Calculating excited states of extended systems in LR-TDDFT

T. J. Zuehlsdorff[1, *] and N. D. M. Hine[2]

[1] *Cavendish Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, UK*

[2] *Department of Physics, University of Warwick, Coventry, CV4 7AL, UK*

(Dated: September 21, 2015)

In this eCSE funded project, a number of improvements have been made to the linear-scaling TDDFT implementation in the `ONETEP` code. These have signficantly expanded its scope as a powerful tool to study spectroscopic properties of large-scale biological systems on high-performance computing facilities like ARCHER. We have carried out extensive benchmarks of the existing Tamm-Dancoff TDDFT implementation, demonstrating excellent parallel scalability using hybrid OpenMP/MPI parallelism. We have furthermore implemented an efficient preconditioning scheme that reduces the walltime of TDDFT calculations performed on large numbers of cores by more than factor of two. In order to improve on the spectra of a number of pigments that are important for biological applications, we have extended the TDDFT formalism beyond the Tamm-Dancoff approximation, allowing for efficient linear-scaling TDDFT calculations of systems containing 10,000 atoms. All code improvements carried out have been made available to the end user in the newest release of `ONETEP`. Finally, we have taken an important first step of implementing excited state gradients in `ONETEP`, which will, with minor future development, allow for the calculation of photoisomerisation processes in large pigment-protein complexes.

I. INTRODUCTION

`ONETEP`[1] and other codes for large-scale DFT calculations have enabled the prediction of ground state properties of a wide range of complex systems, such as biomolecules and nanostructured materials. Ground state calculations of systems containing thousands of atoms can now be routinely performed by non-expert users. However, the prediction of excited states in large biomolecules is still a considerable challenge. While initial developments in the `ONETEP` code have shown the potential for performing calculations predicting spectroscopic properties of large-scale systems[2], further development of methodology and simulation software is needed in order to treat challenging problems like light-harvesting pigment-protein complexes and photoactivated biological molecules like rhodopsin.

The work carried out in this eCSE project is focused on two main tasks. The first task is to test and optimise a linear-scaling implementation of time-dependent density functional theory (TDDFT) in the Tamm-Dancoff approximation already available in the `ONETEP` code[2] in order to ensure its suitability for obtaining spectroscopic properties of systems like large pigment-protein complexes. A special focus was placed on ensuring that the excellent parallel scalability, previously demonstrated for ground state DFT in the `ONETEP` code on high-performance computing (HPC) platforms, is carried over to the TDDFT implementation. The second part of the project is focused on new methodological developments centered around extending the TDDFT implementation to solving the full TDDFT eigenvalue equation and beginning an implementation of excited state gradients, which are crucial for obtaining transition states of photoactivated biomolecules.

---

[*] Electronic address: `tjz21@cam.ac.uk`

All code development and testing is aimed at significantly expanding the scope of the `ONETEP` TDDFT implementation for the ARCHER community, ensuring that it will be a powerful tool for theoretical spectroscopy in large-scale biological systems.

II. THEORETICAL BACKGROUND

In order to detail the work carried out in this eCSE project, it is first necessary to give a short overview over the capabilities of the `ONETEP` code at the start of the project, as well as some background on the theoretical methods used. For this reason, we will give a short introduction to linear-response TDDFT, the most important features of the `ONETEP` code, and the recent implementation of linear-response TDDFT in the Tamm-Dancoff approximation in the code.

A. Linear-response TDDFT

In linear-response TDDFT, the task of finding the lowest excitations of a given system is recast into solving for the eigenvalues of an effective 2-particle Hamiltonian. In the `ONETEP` code, the focus is on treating large isolated systems, in which case the Kohn-Sham DFT orbitals can be chosen to be real. For such an isolated system, the non-hermitian TDDFT eigenvalue equation can be expressed as[3]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & -\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{1}$$

with $\mathbf{A}$ and $\mathbf{B}$ denoting block matrices. If the sets of occupied and unoccupied Kohn-Sham states of spin channel σ are given by $\{\psi_{v\sigma}\}$ and $\{\psi_{c\sigma}\}$ with associated eigenvalues $\{\epsilon^{\mathrm{KS}}_{v\sigma}\}$ and $\{\epsilon^{\mathrm{KS}}_{c\sigma}\}$, the matrix elements of $\mathbf{A}$ and $\mathbf{B}$ can be written as

$$\begin{aligned} A_{cv\sigma,c'v'\sigma'} &= \delta_{\sigma\sigma'}\delta_{cc'}\delta_{vv'}(\epsilon^{\mathrm{KS}}_{c'\sigma'} - \epsilon^{\mathrm{KS}}_{v'\sigma'}) + K_{cv\sigma,c'v'\sigma'} && (2) \\ B_{cv\sigma,c'v'\sigma'} &= K_{cv\sigma,c'v'\sigma'}. && (3) \end{aligned}$$

As can be seen, the TDDFT eigenvalue equation consists of a diagonal part of Kohn-Sham eigenvalue differences and off-diagonal terms due to the matrix $\mathbf{K}$ that couple individual Kohn-Sham transitions. The matrix elements of $\mathbf{K}$ are given by

$$K_{cv\sigma,c'v'\sigma'} = \psi_{c\sigma}(\mathbf{r})\psi_{v\sigma}(\mathbf{r}) \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \left.\frac{\delta E^2_{\mathrm{xc}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0} \right] \psi_{c\sigma'}(\mathbf{r}')\psi_{v'\sigma'}(\mathbf{r}') \tag{4}$$

Here, the second derivative of the exchange-correlation potential is evaluated at the ground-state density ρ_0 of the system and the adiabatic approximation is made, ignoring any frequency dependence of $\mathbf{K}$.

The full TDDFT eigenvalue equation is difficult to solve using standard eigenvalue solvers due to its non-hermitian properties. For this reason, a commonly-made approximation, known as the Tamm-Dancoff approximation[4], is to assume that the off-diagonal coupling matrix elements $\mathbf{B}$ are small. In this approximation, solving for the TDDFT eigenvalues is then reduced to solving the much simpler, hermitian problem of $\mathbf{AX} = \omega\mathbf{X}$. This approximation is known to introduce non-negligible errors in some situations.

B. DFT in the `ONETEP` code

`ONETEP` is a linear-scaling DFT code aimed at treating systems of thousands of atoms. In order to achieve this goal, any representation in form of individual Kohn-Sham states, as would be used in standard cubic scaling plane-wave codes like CASTEP or VASP, must be avoided. Instead, a collective representation of all such states is used, in form of the single particle density matrix $\rho(\mathbf{r},\mathbf{r}')$. The density matrix is expanded in a set of non-orthogonal atom-centered localised orbitals $\{\phi_\alpha\}$ referred to as non-orthogonal generalised Wannier functions that are optimised *in-situ* during a ground state calculation in order to ideally represent the occupied Kohn-Sham space. The density matrix is then expressed as[1]

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{v}^{N_{\mathrm{occ}}} \psi_v^{\mathrm{KS}}(\mathbf{r})\psi_v^{\mathrm{KS}}(\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r})P^{\{v\}\alpha\beta}\phi_\beta(\mathbf{r}') \tag{5}$$

where all spin indices have been dropped for convenience. Since $\{\phi_\alpha\}$ is optimised during a ground state DFT calculation to optimally represent the single-particle density matrix, the occupied Kohn-Sham space is spanned to the same degree of accuracy as in plane-wave wave codes for a suitably chosen localisation radius. Since the occupied Kohn-Sham space is all that is needed to obtain the ground state energy of the system, plane-wave accuracy can be reached while a linear-scaling computational cost with system size is obtained when exploiting the fact that $\mathbf{P}^{\{v\}}$ is exponentially localised and thus sparse in any system with a band gap. The *in-situ* optimisation has the additional advantage that the size of $\{\phi_\alpha\}$ can be kept very small, as compared to other localised orbital representations such as Gaussian-type orbitals.

In TDDFT however, the eigenvector of the effective 2-particle Hamiltonian is expanded in Kohn-Sham transitions from the occupied to the unoccupied space. While $\{\phi_\alpha\}$ forms an ideal representation for the occupied space, the unoccupied space is only very poorly represented. This problem can be overcome by introducing a second set of localised atom-centered orbitals $\{\chi_\beta\}$ that is optimised in a post-processing step to ideally represent a low-energy subset of the conduction space expressed through an effective density matrix $\mathbf{P}^{\{c\}}$[5].

ONETEP has been written from its inception as a parallel code, and a recent implementation of hybrid OpenMP/MPI parallelism for ground state calculations extended the parallel scalability of the code significantly[6]. An excellent parallel efficiency over several thousand cores in ground state calculations on ARCHER has been demonstrated[10].

C. Linear-scaling TDDFT in `ONETEP`

The *in-situ* optimisation of the localised functions $\{\chi_\alpha\}$ and $\{\phi_\beta\}$ allows for a very compact and efficient representation of the valence states and the low-energy conduction states. This can be considered an ideal representation for the low energy excitations into bound states. A fully linear-scaling TDDFT implementation, within the simplified Tamm-Dancoff approximation has been implemented in the `ONETEP` code recently that fully exploits the advantages of the compact double representation of $\{\chi_\alpha\}$ and $\{\phi_\beta\}$[2].

The implementation relies on representing a trial vector $\mathbf{x}$ to the TDDFT eigenvalue equation through an effective density matrix $\mathbf{P}^{\{1\}}$ such that

$$\rho^{\{1\}}(\mathbf{r}) = \sum_{cv} \psi_c^{\mathrm{KS}} x_{cv} \psi_v^{\mathrm{KS}} = \sum_{\alpha\beta} \chi_\alpha(\mathbf{r})P^{\{1\}\alpha\beta}\phi_\beta(\mathbf{r}) \tag{6}$$

where $\rho^{\{1\}}(\mathbf{r})$ is the effective response density associated with trial vector $\mathbf{x}$. Denoting the action of $\mathbf{A}$ acting on trial vector $\mathbf{x}$ as $\mathbf{f}$, it can be seen that $\mathbf{f}$ can be computed for a given trial vector without the explicit computation of $\mathbf{A}$, which cannot be obtained in linear-scaling effort. It can be shown that in mixed $\{\chi_\alpha\}$-$\{\phi_\beta\}$ representation, $\mathbf{f}$ can be written as

$$\mathbf{f}^{\chi\phi} = \mathbf{P}^{\{c\}}\mathbf{H}^{\chi}\mathbf{P}^{\{1\}} - \mathbf{P}^{\{1\}}\mathbf{H}^{\phi}\mathbf{P}^{\{v\}} + \mathbf{P}^{\{c\}}\left(\mathbf{V}_{\mathrm{SCF}}^{\{1\}\chi\phi}\left[\mathbf{P}^{\{1\}}\right]\right)\mathbf{P}^{\{v\}} \tag{7}$$

where $\mathbf{H}^{\chi}$ and $\mathbf{H}^{\phi}$ denote the ground state DFT Hamiltonian in $\{\chi_\alpha\}$ and $\{\phi_\beta\}$ representation respectively, while $\mathbf{V}_{\mathrm{SCF}}^{\{1\}\chi\phi}$ is the result of the trial vector $\mathbf{x}$ acting on $\mathbf{K}$ and can be expressed as a functional of $\rho^{\{1\}}(\mathbf{r})$. If all involved density matrices $\mathbf{P}^{\{1\}}$, $\mathbf{P}^{\{c\}}$ and $\mathbf{P}^{\{v\}}$ can be made sparse for sufficiently large system sizes, $\mathbf{f}^{\chi\phi}$ can be obtained in linear-scaling effort.

Sparsity of the density matrix $\mathbf{P}^{\{v\}}$ for systems with a band gap is guaranteed and can be proven mathematically. By contrast, no such proof exists for $\mathbf{P}^{\{c\}}$ and $\mathbf{P}^{\{1\}}$, but it has been observed empirically that truncation of these matrices is also possible without introducing significant error. Linear-scaling TDDFT calculations are indeed possible on a large class of systems sizes of interest.

The action $\mathbf{f}^{\chi\phi}$ can be used to define the lowest excitation of the system in terms of an effective variational principle such that

$$\omega_{\mathrm{min}}^{\mathrm{TDA}} = \min_{\mathbf{P}^{\{1\}}} \Omega_{\mathrm{TDA}}\left[\mathbf{P}^{\{1\}}\right] = \frac{\mathrm{Tr}\left[\mathbf{P}^{\{1\}\dagger}\mathbf{S}^{\chi}\mathbf{f}^{\chi\phi}\mathbf{S}^{\phi}\right]}{\mathrm{Tr}\left[\mathbf{P}^{\{1\}\dagger}\mathbf{S}^{\chi}\mathbf{P}^{\{1\}}\mathbf{S}^{\phi}\right]} \tag{8}$$

where $\mathbf{S}^{\chi}$ and $\mathbf{S}^{\phi}$ are the overlap matrices of the two NGWF representations. In the `ONETEP` code, the minimisation of the functional is achieved using a conjugate gradients algorithm. Multiple excited states can be converged simultaneously by ensuring orthogonality between them. The algorithm scales fully linearly with system size for a fixed number of excitations. If the number of excitations grows, there is a small component of the computational cost which scales asymptotically with the number of excitations converged due to the orthogonality constraints. Therefore the asymptotic scaling would eventually be as $O(N^2)$. However, the main application for the `ONETEP` TDDFT implementation lies in large scale systems with a small number of well-defined low-energy excitations of interest, like pigment-protein complexes. In these systems, the number of excitations that need to be converged remains constant with system size, such that truly linear-scaling calculations far beyond the capability of standard TDDFT approaches can be achieved.

III. GENERAL CODE IMPROVEMENTS

While the Tamm-Dancoff TDDFT implementation mentioned in the previous section is fully functional in the ONETEP code, a number of crucial improvements have to be carried out in order to increase usability and enable large-scale high-performance computing applications on ARCHER. These improvements are not related to the development of new functionality but are rather focused on parallel scalability and convergence performance, as well as the reusability of the original source code for future applications.

A. Parallel scalability of TDDFT

In order for the linear-scaling TDDFT implementation in `ONETEP` to become a widely used tool in the ARCHER community, it is crucial to ensure that the excellent parallel scalability of ground-state

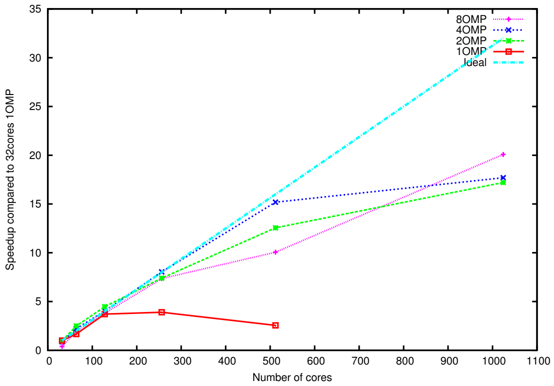

FIG. 1: Parallel efficiency of a single TDDFT conjugate gradient time step for a system of two bacteriochlorophylls in protein environment (≈ 1900 atoms)

DFT carries over to TDDFT. The TDDFT implementation uses many of the same routines as the ground state functionality, so it was comparatively simple to carry over the OpenMP parallelisation to the TDDFT code. To demonstrate this, we have performed extensive benchmarking tests on a realistic test system. The system chosen for benchmark tests is a fragment of the Fenna-Matthews-Olson pigment-protein complex containing 1906 atoms surrounded by an implicit solvent model. The calculations are performed with production-quality convergence settings and the time taken for a single conjugate gradient iteration in the Tamm-Dancoff TDDFT implementation is computed for a range of different processors and varying degrees of OpenMP/MPI parallelism.

The results of the extensive benchmarking test can be found in Fig. 1. As can be seen, the hybrid OpenMP/MPI formalism shows an excellent scaling, even in this non-trivial example of a large pigment-protein complex in implicit solvent. Note that with pure MPI parallelism, it is impossible to scale beyond 128 cores in this system without suffering a significant decrease in performance, while almost perfect parallel efficiency can be reached in the hybrid OpenMP/MPI parallelism scheme for 512 cores. If a small decrease in parallel efficiency is acceptable, the problem scales well up to 1024 cores, thus ensuring that large-scale TDDFT computations are possible that make excellent use of the resources available on ARCHER.

B. Preconditioning of the TDDFT equation

A second focus for improvement of the original `ONETEP` TDDFT implementation is a speedup in the convergence of the conjugate gradient algorithm. In the original version of the code, the reduced Tamm-Dancoff eigenvalue problem is solved using a standard conjugate gradient algorithm without any preconditioner. However, it is known from other algorithms solving the eigenvalue problem in conventional cubic scaling effort that the effective 2-particle matrix $\mathbf{A}$ is poorly conditioned and that effective preconditioners are vital in order to achieve fast convergence. In the `ONETEP` TDDFT implementation the challenge is to find a preconditioner that can be applied in linear-scaling effort.

In order to precondition the eigenvalue problem, the matrix $\mathbf{A}$ is taken to be diagonally dominant in Kohn-Sham space, such that the matrix of Kohn-Sham eigenvalue differences should form an effective

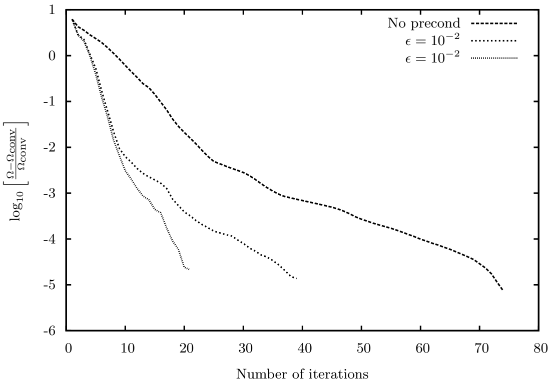

FIG. 2: Convergence of the 2 lowest states of Azobenzene for different degrees of preconditioning applied. "ϵ" describes the tolerance to which the linear system is solved in order to apply the preconditioning.

preconditioner. It is then straightforward to notice that applying such a preconditioner in mixed $\{\chi_\alpha\}$-$\{\phi_\beta\}$ representation can be achieved by solving a linear system of the form

$$\mathbf{P}^{\{c\}}\mathbf{H}^{\chi}\mathbf{G}^{\chi\phi} - \mathbf{G}^{\chi\phi}\mathbf{H}^{\phi}\mathbf{P}^{\{v\}} = \mathbf{g}^{\chi\phi} \tag{9}$$

where $\mathbf{g}^{\chi\phi}$ is the unpreconditioned gradient of the system and $\mathbf{G}^{\chi\phi}$ is the preconditioned version. Note that the linear equation can be solved in linear-scaling effort using standard numerical methods and that the scaling properties of the original algorithm are thus unaltered by the preconditioning.

The performance of the preconditioner for a small test system of Azobenzene can be found in Fig. 2. As can be seen, solving the linear system to different degrees of accuracy during every single conjugate gradient steps yields varying speedups in the convergence. For the most accurate preconditioning, the convergence rate is increased by a factor of 3.5 compared to the unconditioned system, with the computational overhead associated with the preconditioner being less than 1% of the total calculation time. While the computational overhead of the preconditioner is larger for large systems, tests have shown that for a realistic range of system sizes to be studied on ARCHER by a typical user (≈300-3000 atoms), reductions in wall time by at least a factor of 2 can be expected.

C. A general-purpose linear-response module

The linear-response TDDFT operator of Eqn. 7 is a very general operator, as it describes the first order response of the Kohn-Sham system to a perturbation to the ground state density described by $\mathbf{P}^{\{1\}}$. Therefore, the operator has wide applicability that goes beyond linear-response TDDFT. For this reason, the original implementation of Tamm-Dancoff TDDFT has undergone an extensive rewriting in order to create a general-purpose linear-response module that is easily usable in future code developments. Thus, a first step towards a general density-functional perturbation theory module has been taken, with future planned implementations focusing on the linear-response calculation of phonons. It is expected that the creation of a linear-response module carried out in this eCSE project will speed up future code developments in this direction significantly.

IV. TDDFT FORCES

The calculation of analytic excited state gradients in TDDFT is of vital importance in many applications. It allows for the performance of excited state geometry optimisations, yielding photoemission spectra as well as transition states in photoactivated chromophores and is a key ingredient to performing molecular dynamics simulations of the system in the excited state. The implementation of excited states forces in `ONETEP` is a requirement for studying the photoisomerisation processes taking place in a number of large pigment-protein complexes.

A. Analytic TDDFT forces in `ONETEP`

Following the excited states gradient formalism introduced by Furche and Ahlrichs[7], the force due to a given TDDFT excited state can be split into three basic components: A direct term due to the redistribution of charge in the system that can be computed from the electron-hole difference density constructed via the response density $\mathbf{P}^{\{1\}}$, a relaxation term referred to as the Z-vector term that is due to the orthogonality constraint between the occupied and the unoccupied Kohn-Sham subspace and Pulay force terms that are due to the incompleteness and locality of the given basis set representation.

Since the NGWF representation in `ONETEP` is optimised to ideally represent the ground state density matrix, ground state Pulay forces are vanishingly small and can often be ignored in practical calculations. Given that in the TDDFT implementation, the excitation is represented by a compact NGWF representation of $\{\chi_\alpha\}$ and $\{\phi_\beta\}$ that ideally spans the relevant subspace for low-lying excitations, it is vital to investigate carefully to what degree Pulay forces are an important contribution to the excited state gradient and how they can be efficiently evaluated. For this reason, at this time we have implemented only the direct and the Z-vector force terms in the `ONETEP` code, as a first step.

In the mixed $\{\chi_\alpha\}$ and $\{\phi_\beta\}$ representation, the direct term can be written as

$$\mathbf{F}_{\text{direct}} = \frac{\partial \Omega}{\partial \mathbf{R}_\gamma} = \text{Tr}\left[\frac{\partial \mathbf{H}^\chi}{\partial \mathbf{R}_\gamma}\mathbf{P}^{\{\text{elec}\}}\right] - \text{Tr}\left[\frac{\partial \mathbf{H}^\phi}{\partial \mathbf{R}_\gamma}\mathbf{P}^{\{\text{hole}\}}\right]. \tag{10}$$

where the effective electron and hole density matrices are constructed via

$$\mathbf{P}^{\{\text{elec}\}} = \mathbf{P}^{\{1\}}\mathbf{S}^\phi\mathbf{P}^{\{1\}\dagger} \tag{11}$$

$$\mathbf{P}^{\{\text{hole}\}} = \mathbf{P}^{\{1\}\dagger}\mathbf{S}^\chi\mathbf{P}^{\{1\}} \tag{12}$$

Following the formalism introduced by Furche and Ahlrichs[7], the Z-vector force term can be evaluated by introducing the effective Z-vector density matrix $\mathbf{P}^{\{z\}}$, which in `ONETEP`, similarly to $\mathbf{P}^{\{1\}}$, is written in mixed conduction-valence NGWF representation. It is then possible to write the Z-vector force term as

$$\mathbf{F}_{\text{z}} = \text{Tr}\left[\frac{\partial \mathbf{H}^{\chi\phi}}{\partial \mathbf{R}_\gamma}\mathbf{P}^{\{z\}}\right] \tag{13}$$

where $\mathbf{P}^{\{z\}}$ can be obtained via solving the linear system

$$\mathbf{P}^{\{c\}}\mathbf{H}^\chi\mathbf{P}^{\{z\}} - \mathbf{P}^{\{z\}}\mathbf{H}^\phi\mathbf{P}^{\{v\}} + 2\mathbf{P}^{\{c\}}\mathbf{V}_{\text{SCF}}^{\{1\}\chi\phi}\left[\mathbf{P}^{\{z\}}\right]\mathbf{P}^{\{v\}} = -\mathbf{R}^{\chi\phi} \tag{14}$$

with the matrix $\mathbf{R}^{\chi\phi}$ given by

$$\begin{aligned}\mathbf{R}^{\chi\phi} &= \mathbf{P}^{\{c\}}\mathbf{V}_{\text{SCF}}^{\{1\}\chi\chi}\left[\mathbf{P}^{\{1\}}\right]\mathbf{P}^{\{1\}} - \mathbf{P}^{\{1\}}\mathbf{V}_{\text{SCF}}^{\{1\}\phi\phi}\left[\mathbf{P}^{\{1\}}\right]\mathbf{P}^{\{v\}} \\ &\quad + \mathbf{P}^{\{c\}}\mathbf{V}_{\text{SCF}}^{\{1\}\chi\phi}\left[\mathbf{P}^{\{x\}}\right]\mathbf{P}^{\{v\}} + \mathbf{P}^{\{c\}}\mathbf{V}_{\text{SCF}}^{\{2\}\chi\phi}\left[\mathbf{P}^{\{1\}}\right]\mathbf{P}^{\{v\}}.\end{aligned} \tag{15}$$

	x	y	z
Si	0.0	5.39	0.0
H1	0.47	0.33	0.0
H2	-0.47	0.33	0.0
H3	0.0	-2.88	1.60
H4	0.0	-2.90	1.64

TABLE I: Forces in eV/Å of the S_1 state of Silane as calculated with a minimum basis set of 4 NGWFs per Si and 1 NGWF per H

	x	y	z
Si	0.0	4.14	0.05
H1	0.46	0.09	0.0
H2	-0.46	0.09	0.0
H3	0.0	-2.14	0.89
H4	0.0	-2.16	0.92

TABLE II: Forces in eV/Å of the S_1 state of Silane as calculated with a large basis set of 34 NGWFs per Si and 34 NGWF per H

Here, $\mathbf{P}^{\{x\}} = \mathbf{P}^{\{\mathrm{elec}\}} - \mathbf{P}^{\{\mathrm{hole}\}}$ and $\mathbf{V}^{\{2\}\chi\phi}_{\mathrm{SCF}}$ denotes second order response potential of the system. Note that the linear system can be solved for in linear-scaling effort as long as all density matrices become sparse in sufficiently large systems, thus preserving the linear-scaling properties of the original TDDFT algorithm. Furthermore, note that the linear system can be preconditioned in exactly the same way as the TDDFT eigenvalue equation, thus yielding very fast convergence for the Z-vector matrix.

B. Preliminary results

The direct and the z-vector relaxation term have been implemented in `ONETEP` and we have performed a number of preliminary tests on small molecules like Silane (SiH_4) with a single compressed Si-H bond to break the 3-fold degeneracy of the lowest excited state (See Tables I, II, III for the results obtained in these calculations). Tables I and II contain benchmark calculations with different sizes of the NGWF representation, while Table III shows the forces obtained from a finite difference scheme.

As can be seen, the finite difference forces differ significantly from the results obtained with the smallest `ONETEP` representation but are in much better agreement when compared to the calculation using a very large NGWF representation. From this result, the conclusion can be drawn that unlike in ground state DFT, Pulay forces are very significant in the `ONETEP` TDDFT implementation and have to be accounted for. In the very large basis set limit, the Pulay forces become relatively small as would be expected. However, for realistic sizes of NGWF representations, they are certainly too large to be ignored in any practical calculation.

In order to test that the direct and the z-vector term are implemented correctly in `ONETEP`, we can perform a further test. Finite difference force calculations are repeated using the smallest NGWF representation but keeping the NGWF orbitals fixed in space while moving the atoms. This removes any Pulay forces from the calculations, and the finite difference forces obtained in this way agree perfectly with the forces obtained in Table I. It can be summarised that a first step to implementing analytic TDDFT forces has been taken, however it is clear that a correct treatment of Pulay forces

	x	y	z
Si	0.0	4.10	0.0
H1	0.58	0.17	0.0
H2	-0.58	0.17	0.0
H3	0.0	-2.22	1.06
H4	0.0	-2.22	1.06

TABLE III: Forces in eV/Å of the S_1 state of Silane as calculated with a finite-difference technique

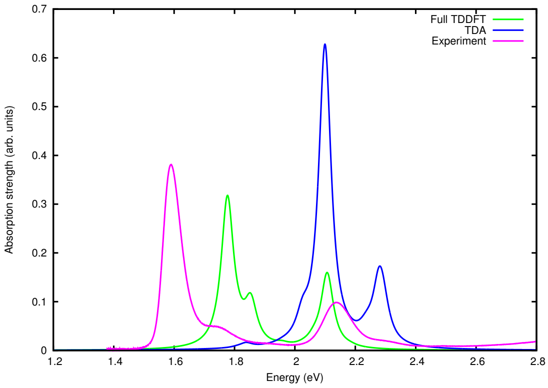

FIG. 3: Absorption spectrum of Bacteriochlorophyll in Toluene as computed with TDA and full TDDFT and in comparison with experimental results

is vital in realistic NGWF representations. Implementing these Pulay forces following the formalism detailed in the work by Furche and Ahlrichs[7] is an area of ongoing code development.

V. FULL TDDFT IN `ONETEP`

The original implementation of TDDFT in `ONETEP` is only available in the simplified Tamm-Dancoff approximation. This simplified approximation makes the eigenvalue equation hermitian and thus simpler to solve, but breaks time reversal symmetry and important oscillator sum rules. Furthermore, it is found that the approximation introduces significant errors in some widely studied pigments like bacteriochlorophyll (BChl)[9]. See Fig. 3 for a comparison of the performance of full TDDFT and the TDA for BChl in Toluene as compared with experimental data. As it can be seen the main peak is overestimated in energy by around 0.5 eV in the TDA and the spectral shape is poorly represented. Since a main focus of this eCSE proposal is to enable the calculation of optical properties of widely studied pigment-protein complexes for the ARCHER community, it is found that a full TDDFT implementation is highly desirable in order to make `ONETEP` an effective tool for this class of systems.

A. Details of the implementation

While the full TDDFT eigenvalue equation is non-hermitian, it is still possible to formulate a variational principle in the spirit of Eqn. 8 that yields the smallest *positive* eigenvalue of the system. It was demonstrated by Tsiper[8] that after an appropriate change in variables $\mathbf{p} = \mathbf{X} - \mathbf{Y}$ and $\mathbf{q} = \mathbf{X} + \mathbf{Y}$, an effective functional in conduction-valence Kohn-Sham space can be written as

$$\omega_{\text{min}} = \min_{(\mathbf{p}\,\mathbf{q})} \Omega_{\text{Tsip}}(\mathbf{p}, \mathbf{q}) = \min_{(\mathbf{p}\,\mathbf{q})} \frac{\begin{pmatrix} \mathbf{p}^{\dagger} & \mathbf{q}^{\dagger} \end{pmatrix} \begin{pmatrix} \mathbf{A} - \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} + \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix}}{|\mathbf{p}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{p}|}.$$

In order to carry out the minimisation of the functional of Eqn. 16, it becomes necessary to reformulate it in mixed $\{\chi_{\alpha}\}$-$\{\phi_{\beta}\}$ representation, just like in the original Tamm-Dancoff implementation. This can be achieved by introducing the effective response density matrices $\mathbf{P}^{\{p\}}$ and $\mathbf{P}^{\{q\}}$, which allow the functional to be reformulated as[9]

$$\omega_{\text{min}} = \min_{\{\mathbf{P}^{\{p\}}, \mathbf{P}^{\{q\}}\}} \left\{ \frac{\text{Tr}\left[\mathbf{P}^{\{p\}\dagger}\mathbf{S}^{\chi}\mathbf{f}^{\chi\phi}_{\{p\}}\mathbf{S}^{\phi}\right]}{2\left|\text{Tr}\left[\mathbf{P}^{\{p\}\dagger}\mathbf{S}^{\chi}\mathbf{P}^{\{q\}}\mathbf{S}^{\phi}\right]\right|} + \frac{\text{Tr}\left[\mathbf{P}^{\{q\}\dagger}\mathbf{S}^{\chi}\mathbf{f}^{\chi\phi}_{\{q\}}\mathbf{S}^{\phi}\right]}{2\left|\text{Tr}\left[\mathbf{P}^{\{p\}\dagger}\mathbf{S}^{\chi}\mathbf{P}^{\{q\}}\mathbf{S}^{\phi}\right]\right|} \right\}$$

where

$$\underline{\mathbf{f}}^{\chi\phi}_{\text{Tsip}} = \begin{pmatrix} \mathbf{f}^{\chi\phi}_{\{p\}} \\ \mathbf{f}^{\chi\phi}_{\{q\}} \end{pmatrix} = \begin{pmatrix} \mathbf{P}^{\{c\}}\mathbf{H}^{\chi}\mathbf{P}^{\{p\}} - \mathbf{P}^{\{p\}}\mathbf{H}^{\phi}\mathbf{P}^{\{v\}} \\ \mathbf{P}^{\{c\}}\mathbf{H}^{\chi}\mathbf{P}^{\{q\}} - \mathbf{P}^{\{q\}}\mathbf{H}^{\phi}\mathbf{P}^{\{v\}} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ 2\mathbf{P}^{\{c\}}\left(\mathbf{V}^{\{1\}\chi\phi}_{\text{SCF}}\left[\mathbf{P}^{\{q\}}\right]\right)\mathbf{P}^{\{v\}} \end{pmatrix}.$$

Minor changes to the generalised linear-response module allow for the construction of $\underline{\mathbf{f}}^{\chi\phi}_{\text{Tsip}}$ from two effective response density matrices $\mathbf{P}^{\{p\}}$ and $\mathbf{P}^{\{q\}}$. Note that for the same reason the linear-response operator can be evaluated in linear scaling effort in standard Tamm-Dancoff TDDFT, an evaluation of $\underline{\mathbf{f}}^{\chi\phi}_{\text{Tsip}}$ can be achieved in linear-scaling effort if all density matrices involved become sparse for sufficiently large system sizes. The functional in Eqn. 16 is then minimised using the same preconditioned conjugate gradient algorithm as in the Tamm-Dancoff TDDFT such that the calculation of full TDDFT eigenstates in `ONETEP` is enabled following only minor changes to existing code.

B. Benchmark results

In order to test the implementation of full TDDFT, we have performed a number of benchmark calculations on small molecules and compared to existing codes. In Table IV, results for a benchmark calculation on alizarin can be found in comparison with results obtained from a large diffuse gaussian basis set. As can be seen, the agreement between `ONETEP` and the gaussian basis set results for full TDDFT is of the same quality as the agreement in the Tamm-Dancoff approximation, demonstrating that the implementation works as expected.

Furthermore, we have performed a fully linear-scaling test of a classical molecular dynamics snapshot of BChl in Toluene, with the largest system size studied containing almost 7000 atoms. The results of the test can be found in fig. 4. As can be seen, applying the TDDFT operator does scale fully linearly and system sizes of the same order of magnitude as full pigment-protein complexes can be reached in practical calculations.

While full TDDFT comes with a computational overhead compared to the simplified Tamm-Dancoff approximation, this overhead is relatively small for most practical calculations. Tests performed on smaller BChl in Toluene systems of 770 atoms have shown a total increase in wall time by a factor of 1.7 when using full TDDFT instead of the TDA. However, given the significant improvement in spectral properties of pigments that are of vital importance in many biological systems, we can conclude that full TDDFT calculations are necessary in this class of systems. The extension of the capabilities of the

	`ONETEP` TDA	NWChem TDA	`ONETEP` RPA	NWChem RPA
1	2.233	2.192	2.184	2.149
2	3.520(0.047)	3.524(0.091)	3.516(0.286)	3.518(0.186)
3	3.536(0.001)	3.546(0.001)	3.489(0.002)	3.499(0.003)
4	3.720(1.094)	3.681(1.025)	3.408(0.499)	3.379(0.751)
5	3.822	3.866	3.821	3.865
6	3.875(0.003)	3.923(0.001)	3.874	3.922
7	4.234(0.001)	4.268(0.001)	4.229(0.001)	4.262(0.001)
8	4.315	4.305	4.241	4.230

TABLE IV: Lowest eight excited states of azobenzene, as computed with `ONETEP` and NWChem. Energies are given in eV, oscillator strengths are shown in brackets. States without specified oscillator strengths are dark. The states have been reordered according to their character, such that the order is the same as for the `ONETEP` TDA results. For the NWChem calculations, an aug-cc-pVTZ Gaussian basis set is used. TDA denotes the Tamm-Dancoff approximation, while RPA is used to denote a solution to the full TDDFT equation.

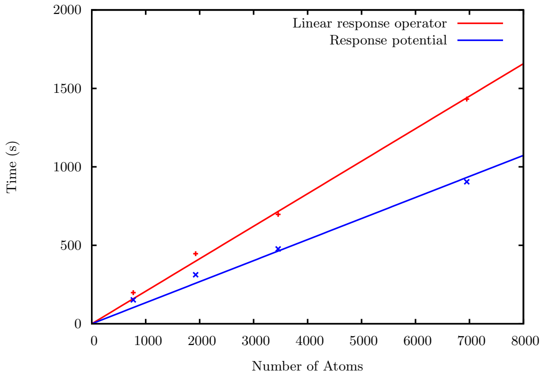

FIG. 4: Time taken for applying the TDDFT operator with number of atoms in the system

`ONETEP` code to perform these calculations as described here thus ensures that the code can become a valuable tool for the wider ARCHER user base for studying a wide range of biological systems such as pigment-protein complexes.

VI. CONCLUSION AND OUTLOOK

In conclusion, it can be stated that the scope of `ONETEP` as a tool for theoretical spectroscopy studies using high-performance computing facilities has been significantly expanded thanks to the funding provided by the ARCHER eCSE project. A large-scale benchmark test on a realistic test system has confirmed an excellent parallel scalability using the hybrid OpenMP/MPI parallelism approach, allowing for calculations of system sizes of up to 10,000 atoms on thousands of cores, while making excellent use of the computational resources available on ARCHER. An effective preconditioning scheme has been implemented that will result in a reduction of the total wall time in standard calculations by at least a factor of two. Furthermore, a first step has been taken in implementing the analytic evaluation excited state gradients in `ONETEP`, which, with the help of future development regarding the evaluation of Pulay correction terms, will enable to computation of transition states in large photoactivated pigment-protein complexes. Furthermore, the simplified Tamm-Dancoff TDDFT approach has been extended to allow for the computation of excited states using full TDDFT, which, with a minor increase in computational cost, can yield a significant improvement of spectra of pigments that are important in many biological systems. With the exception of the analytical evaluation of the TDDFT forces, all changes reported here have been included in the newest release version of the `ONETEP` code and thus have been made directly available to the ARCHER community.

Acknowledgments

This work was funded under the embedded CSE programme of the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk)

---

[1] P. D. Haynes, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, J. Chem. Phys. **122**, 084119 (2005)

[2] T. J. Zuehlsdorff, N. D. M. Hine, J. S. Spencer,N. M. Harrison, D. J. Riley, and P. D. Haynes, J. Chem. Phys. **139**, 064104 (2013)

[3] M. E. Casida, *Recent Advances in Density Functional Methods*, In: Chong, D.P. (ed.) vol. 1. World Scientific, Singapore (1995)

[4] S. Hirata, and M. Head-Gordon, Chem. Phys. Lett. **314**, 291-299 (1999)

[5] L. E. Ratcliff, N. D. M. Hine, and P. D. Haynes, Phys. Rev. B **84**, 165131 (2011)

[6] K. A. Wilkinson, N. D. M. Hine, and C.-K. Skylaris, J. Chem. Theory Comput. **10**, 4782 (2014)

[7] F. Furche and R. Ahlrichs, J. Chem. Phys. **117**, 7433 (2002)

[8] E. V. Tsiper, J. Phys. B **34**, L401 (2001)

[9] T. J. Zuehlsdorff, N. D. M. Hine, M. C. Payne, P. D. Haynes, Submitted (2015), http://arxiv.org/abs/1507.08157

[10] See www.onetep.org/Main/Archer