns the ownership of a Wannier function to the subsystem to which its center is closest. To investigate the ramifications of this phenomenon, we run a set of RT-TDDFT simulations of the water dimer at various displacements along the H-bond axis, starting from equilibrium distance r_0 (defined as 2.18 Å as shown in Fig. 4) up to $2.4r_0$. The effects are depicted in Fig. 4: at equilibrium distance, the algorithm places full ownership of e.g. the ≈ 19 eV excitation on the in-plane water (with respect to the mirror plane). As the distance increases, the peaks are distributed more evenly between the monomers. The same downside is also found by the embedding approach, where the subdensities are divided just as strictly. Investigation of shared excitation thus requires special care. For example, a slight change of the algorithm would make it possible to treat Wannier functions whose ownership changes frequently separately to form an additional subsystem (without any nuclei) to capture shared excitations.

As we move towards the strongly interacting water solution of *cis*-diimide, we observe that while the subsystem of diimide is separated well from the bulk spectrum, the noise floor is raised significantly in all methods. Still, the characteristic peaks of diimide at around 10 eV and at 17 eV are well recognizable (Fig. 3(c)). In the case of the DMC/EC mixture, the embedding run performed almost on par with the conventional method (Fig 3(d)). Both runs reproduce the reference spectrum reasonably close, with a slight increase of noise in the embedding case. The predicted DMC subspectra are also in agreement between the electric dipole analysis methods, clearly attributing large parts of the regions around 8 eV and 16 eV to the ethylene carbonates.

Although we were able to obtain high quality spectra for aqueous solutions, the DFT embedding method employed here is expected to perform worse if one intends to use it in RT-TDDFT absorption simulations for very strongly interacting subsystems because, among others, relatively small errors in the initial electronic structure can add up over time since RT-TDDFT relies on extrapolation. Additional error might also be mitigated by choosing a smaller number of subsystems, e.g. by grouping nearby water molecules together, or fine-tuning of the kinetic energy functional used for embedding.

Figure 4: Absorption spectra of the water dimer near and above the ionization energy of water at various displacements along the H-bond axis, starting at the equilibrium distance r_0 . The dashed lines represent the in-plane H_2O molecule, obtained from the embedding ("emb", Eq. 15a, in green) and the localization ("loc", Eq. 15b, in red) approaches.

B. AIMD

The single molecule spectra reproduce the reference spectra perfectly, as expected (see Figs. 5-8 in the SI). Once again, the dimer systems produce cleanly separated subspectra, with only slight differences between the methods, most notably the exaggerated intensities of the 1500cm^{-1} and 3100cm^{-1} vibrations, which are entirely attributable to the diimide. It is also apparent that DFT embedding performs slightly worse for the water dimer where the monomers are interacting more strongly, as the total spectrum appears with significantly altered intensities from the reference spectra.

In the bulk phase, this trend becomes more pronounced. While all tested methods perform perfectly in the weakly interacting DMC/EC mixture (with the EC subspectrum predicted prior¹² being recovered as well), the aqueous solution of *cis*-diimide shows significant alterations of the intensities, with lower energy vibrations being exaggerated while higher energy vibrations appear less intense than expected from the reference and localization-based methods. This is most prominent in the $3000\text{-}3500\text{ cm}^{-1}$ range, where the spectrum is dominated by the O-H stretching vibration, broadened by the spectral band of H-bonds. Nevertheless, the single diimide produces a subspectrum entirely on par with the localization approaches.

It is clear that AIMD performs well with DFT embedding analysis of local electric dipole moments

Figure 5: IR spectra simulated via AIMD with the total spectrum given in solid lines and a selected subspectrum from the respective method given in black dashed lines. For each simulation, every single molecule was treated as an own subsystem. (a) the water dimer (subspectrum: in-plane water), (b) the *cis*-diimide-water dimer (subspectrum: diimide), (c) one *cis*-diimide molecule solvated in water (subspectrum: diimide), and (d) a 1:1 liquid mixture of ethylene carbonate and dimethyl carbonate (subspectrum: EC). "ref" refers to the Berry phase dipole formulation used as reference, "wc" refers to the Wannier centers method (Eq. 12), "emb" to the DFT embedding approach (Eq. 15a), and "loc" to the localization approach (Eq. 15b).

for the systems studied here and their AIMD trajectories, albeit with more pronounced differences to reference data than RT-TDDFT when hydrogen-bonding is relevant. Still, it remains a competitive alternative to localization methods in this use case, especially considering that it can lead to significantly faster wavefunction optimization and does not require the additional

localization procedure. While this validates the tested approach to electronic subsystem dipoles, concern remains about the accuracy of a fully embedded simulation, i.e. one where the trajectory is obtained from the equations of motion on a DFT embedding density. Previous results by Genova et al.⁷⁶ showed promising results for accurate trajectories using frozen density embedding, where the expected radial and angular distribution functions as well the diffusion coefficient are recovered for small water boxes and the solvated hydroxy radical. Within the embedding framework used in CP2K, the reduced quality of the $\text{N}_2\text{H}_2(\text{aq})$ spectrum can also indicate that strongly interacting subsystems are better investigated with localization-based methods. It should also be noted that fine-tuning of the embedding parameters like different kinetic energy functionals or a different choice of water subsystems may alleviate some of the error. For now, the addition of Eq. 15b (aside from Eq. 12) provides us with a powerful tool to investigate such systems with a periodic, localization-based subsystem approach. We have shown that the additional approximation that Eq. 15b makes over Eq. 12 is small enough in the tested systems. The resulting formulation uses the traditional position operator while retaining the periodic property expected of bulk polarization.

V. OUTLOOK

We have investigated approaches for the calculation of local electric dipole moments within PBC using MLWF and subsystem DFT respectively. For this purpose, we successfully adapted the local electric dipole approach to a Wannier function approach (Eq. 15b) and implemented the real-time propagation of MLWFs into CP2K, enabling their use in a widely used quantum chemistry code. We have shown that the approximation to the Berry phase formulation of electric dipole moments which was previously applied to simulations that use DFT embedding to separate subsystems also performs well when used on MLWFs and have implemented these novel analysis approaches into the popular quantum chemistry program CP2K. While it comes with an increase in computational cost due to the additional localization procedure at each sampled step as well as the loss of computational savings due to DFT embedding, it makes it possible to investigate local electric dipole moments in subsystems where DFT embedding may give lackluster results, particularly when subsystems are connected via covalent bonds or when strong interatomic interactions are relevant. While we obtained sufficiently accurate spectra for aqueous test systems, the option to

use localization instead of embedding may prove advantageous in cases where DFT embedding performs poorly. It might also provide a promising route for obtaining local higher electric multipole and magnetic moments. The success of the method in both AIMD and RT-TDDFT simulations also suggests that its application to EMD can provide insights into coupled electronic and nuclear dynamics of subsystems given the necessary computational resources.

VI. SUPPLEMENTARY MATERIAL

The supplementary material provides the single molecule absorption and IR spectra.

VII. ACKNOWLEDGEMENTS

The work was supported by the University of Zurich and the Swiss National Science Foundation (Grant No. PP00P2 170667). We thank the Swiss National Supercomputing Center for computing resources (Project Nos. s788, s875, s1001, and s1036).

VIII. DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request

REFERENCES

- ¹R. Car and M. Parrinello, Physical review letters **55**, 2471 (1985).
- ²D. Marx and J. Hutter, *Ab initio molecular dynamics: basic theory and advanced methods* (Cambridge University Press, 2009).
- ³R. King-Smith and D. Vanderbilt, Physical Review B **47**, 1651 (1993).
- ⁴R. Resta, EPL (Europhysics Letters) **22**, 133 (1993).
- ⁵D. Vanderbilt and R. King-Smith, Physical Review B **48**, 4442 (1993).
- ⁶I. Souza, J. Íñiguez, and D. Vanderbilt, Physical review letters **89**, 117602 (2002).
- ⁷I. Souza, J. Íñiguez, and D. Vanderbilt, Physical Review B **69**, 085106 (2004).
- ⁸S. Ono, L. Trifunovic, and H. Watanabe, Physical Review B **100**, 245133 (2019).

- ⁹N. Marzari and D. Vanderbilt, Physical review B **56**, 12847 (1997).
- ¹⁰I. Souza, N. Marzari, and D. Vanderbilt, Physical Review B **65**, 035109 (2001).
- ¹¹N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Reviews of Modern Physics **84**, 1419 (2012).
- ¹²S. Lubert, The Journal of chemical physics **141**, 234110 (2014).
- ¹³R. G. Gordon and Y. S. Kim, The Journal of Chemical Physics **56**, 3122 (1972).
- ¹⁴R. Gordon and Y. Kim, J. Chem. Phys **60**, 1842 (1974).
- ¹⁵G. Senatore and K. Subbaswamy, Physical Review B **34**, 5754 (1986).
- ¹⁶P. Cortona, Physical Review B **44**, 8454 (1991).
- ¹⁷T. A. Wesolowski and A. Warshel, The Journal of Physical Chemistry **97**, 8050 (1993).
- ¹⁸M. E. Casida and T. A. Wesolowski, International journal of quantum chemistry **96**, 577 (2004).
- ¹⁹P. Hohenberg and W. Kohn, Physical review **136**, B864 (1964).
- ²⁰E. Runge and E. K. Gross, Physical Review Letters **52**, 997 (1984).
- ²¹C. A. Ullrich, *Time-dependent density-functional theory: concepts and applications* (OUP Oxford, 2011).
- ²²M. Marques, A. Rubio, E. K. Gross, K. Burke, F. Nogueira, and C. A. Ullrich, *Time-dependent density functional theory*, Vol. 706 (Springer Science & Business Media, 2006).
- ²³K. Lopata and N. Govind, Journal of chemical theory and computation **7**, 1344 (2011).
- ²⁴W. Liang, C. T. Chapman, and X. Li, The Journal of chemical physics **134**, 184102 (2011).
- ²⁵M. Casida, Recent Developments and Applications in Modern Density Functional Theory **4** (1996).
- ²⁶M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, The Journal of chemical physics **108**, 4439 (1998).
- ²⁷R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, The Journal of chemical physics **109**, 8218 (1998).
- ²⁸D. C. Yost, Y. Yao, and Y. Kanai, The Journal of chemical physics **150**, 194113 (2019).
- ²⁹R. Car and M. Parrinello, Phys. Rev. Lett **55**, 2471 (1985).
- ³⁰R. Futrelle and D. McGinty, Chemical Physics Letters **12**, 285 (1971).
- ³¹“Cp2k, <http://www.cp2k.org/> (accessed 2020-12-04),” .
- ³²F. Gygi, J.-L. Fattebert, and E. Schwegler, Computer physics communications **155**, 1 (2003).

- ³³A. Schleife, E. W. Draeger, Y. Kanai, and A. A. Correa, The Journal of chemical physics **137**, 22A546 (2012).
- ³⁴F. Gygi, IBM Journal of Research and Development **52**, 137 (2008).
- ³⁵S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, Reviews of Modern Physics **73**, 515 (2001).
- ³⁶E. Yaschenko, L. Fu, L. Resca, and R. Resta, Physical Review B **58**, 1222 (1998).
- ³⁷G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, and M. Parrinello, Physical Review B **61**, 10040 (2000).
- ³⁸W. A. Benalcazar, B. A. Bernevig, and T. L. Hughes, Science **357**, 61 (2017).
- ³⁹W. A. Benalcazar, B. A. Bernevig, and T. L. Hughes, Physical Review B **96**, 245115 (2017).
- ⁴⁰F. Bloch, Zeitschrift für physik **52**, 555 (1929).
- ⁴¹J. Foster and S. Boys, Reviews of Modern Physics **32**, 300 (1960).
- ⁴²R. Resta, Physical Review Letters **80**, 1800 (1998).
- ⁴³T. A. Wesolowski, S. Shedge, and X. Zhou, Chemical reviews **115**, 5891 (2015).
- ⁴⁴C. Huang, M. Pavone, and E. A. Carter, The Journal of chemical physics **134**, 154110 (2011).
- ⁴⁵Q. Wu and W. Yang, The Journal of chemical physics **118**, 2498 (2003).
- ⁴⁶J. Kolafa, Journal of computational chemistry **25**, 335 (2004).
- ⁴⁷M. Iannuzzi, B. Kirchner, and J. Hutter, Chemical physics letters **421**, 16 (2006).
- ⁴⁸M. E. Fornace, J. Lee, K. Miyamoto, F. R. Manby, and T. F. Miller III, Journal of chemical theory and computation **11**, 568 (2015).
- ⁴⁹A. Krishtal, D. Ceresoli, and M. Pavanello, The Journal of chemical physics **142**, 154116 (2015).
- ⁵⁰K. J. Koh, T. S. Nguyen-Beck, and J. Parkhill, Journal of chemical theory and computation **13**, 4173 (2017).
- ⁵¹F. Ding, F. R. Manby, and T. F. Miller III, Journal of Chemical Theory and Computation **13**, 1605 (2017).
- ⁵²M. De Santis, L. Belpassi, C. R. Jacob, A. Severo Pereira Gomes, F. Tarantelli, L. Visscher, and L. Storchi, Journal of chemical theory and computation **16**, 5695 (2020).
- ⁵³S. Andermatt, J. Cha, F. Schiffmann, and J. VandeVondele, *Journal of Chemical Theory and Computation* **12**, 3214 (2016).
- ⁵⁴A. Krishtal, D. Sinha, A. Genova, and M. Pavanello, *Journal of Physics: Condensed Matter* **27**, 183202 (2015).

- ⁵⁵S. Goedecker, M. Teter, and J. Hutter, *Physical Review B* **54**, 1703 (1996).
- ⁵⁶C. Hartwigsen, S. Goedecker, and J. Hutter, *Physical Review B* **58**, 3641 (1998).
- ⁵⁷M. Krack, *Theoretical Chemistry Accounts* **114**, 145 (2005).
- ⁵⁸A. Castro, M. A. Marques, and A. Rubio, *The Journal of chemical physics* **121**, 3425 (2004).
- ⁵⁹J. Mattiat and S. Lubner, *The Journal of Chemical Physics* **149**, 174108 (2018).
- ⁶⁰J. Mattiat and S. Lubner, *Chemical Physics* **527**, 110464 (2019).
- ⁶¹J. Mattiat and S. Lubner, *The Journal of Chemical Physics* **151**, 234110 (2019).
- ⁶²J. Mattiat and S. Lubner, *Journal of Chemical Theory and Computation* **17**, 344 (2020).
- ⁶³P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *Nature Methods* **17**, 261 (2020).
- ⁶⁴J. D. Hunter, *Computing in Science & Engineering* **9**, 90 (2007).
- ⁶⁵M. Thomas, M. Brehm, R. Fligg, P. Vöhringer, and B. Kirchner, *Physical Chemistry Chemical Physics* **15**, 6608 (2013).
- ⁶⁶K. Yabana and G. F. Bertsch, *International Journal of Quantum Chemistry* **75**, 55 (1999).
- ⁶⁷C. Eckart, *Physical Review* **47**, 552 (1935).
- ⁶⁸C. F. Van Loan and G. H. Golub, *Matrix computations* (Johns Hopkins University Press Baltimore, 1983).
- ⁶⁹J.-F. Cardoso and A. Souloumiac, *SIAM journal on matrix analysis and applications* **17**, 161 (1996).
- ⁷⁰E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).
- ⁷¹Z. Xianyi and M. Kroeker, “OpenBLAS: An optimized BLAS library,” .
- ⁷²I. Corporation, “Intel® math kernel library,” .
- ⁷³L. Blackford, J. Choi, A. Cleary, E. D’Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and W. RC, *ScaLAPACK Users' Guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997).

- ⁷⁴P. Linstrom, “[Nist chemistry webbook, nist standard reference database 69](#),” (1997).
- ⁷⁵M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, [Journal of Chemical Theory and Computation](#) **11**, 5665 (2015).
- ⁷⁶A. Genova, D. Ceresoli, and M. Pavanello, [The Journal of Chemical Physics](#) **144**, 234105 (2016).