Efficient Algorithms for Estimating the Absorption Spectrum within Linear Response TDDFT

Jiri Brabec,*,† Lin Lin,*,†,‡ Meiyue Shao,† Niranjan Govind,¶ Chao Yang,*,†
Yousef Saad,§ and Esmond G. Ng†

† Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA
94720

‡ Department of Mathematics, University of California, Berkeley, California 94720, USA

¶ Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory,
Richland, WA 99352

§ Department of Computer Science and Engineering, University of Minnesota, Twin Cities,
Minneapolis, MN 55455

E-mail: jbrabec@lbl.gov; lin.lin@lbl.gov; cyang@lbl.gov

Abstract

We present a special symmetric Lanczos algorithm and a kernel polynomial method (KPM) for approximating the absorption spectrum of molecules within the linear response time-dependent density functional theory (TDDFT) framework in the product form. In contrast to existing algorithms, the new algorithms are based on reformulating the original non-Hermitian eigenvalue problem as a product eigenvalue problem and the observation that the product eigenvalue problem is self-adjoint with respect to

an appropriately chosen inner product. This allows a simple symmetric Lanczos algorithm to be used to compute the desired absorption spectrum. The use of a symmetric Lanczos algorithm only requires half of the memory compared with the non-symmetric variant of the Lanczos algorithm. The symmetric Lanczos algorithm is also numerically more stable than the non-symmetric version. The KPM algorithm is also presented as a low-memory alternative to the Lanczos approach, but may require more matrix-vector multiplications in practice.

We discuss the pros and cons of these methods in terms of their accuracy as well as their computational and storage cost. Applications to a set of small and medium-sized molecules are also presented.

1 Introduction

Time-Dependent Density Functional Theory (TDDFT)[1,2] has emerged as an important tool for reliable excited-state calculations for a broad spectrum of applications from molecular to materials systems. The most common formulation of TDDFT in quantum chemistry is in the frequency domain via linear response theory or the Casida formulation.[3,4] This approach is also known as linear-response (LR) TDDFT and is widely used to calculate absorption spectra. Computing the absorption spectrum with this approach involves solving a non-Hermitian eigenvalue problem. Formally, the numerical cost to diagonalize the full LR-TDDFT matrix equations scales as $\mathcal{O}(N^6)$, where N is the total number of molecular orbitals (MO).[5] As a result, for large systems, this approach becomes expensive if a large number of excitations ($\sim 10^3$–10^4) are needed. It can be shown, that the original matrix equation can be unitarily transformed into a form that decouples into two equivalent product eigenvalue problems half of the size of the original problem.[6] The iterative eigensolvers are typically used to solve lowest lying excited states[7] or the excited states within a given energy window.[8,9] In the worst scenario these algorithms scale as $O(N^5)$, more advanced techniques including Krylov subspace approaches and linear-scaling methods can reduce the cost to

$O(N^3)$ or less.[5,10] We refer the reader to more comprehensive reviews of LR-TDDFT.[3,4,11–13]

In many cases like large molecular complexes and high density of states (DOS) materials, excitations over a wider energy range may be required. This often results in a very demanding calculation[14,15] even when iterative eigensolvers are used. Over the years several approaches have been developed to tackle this problem including the complex polarization,[16] damped response approaches,[17,18] multishift linear solvers,[19] variational DFT approach of Ziegler and co-workers,[20] the simplified approaches from Grimme and co-workers[21,22] and the efficient LR approach of Neuhauser and Baer.[23] Alternative approaches like real-time time-dependent density functional theory (RT-TDDFT)[24,25] in combination with a weak delta-function field have also been used to tackle this problem in the time domain. Despite these algorithmic developments, it is desirable to look for novel approximate, yet accurate, excited-state approaches for large systems.

In this paper, we explore efficient ways to estimate the absorption spectrum of a finite system in the frequency domain. Specifically, we are interested in methods that do not explicitly compute the eigenvalues and eigenvectors of the full LR-TDDFT matrix. Within the Tamm–Dancoff approximation (TDA),[26] the matrix to be diagonalized becomes Hermitian. In this case, the absorption spectrum can be approximated by the *Kernel Polynomial Method* (KPM), originally proposed to estimate the density of states of a symmetric matrix.[27] This approach has been discussed in Ref.[28] By casting the full LR-TDDFT eigenvalue problem as a product eigenvalue problem, we show that the KPM can be extended to the full LR-TDDFT equations. To the best of our knowledge, the product form has not been utilized in the computation of absorption spectra.

A "two-sided" Lanczos procedure was proposed in Ref.[29] to approximate the absorption spectrum in the full LR-TDDFT framework. This approach treats the Casida Hamiltonian as a non-Hermitian matrix and can be numerically unstable. We show that it is possible to use a more standard Lanczos algorithm with a properly chosen inner product to obtain an accurate approximation of the absorption spectrum. This approach improves the numerical

stability and reduces the memory and computational cost compared with a two-sided Lanczos procedure.

The rest of the paper is organized as follows. For completeness, we first derive the expression to be computed in the absorption spectrum estimation using matrix notation. Although this result is well known (see, e.g., Ref. [29]), our derivation highlights the relationship between several quantities associated with the LR-TDDFT eigenvalue problem. We then present the Lanczos and KPM algorithms for estimating the absorption spectrum in Section. 3. We also discuss how to use the Lanczos and KPM algorithms to compute the density of states (DOS) of the LR-TDDFT eigenvalue problem. Computational results that demonstrate the effectiveness of the Lanczos algorithm are presented in Section. 6, where we compare the accuracy and cost of the Lanczos and KPM approaches. All algorithms discussed in this paper have been implemented in a development version of the NWChem [30] program.

2   Linear Response and Absorption Spectrum

In the linear response regime of TDDFT, the optical absorption spectrum of a finite system can be obtained from the trace of the 3×3 dynamic polarizability tensor $\alpha_{\mu,\nu}$ defined as

$$\alpha_{\mu,\nu}(\omega) = \langle\mu| - \frac{1}{\pi}\mathrm{Im}\chi(\omega)|\nu\rangle = -\frac{1}{\pi}\mathrm{Im}\langle\mu|\chi(\omega)|\nu\rangle, \tag{1}$$

where μ and ν are one of the coordinate variables x, y and z, and $\chi(\omega)$ characterizes the linearized charge density response $\Delta\rho$ to an external frequency dependent potential perturbation $v_{\mathrm{ext}}(r;\omega)$ of the ground state Kohn–Sham Hamiltonian in the frequency domain, i.e.

$$\Delta\rho(r;\omega) = \int \chi(r,r';\omega)v_{\mathrm{ext}}(r';\omega)dr'.$$

The symbol $\mathrm{Im}\chi$ denotes the imaginary part of χ, which is also called the spectral function of χ. Further discussion about the imaginary part is given in Appendix. A.

It is well known[11] that χ can be expressed as

$$\chi(r, r'; \omega) = \int \varepsilon^{-1}(r, r'''; \omega)\chi_0(r''', r'; \omega)dr''', \tag{2}$$

where ε is called the dielectric operator defined as

$$\varepsilon(r, r'''; \omega) = \delta(r, r''') - \int \chi_0(r, r''; \omega)f_{Hxc}(r'', r''')dr'' \tag{3}$$

and ε^{-1} is the inverse of the dielectric operator in the operator sense. In TDDFT, within the adiabatic approximation, the Hartree–exchange–correlation kernel f_{Hxc} is frequency-independent, and is defined as

$$f_{Hxc}(r, r') = \frac{1}{|r - r'|} + \frac{\delta v_{xc}[\rho](r)}{\delta \rho(r')},$$

with v_{xc} being the static exchange–correlation potential. The retarded irreducible independent-particle polarization function χ_0 is defined by

$$\chi_0(r, r'; \omega) = \sum_j \sum_a \frac{\phi_j(r)\phi_a^*(r)\phi_j^*(r')\phi_a(r')}{\omega - \Delta\varepsilon_{a,j} + \mathrm{i}\eta} - \frac{\phi_j^*(r)\phi_a(r)\phi_j(r')\phi_a^*(r')}{\omega + \Delta\varepsilon_{a,j} + \mathrm{i}\eta}, \tag{4}$$

where (ε_j, ϕ_j) and (ε_a, ϕ_a) are eigenpairs of the ground state self-consistent Kohn–Sham Hamiltonian. Here j and a are indices of the occupied and virtual Kohn–Sham eigenfunctions (orbitals), $\Delta\varepsilon_{a,j} \equiv \varepsilon_a - \varepsilon_j$, and $\Delta\varepsilon_{a,j} > 0$ by definition. Also $\eta > 0$ is an infinitesimally small constant to keep Eq. (4) well defined for all ω.

Since we consider the excitation properties of finite systems, without loss of generality we can assume that all Kohn–Sham eigenfunctions ϕ_j and ϕ_a are real. To simplify the notation, we will use $\Phi(r)$ to denote a matrix that contains all products of occupied and virtual wavefunction pairs, i.e., $\Phi(r) = \left[\phi_j(r)\phi_a^*(r), \dots\right]$ and D_0 to denote a diagonal matrix

that contains $\Delta\varepsilon_{a,j}$ on its diagonal. Using this notation, we can rewrite χ_0 as

$$\chi_0(r, r'; \omega) = [\Phi(r), \Phi(r)] \begin{bmatrix} (\omega + i\eta)I - D_0 & 0 \\ 0 & -(\omega + i\eta)I - D_0 \end{bmatrix}^{-1} \begin{bmatrix} \Phi(r') \\ \Phi(r') \end{bmatrix}$$

To simplify the notation further, we define

$$C = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \quad D = \begin{bmatrix} D_0 & 0 \\ 0 & D_0 \end{bmatrix}, \quad \text{and} \quad \hat{\Phi}(r) = [\Phi(r), \Phi(r)],$$

which allows us to rewrite χ_0 succinctly as

$$\chi_0(r, r'; \omega) = \hat{\Phi}(r) \left[(\omega + i\eta)C - D \right]^{-1} \hat{\Phi}(r'). \tag{5}$$

We denote by $\Phi, \hat{\Phi}$ the finite dimensional matrices obtained by discretizing $\Phi(r), \hat{\Phi}(r)$ on real space grids, respectively. Similarly, the discretized $\hat{\Phi}(r')$ can be viewed as the matrix transpose of $\hat{\Phi}$. Replacing all integrals with the matrix–matrix multiplication notation and making use of the Sherman–Morrison–Woodbury formula for manipulating a matrix inverse, we can show that (see Appendix. B for a detailed derivation)

$$\chi(\omega) = \hat{\Phi} \left[(\omega + i\eta)C - \Omega \right]^{-1} \hat{\Phi}^T \tag{6}$$

where

$$\Omega \equiv \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \begin{bmatrix} D_0 + \Phi^T f_{Hxc} \Phi & \Phi^T f_{Hxc} \Phi \\ \Phi^T f_{Hxc} \Phi & D_0 + \Phi^T f_{Hxc} \Phi \end{bmatrix}. \tag{7}$$

Here $\Phi^T f_{Hxc} \Phi$ is an $n_o n_v \times n_o n_v$ matrix commonly known as the *coupling matrix*, where n_o and n_v are the number of occupied and virtual states, respectively. The $(j, a; j', a')$th

element of the matrix is evaluated as

$$\int dr dr' \phi_j(r)\phi_a(r) f_{Hxc}(r, r')\phi_j'(r')\phi_a'(r').$$

It follows from (6) that

$$\langle x|\chi(\omega)|x\rangle = \hat{x}^T \left[(\omega + \mathrm{i}\eta)C - \Omega\right]^{-1} \hat{x} \tag{8}$$

where $\hat{x} = \left[\hat{x}_1^T, \hat{x}_1^T\right]^T$, and $\hat{x}_1 = \Phi^T x$ is a column vector of size $n_o n_v$. The (j, a)th element of $\hat{x}_1$ can be evaluated as

$$\int x\phi_j(r)\phi_a(r)dr. \tag{9}$$

It can be easily verified that

$$\langle x|\chi(\omega)|x\rangle = \hat{x}^T \left[(\omega + \mathrm{i}\eta)I - H\right]^{-1} C\hat{x}, \tag{10}$$

where

$$H = \begin{bmatrix} A & B \\ -B & -A \end{bmatrix} \tag{11}$$

and is also referred to as the Casida or LR-TDDFT matrix equations. Although H is non-Hermitian, it has a special structure that has been examined in detail in previous work.[12,31,32] In particular, when both $K \equiv A-B$ and $M \equiv A+B$ are positive definite, it can be shown that the eigenvalues of H come in positive and negative pairs $(-\lambda_i, \lambda_i)$, $\lambda_i > 0$, $i = 1, 2, \ldots, n_o n_v$. If $\left[u_i^T, v_i^T\right]^T$ is the right eigenvector associated with λ_i, the left eigenvector associated with the same eigenvalue is $\left[u_i^T, -v_i^T\right]^T$.

It follows from the eigendecomposition of H (see Appendix. C) and (10) that

$$\langle x|\chi(\omega)|x\rangle = \sum_{i=1}^{n_o n_v} \left(\frac{\left[\hat{x}_1^T(u_i + v_i)\right]^2}{\omega - \lambda_i + \mathrm{i}\eta} - \frac{\left[\hat{x}_1^T(u_i + v_i)\right]^2}{\omega + \lambda_i + \mathrm{i}\eta} \right), \tag{12}$$

where the eigenvectors satisfy the normalization condition $u_i^T u_i - v_i^T v_i = 1$. In the limit of $\eta \to 0^+$, $\alpha_{x,x} = -\frac{1}{\pi} \text{Im} \langle x | \chi(\omega) | x \rangle$ becomes

$$\sum_{i=1}^{n_o n_v} \left[\hat{x}_1^T (u_i + v_i) \right]^2 \left[\delta(\omega - \lambda_i) - \delta(\omega + \lambda_i) \right] = \hat{x}^T \delta(\omega I - H) C \hat{x}.$$

Hence, the absorption spectrum has the form

$$\sigma(\omega) = -\frac{1}{3\pi} \lim_{\eta \to 0^+} \text{Im} \left(\langle x | \chi(\omega) | x \rangle + \langle y | \chi(\omega) | y \rangle + \langle z | \chi(\omega) | z \rangle \right) = \frac{1}{3} \sum_i f_i^2 \left[\delta(\omega - \lambda_i) - \delta(\omega + \lambda_i) \right],$$
(13)

where $f_i^2 = \left[\hat{x}_1^T (u_i + v_i) \right]^2 + \left[\hat{y}_1^T (u_i + v_i) \right]^2 + \left[\hat{z}_1^T (u_i + v_i) \right]^2$ is known as the *oscillator strength*.

It is not difficult to show that $w_i \equiv u_i + v_i$ is the ith left eigenvector of the $n_o n_v \times n_o n_v$ matrix MK, associated with the eigenvalue λ_i^2. The corresponding right eigenvector is in the direction of $u_i - v_i$. Therefore, the absorption spectrum can be obtained by computing the eigenvalues of KM, which is half of the size of H. It can be shown that $\alpha_{x,x}$ can be expressed as

$$\alpha_{x,x}(\omega) = 2 \, \text{sign}(\omega) \hat{x}_1^T K \delta(\omega^2 I - MK) \hat{x}_1,$$
(14)

where the matrix function $\delta(\omega^2 I - MK)$ should be understood in terms of the eigendecomposition of MK, i.e., $\delta(\omega^2 I - MK) = \sum_{i=1}^{n_o n_v} \delta(\omega^2 - \lambda_i^2)(u_i - v_i)(u_i + v_i)^T$. The derivation can be found in the Appendix. C. In (13), the value of $\omega = \lambda_j > 0$ is often referred to as the jth excitation energy, whereas $\omega = -\lambda_j < 0$ is often referred to as the jth deexcitation energy. Because the oscillator strength factors associated with these energy levels are identical in magnitude and opposite in signs, it is sufficient to focus on just one of them. Here, we will only be concerned with excitation energies, i.e., *we assume $\omega > 0$*.

8

3 Algorithms for Approximating the Absorption Spectrum

If the eigenvalues and eigenvectors of H or MK are known, the absorption spectrum defined in (13) is easy to construct. However, because the dimensions of H and MK are $\mathcal{O}(n_o n_v)$, computing all eigenvalues and eigenvectors of these matrices is prohibitively expensive for large systems due to the $\mathcal{O}(n_o^3 n_v^3)$ complexity.

If we only need the excitation energies and oscillator strengths of the first few low excited states, we may use an iterative eigensolver such as Davidson's method or a variant of the locally optimal block preconditioned gradient method (LOBPCG) to compute the lowest few eigenpairs of MK.[12,33,34] These methods only require the user to provide a procedure for multiplying A and B with a vector. The matrix A or B does not need to be explicitly constructed. However, these methods tend to become prohibitively expensive when the number of eigenvalues within the desired energy range becomes large.

In this section, we introduce two methods that do not require computing eigenvalues and eigenvectors of the Casida Hamiltonian explicitly. In addition, these methods only require a procedure for multiplying A and B with a vector. Both approaches provide a satisfactory approximation to the overall shape of the absorption spectrum from which the position and the height of each major peak with a desired energy range can be easily identified.

3.1 The Lanczos Method

One way to estimate $\alpha_{x,x}$ is to use the Lanczos algorithm. Because MK is self-adjoint with respect to the inner produced induced by K, i.e.,

$$\langle v, MKv \rangle_K = v^T KMKv = \langle MKv, v \rangle_K, \tag{15}$$

we can use the K-inner product to generate a k-step Lanczos factorization of the form

$$MKQ_k = Q_kT_k + f_ke_k^T, \tag{16}$$

where

$$Q_k^TKQ_k = I \quad \text{and} \quad Q_k^TKf_k = 0, \tag{17}$$

and T_k is a tridiagonal matrix of size $k \times k$. The major steps of the K-inner product Lanczos procedure is shown in Algorithm 1. We use the MATLAB notation $Q(:, j)$ to denote the jth column of the matrix Q. In principle, because T_k is tridiagonal in exact arithmetic, columns of Q_k can be generated via a three-term recurrence. However, it is well known that as some of the eigenvalues of T_k converge to those of MK, and loss of orthogonality among columns of Q_k can occur due to potential instability in the numerical procedure. As a result, T_k may become singular with several spurious eigenvalues near zero.[35] To avoid loss of orthogonality, we perform full reorthogonalization as shown in steps 5–7 of Algorithm 1. Unless k is extremely large, the cost of full reorthogonalization is relatively small compared to the cost of multiplying A and B with vectors. However, full reorthogonalization does require keeping all columns of Q_k in memory.

If we choose the starting vector for the Lanczos iteration to be $\hat{x}_1$, i.e., $Q_ke_1 = \hat{x}_1/\sqrt{\hat{x}_1^TK\hat{x}_1}$, then $e_1^T\delta(\omega^2I - T_k)e_1^T$ serves as a good approximation of $\alpha_{x,x}(\omega)$.

To see why this is the case, let us first assume that $\delta(\omega^2I - MK)$ can be formally approximated by a k-th degree polynomial of the form $p_k(MK; \omega^2)$, i.e.

$$\delta(\omega^2I - MK) \approx p_k(MK; \omega^2) = \sum_{i=0}^{k} \gamma_i(\omega^2)(MK)^i.$$

It follows from (16) that

$$p_k(MK; \omega^2)Q_k = Q_kp_k(T_k; \omega^2) + R_k,$$

Algorithm 1: A k-step Lanczos method in the $(A - B)$-inner product.

Input: Real symmetric matrices A and B defined in (7); A start-
ing vector v_0; The number of steps k.

Output: Matrices Q_k and T_k that satisfy (16) and (17)

1: $Q_k(:, 1) = v_0/\sqrt{v_0^T(A - B)v_0}$;
2: **for** $j = 1, \ldots, k$ **do**
3:     $w \leftarrow (A - B)Q_k(:, j)$;
4:     $w \leftarrow (A + B)w$;
5:     $h \leftarrow Q_k(:, 1 : j)^T(A - B)w$;
6:     $H(1 : j, j) = h$;
7:     $w \leftarrow w - Q_k(:, 1 : j)h$;
8:     **if** $j < k$ **then**
9:        $Q_k(:, j + 1) = w/\sqrt{w^T(A - B)w}$;
10:    **end if**
11: **end for**

---

where R_k is a residual matrix that vanishes when $k = n_v n_o$.

Consequently, we can show that

$$
\begin{aligned}
\lim_{\eta \to 0^+} -\frac{1}{\pi}\mathrm{Im}\langle x|\chi(\omega)|x\rangle &= \hat{x}_1^T K \delta(\omega^2 I - MK)\hat{x}_1 \\
&\approx e_1^T Q_k^T K p_k(MK; \omega^2)Q_k e_1 \\
&= e_1^T Q_k^T K \left[Q_k p_k(T_k; \omega^2) + R_k\right] e_1 \\
&\approx e_1^T p_k(T_k; \omega^2)e_1 \quad\quad\quad\quad (18) \\
&\approx e_1^T \delta(\omega^2 I - T_k)e_1 \\
&= \sum_{j=1}^{k} \tau_j^2 \delta(\omega^2 - \theta_j), \quad\quad\quad (19)
\end{aligned}
$$

where $\theta_j > 0$ is the jth eigenvalue of the $k \times k$ tridiagonal matrix T_k sorted in increasing order and τ_j is the first component of the jth eigenvector of T_k.

As shown in the Appendix. C, $\delta(\omega^2 - \theta_j)$ can be rewritten as

$$
\delta(\omega^2 - \theta_j) = \frac{1}{2\sqrt{\theta_j}}\delta\big(\omega - \sqrt{\theta_j}\big), \quad\quad (\omega > 0). \quad\quad\quad (20)
$$

11

Therefore,

$$\alpha_{x,x}(\omega) \approx \sum_{j=1}^{k} \tau_j^2 \delta(\omega^2 - \theta_j) = \sum_{j=1}^{k} \frac{\tau_j^2}{2\sqrt{\theta_j}} \delta(\omega - \sqrt{\theta_j}) \tag{21}$$

holds.

The cost of the Lanczos algorithm is proportional to the number of steps (k) in the Lanczos iterations. We would like to keep k as small as possible without losing desired information in Eq. (19). However, when k is small, (21) gives a few spikes at the square root of the eigenvalues of T_k, also called *Ritz values*. No absorption intensity is given at other frequencies. To estimate the intensity of the absorption at these frequencies, we replace $\delta(\omega - \sqrt{\theta_j})$ in (19) by either a Gaussian or a Lorentzian. The use of these regularization functions allows us to interpolate the absorption intensity from the square root of the Ritz values to any frequency. Replacing $\delta(\omega^2 - \theta_j)$ with a Lorentzian is equivalent to not taking the $\eta \to 0^+$ limit in (19). It is also equivalent to computing the $(1,1)$ entry of the matrix inverse $\left[(\omega^2 + i\eta)I - T_k\right]^{-1}$. This entry can be computed recursively by using a recursive expression that is related to continued fractions.[36] This is the method of Haydock.[37] However, since the cost of computing the inverse and the eigenvalue decomposition of a small tridiagonal matrix is negligibly small, we do not gain much by using the Haydock's recursion.

3.2 Kernel Polynomial Method

Another approach follows the kernel polynomial method (KPM) reviewed in Ref.[27] To use this method, it is convenient to consider the case of $\omega > 0$ and view $\alpha_{x,x}(\omega)$ as a function of $\varpi := \omega^2$. The general idea relies in expressing $\alpha_{x,x}$ (and similarly $\alpha_{y,y}$ and $\alpha_{z,z}$) formally by a polynomial expansion of the form

$$\hat{\alpha}_{x,x}(\varpi) = \sum_{k=0}^{\infty} \gamma_k \mathcal{T}_k(\varpi), \tag{22}$$

where

$$\hat{\alpha}_{x,x}(\varpi) = \frac{\sqrt{1-\varpi^2}}{2}\alpha_{x,x}(\omega) = \sqrt{1-\varpi^2}\,\hat{x}_1^T K\delta(\varpi I - MK)\hat{x}_1 = \sqrt{1-\varpi^2}\sum_i \lambda_i(\hat{x}_1^T w_i)^2\delta(\varpi - \lambda_i^2),$$

with w_i being the ith left eigenvector of MK and $\{\mathcal{T}_k(\varpi)\}$ being a set of orthogonal polynomials of degree k.

For instance, we choose $\{\mathcal{T}_k(\varpi)\}$ to be Chebyshev polynomials. Using the identity $\int \mathcal{T}_k(\varpi)\delta(\varpi - \lambda_i^2) = \mathcal{T}_k(\lambda_i^2)$, we can compute the expansion coefficients γ_k's by

$$
\begin{aligned}
\gamma_k &= \frac{2-\delta_{k0}}{\pi}\int_{-1}^{1}\frac{1}{\sqrt{1-\varpi^2}}\mathcal{T}_k(\varpi)\hat{\alpha}_{x,x}(\varpi)d\varpi \\
&= \frac{2-\delta_{k0}}{\pi}\hat{x}_1^T\left[\sum_{i=1}^{n_o n_v}\lambda_i w_i \mathcal{T}_k(\lambda_i^2)w_i^T\right]\hat{x}_1 \\
&= \frac{2-\delta_{k0}}{\pi}\hat{x}_1^T\left[\sum_{i=1}^{n_o n_v} K(u_i - v_i)\mathcal{T}_k(\lambda_i^2)(u_i + v_i)^T\right]\hat{x}_1 \\
&= \frac{2-\delta_{k0}}{\pi}\hat{x}_1^T K\mathcal{T}_k(MK)\hat{x}_1.
\end{aligned}
\tag{23}
$$

Here δ_{ij} is the Kronecker δ symbol so that $2-\delta_{k0}$ is equal to 1 when $k=0$ and to 2 otherwise. In deriving (23) from the line before, we use the property that $w_i = u_i + v_i$ is a left eigenvector of MK and the corresponding right eigenvector is in the direction of $u_i - v_i = \lambda_i K^{-1}(u_i + v_i)$.

Because Chebyshev polynomials $\mathcal{T}_k(\varpi)$ can be generated recursively using the three-term recurrence

$$\mathcal{T}_0(\varpi) = 1, \quad \mathcal{T}_1(\varpi) = \varpi, \quad \mathcal{T}_{k+1}(\varpi) = 2\mathcal{T}_k(\varpi) - \mathcal{T}_{k-1}(\varpi),$$

for $\varpi \in [-1,1]$, we do not need to construct $\mathcal{T}_k(MK)$ explicitly. We can apply $\mathcal{T}_k(MK)\hat{x}_1$ recursively using a 3-term recurrence also. This three-term recurrence allows us to implement the KPM by storing 3 or 4 vectors depending on whether an intermediate vector is used to hold intermediate matrix vector products. Pseudocode for the KPM is given in Algorithm 2. The cost of the KPM is dominated by the multiplication of the matrices A and B with vectors, and proportional to the degree of the expansion polynomial or the number of expansion terms

in (22).

The derivation above requires $\varpi \in [-1, 1]$. To generate Chebyshev polynomials on an interval $[a, b]$ where a and b are the estimated lower and upper bounds of the eigenvalues of MK, a proper linear transformation should be used to map ϖ to $[-1, 1]$ first.

In (22), we assume that the $\eta \to 0^+$ limit has been taken. Because the left hand side is a sum of Dirac-δ distributions, which is discontinuous, a finite least squares polynomial expansion will produce the well known Gibbs oscillations as shown in Ref.,[38] and lead to larger errors near the point of discontinuity. This effect is more pronounced for molecules of which excitation energies are well isolated, and is less severe for solids whose excitation energies are more closely spaced. To reduce the effects of the Gibbs oscillation, we can set η to a small positive constant (instead of taking the $\eta \to 0^+$ limit.) This is equivalent to replacing each Dirac-δ distribution with a Lorentzian of the form

$$g_\eta(\omega - \lambda_i) = \frac{1}{\pi} \frac{\eta}{(\omega - \lambda_i)^2 + \eta^2}.$$

In this case, the expansion coefficients γ_k's need to be computed in a different way. We may also replace the Dirac-δ distributions on the left hand side of (22) with Gaussians of the form

$$g_\sigma(\omega - \lambda_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\omega - \lambda_i)^2/(2\sigma^2)}, \tag{24}$$

where σ is a smoothing parameter that should be chosen according to the desired resolution of $\alpha_{x,x}$. We refer to the technique of replacing Dirac-δ distributions with Lorentzians or Gaussians as *regularization*. In Ref.,[38] we show how the expansion coefficients γ_k's can be computed recursively when Dirac-δ distributions are replaced with Gaussians, and $\mathcal{T}_k(\varpi)$ is taken to be the kth degree Legendre polynomial. We referred to this particular expansion as the Delta–Gaussian–Legendre (DGL) expansion.[38] Another technique for reducing the Gibbs phenomenon is the use of *Jackson damping*.[39] However, it has been shown in Ref.[38] that Jackson damping can lead to over-regularized spectrum.

Algorithm 2: The kernel polynomial method for approximating the absorption spectrum of a finite system.

Input: Real symmetric matrices A and B defined in (7); A set of points $\{\omega_i\}$, $i = 1, 2, \ldots, m$ at which the absorption spectrum is to be evaluated, the dipole vector $\hat{x}_1$ defined in (9), the degree k of the expansion polynomial used in KPM.

Output: Approximate absorption spectrum $\sigma(\omega_i)$ for $i = 1, 2, \ldots, m$.

1: Estimate the upper and lower bounds (λ_{ub} and λ_{lb}) of the eigenvalues of MK by a few step of the Lanczos iteration.

2: Let $c = (\lambda_{lb} + \lambda_{ub})/2$; $d = (\lambda_{ub} - \lambda_{lb})/2$;

3: Set $\zeta_j = 0$ for $j = 0, \ldots, k$;

4: Let $q_0 \leftarrow \hat{x}_1$;

5: **for** $j = 0, 1, \ldots, k$ **do**

6:     $w_A \leftarrow Aq_j$;

7:     $w_B \leftarrow Bq_j$;

8:     Compute $\zeta_j \leftarrow \zeta_j + q_0^T(w_A + w_B)$;

9:     Compute q_{j+1} via the three-term recurrence

$$q_{j+1} = 2\left[(A - B)(w_A + w_B) - c(w_A + w_B)\right]/d - q_{j-1}$$

(for $j = 0$, $q_1 = \left[(A - B)(w_A + w_B) - c(w_A + w_B)\right]/d$);

10: **end for**

11: Set $\gamma_j \leftarrow \dfrac{2 - \delta_{j0}}{n_o n_v \pi} \zeta_j$ for $j = 0, 1, \ldots, k$;

12: Evaluate

$$\alpha_{x,x}(\omega_i) = 2\sum_{j=0}^{k} \gamma_j \mathcal{T}_j\left(\frac{\omega_i^2 - c}{d}\right) \Big/ \sqrt{1 - \left(\frac{\omega_i^2 - c}{d}\right)^2}$$

for $i = 1, 2, \ldots, m$;

13: Set $q_0 = \hat{y}_1$ and $q_0 = \hat{z}_1$ and repeat Steps 5–10 to obtain $\alpha_{y,y}$ and $\alpha_{z,z}$;

14: $\sigma(\omega_i) = \left[\alpha_{x,x}(\omega_i) + \alpha_{y,y}(\omega_i) + \alpha_{z,z}(\omega_i)\right]/3$ for $i = 1, 2, \ldots, m$.

4 Estimating the Density of States of the LR-TDDFT Matrix Equations

In some cases, the oscillator strength associated with some of the excitation energies are negligibly small. These are sometimes referred to as *dark states*.[11] To reveal these states, it is sometimes useful to examine the density of states (DOS) associated with the LR-TDDFT matrix equations (11).

In the review paper,[38] several numerical algorithms for estimating the DOS of symmetric matrices are presented and compared. To apply these techniques, we make use of the fact that eigenvalues of H can be obtained from that of MK or $K^{1/2}MK^{1/2}$ which is Hermitian. As a result, the DOS can be written as

$$\phi(\omega) = \text{trace}(\omega^2 I - K^{1/2}MK^{1/2}). \tag{25}$$

We assume that $\omega \geq 0$ since $\phi(-\omega) = \phi(\omega)$.

If we use the KPM to approximate (25), i.e.,

$$\phi(\omega) \approx \sum_{k=0}^{\infty} \gamma_k \mathcal{T}_k(\omega^2), \tag{26}$$

the expansion coefficients γ_k can be computed as

$$\gamma_k = \frac{2 - \delta_{k0}}{n\pi} \text{trace}\left[\mathcal{T}_k(K^{1/2}MK^{1/2})\right] = \frac{2 - \delta_{k0}}{n\pi} \text{trace}\left[\mathcal{T}_k(MK)\right]. \tag{27}$$

Thus, apart from the scaling factor $(2 - \delta_{k0})/(n\pi)$, γ_k is the trace of $\mathcal{T}_k(MK)$, and such a trace can be estimated through a statistical sampling technique that involves computing

$$\zeta_k = \frac{1}{n_{\text{vec}}} \sum_{l=1}^{n_{\text{vec}}} \left(q_0^{(l)}\right)^T K \mathcal{T}_k(MK) q_0^{(l)}, \tag{28}$$

16

for a set of randomly generated vectors $q_0^{(1)}$, $q_0^{(2)}$, ..., $q_0^{(n_{\text{vec}})}$. One possible choice of the random vectors is that each entry of $q_0^{(l)}$ follows an independent Gaussian distribution $\mathcal{N}(0,1)$ for all $l = 1, \ldots, n_{\text{vec}}$. The method only requires multiplying A and B with a number of vectors, and the multiplication of $\mathcal{T}_k(MK)$ with a vector can be implemented through a three-term recurrence.

To estimate the DOS by the Lanczos method, we can simply use the K-orthogonal Lanczos factorization shown in (16) to obtain a sequence of tridiagonal matrices $T_k^{(l)}$, $l = 1, 2, \ldots, n_{\text{vec}}$, using a set of randomly generated starting vectors $q_0^{(1)}$, $q_0^{(2)}$, ..., $q_0^{(n_{\text{vec}})}$ as mentioned above. The approximate DOS can be expressed as

$$\phi(\omega) \approx \frac{1}{n_{\text{vec}}} \sum_{l=1}^{n_{\text{vec}}} \sum_i (\tau_i^{(l)})^2 \delta(\omega^2 - \theta_i^{(l)}), \tag{29}$$

where $\theta_i^{(l)}$ is the ith eigenvalue of $T_k^{(l)}$, $\tau_i^{(l)}$ is the first component of the corresponding eigenvector. To regularize the approximate DOS, we rewrite $\delta(\omega^2 - \theta_i^{(l)})$ in the form of (20), and replace $\delta(\omega - \sqrt{\theta_i^{(l)}})$ with properly defined Gaussian $g_\sigma(\omega - \sqrt{\theta_i^{(l)}})$.

5   Frozen Orbital and Tamm–Dancoff Approximation

When ω is relatively small compared with $\Delta\varepsilon_{a,j}$, the contribution of the corresponding term in (4) is relatively insignificant. Thus, it is reasonable to leave out these terms which constitute indices j's that correspond to the lowest occupied orbitals often known as the *core orbitals*. By using this *frozen core approximation* (FCA),[40] we effectively reduces the dimension of A and B matrices in (7) from $n_o n_v$ to $(n_o - n_f)n_v$, where n_f is the number of lowest occupied orbitals that are excluded (or "frozen") from (4). It is well known that the FCA preserves the low end of the absorption spectrum while removing poles of $\chi(\omega)$ at higher frequencies. Therefore, it is extremely useful to combine FCA with the Lanczos algorithm to obtain an accurate approximation to the absorption spectrum in the low excitation energy range at a reduced cost, as we will show in the next section.

Another commonly used technique to reduce the computational cost of estimating the absorption spectrum is to set the matrix B in (7) to zero. This approximation is often referred to as the Tamm–Dancoff approximation (TDA). It reduces the eigenvalue problem to a symmetric eigenvalue problem that involves the matrix A only. We do not consider the TDA in this paper.

6   Computational Results

We now present computational results to demonstrate the quality of the approximate absorption spectrum obtained from the algorithms presented in Section 3 and the efficiency of these algorithms. In particular, we compare the absorption spectrum obtained from the new algorithms with that obtained from a traditional diagonalization based approach and that obtained from a real-time TDDFT (RT-TDDFT) simulation that can be readily performed with the released version of the NWChem open source code. It has been shown that absorption spectra from RT-TDDFT simulations in the weak delta-function field limit are consistent with those obtained from LR-TDDFT calculations for both UV/Vis and core-level excitations.[9,15] In addition, RT-TDDFT has been shown to provide a reasonable approximation of the entire spectrum for large molecular complexes and high DOS systems. For these reasons and since our goal is to demonstrate the applicability of our new approaches to large systems, we have chosen the RT-TDDFT spectra as our reference where full diagonalization is prohibitive. For details of the RT-TDDFT approach, we refer the reader to Refs.[24,25] We also compare the efficiency of the new algorithms with the aforementioned approaches and show that the algorithms presented in Section. 3 are much faster.

6.1   Test Systems

Our test systems include two small dyes: 2,3,5-trifluorobenzaldehyde (TFBA) and 4'-hydroxybenzylidene-2,3-dimethylimidazoline (HBDMI) as well as two relatively large molecules F2N12S and

P3B2, respectively. The atomic configurations of these molecular systems are shown in Figure 1. All geometries were optimized in the ground state DFT calculation using the B3LYP[41] functional. We use the 6-31G(d)[42,43] basis set for all systems except F2N12S where the cc-pVDZ basis set[44] was used.
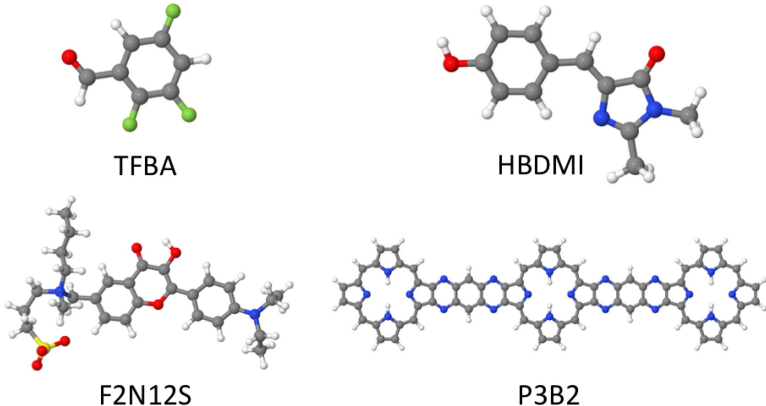

Figure 1: The molecules TFBA, HDMBI, F2N12S and P3B2.

In Table 1, we list n_o, n_v as well as the dimension of the excitation matrix $(n_o n_v)$ for each problem. We also list the number of frozen orbitals (n_f) when FCA is used, and the corresponding reduced $((n_o - n_f)n_v)$.

Table 1: The size of systems used in calculations, n_f denotes the number of frozen (core) orbitals. Geometries can be found in Supplementary information.[45]

System	n_o	n_v	$n_o n_v$	n_f	$(n_o - n_f)n_v$
TFBA	40	120	4,800	11	3,480
HBDMI	57	191	10,887	16	7,831
F2N12S	142	456	64,752	41	46,056
P3B2	305	1,059	322,995	92	225,567

The TFBA and HBDMI problems are relatively small. Thus, we can generate the matrices for these problems explicitly in NWChem and compute all eigenvalues and eigenvectors of these Hamiltonians. These eigenvalues and eigenvectors are used in the expression (13) to construct an "exact" linear response (LR) TDDFT absorption spectrum.

6.2 Approximation for Small Dyes

In Figure 2, we compare the approximate absorption spectrum obtained from the Lanczos algorithm with that obtained with traditional diagonalization for both the TFBA and HB-DMI molecules. We only show results in the $[0, 20]$ eV energy range for TFBA and in the $[0, 30]$ eV energy range for HBDMI, since these are often the ranges of interest in practice.

We observe that the absorption spectrum obtained from $k = 1200$ steps of Lanczos iterations is nearly indistinguishable from the exact solution for TFBA. All major peaks are correctly captured. Similarly, for HBDMI, 1200 Lanczos iterations are required to achieve the same level of accuracy.

(a)   (b)
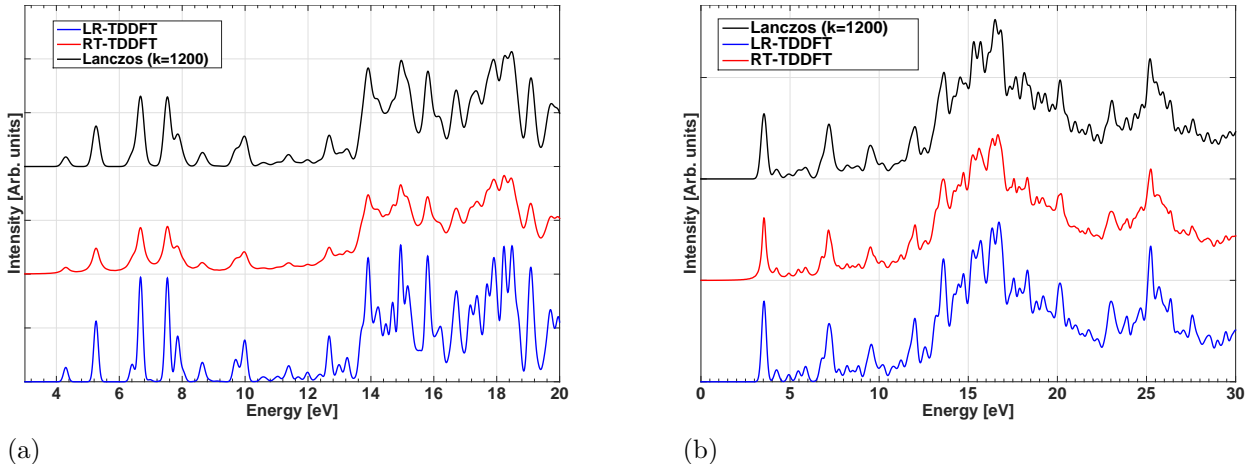
Figure 2: Comparison of the approximate absorption spectrum obtained from the Lanczos algorithm with that obtained from an exact LR-TDDFT calculation and RT-TDDFT calculation for (a) TFBA (b) HBDMI.

We also plot, in Figure 2, the absorption spectrum obtained from a real-time (RT) TDDFT calculation reported in Ref.[15] In this calculation, the time-dependent dynamic polarizability was calculated by propagating the solution to the time-dependent Kohn–Sham equation after an electric field pulse in the form of a small δ-kick and field strength 2×10^{-5} a.u. was applied at $t = 0$. The Fourier transform of the time-dependent polarizability yields the approximate absorption spectrum. Numerically, the RT-TDDFT calculations were carried out using a time step $\Delta t = 0.2$ a.u. (4.8 attoseconds), the size of the simulated trajectory

is 1000 a.u. (24.2 fs, 5000 steps). We can see from this figure that the LR-TDDFT absorption spectrum matches well with that obtained from RT-TDDFT for both TFBA and HBDMI. The nearly perfect match indicates the validity of linear response approximation. Therefore, it seems reasonable to compare the absorption spectrum obtained from the Lanczos algorithm with that obtained from RT-TDDFT directly for larger problems where the traditional diagonalization approach is prohibitively expensive.

6.3   The effect of FCA

In Figure 3, we illustrate the effect of using the frozen core approximation (FCA) for TFBA by setting n_f to 11. That effectively reduces the dimension of the A and B matrices in (11) from $n_o n_v = 4800$ to $(n_o - n_f)n_v = 3480$. We can clearly see from the DOS shown in Figure 3(a) that FCA preserves the small eigenvalues of the original matrix, but the largest eigenvalues are absent. The absence of these eigenvalues allows the Lanczos algorithm to obtain more Ritz values in the low energy range (e.g. [0,20] eV) in fewer iterations, thereby producing an accurate absorption spectrum with lower cost. The cost reduction in FCA results from both the reduction in the dimension of the matrix and the reduction in the steps of Lanczos iterations required to achieve a desired resolution in the absorption spectrum.

(a)

(b)
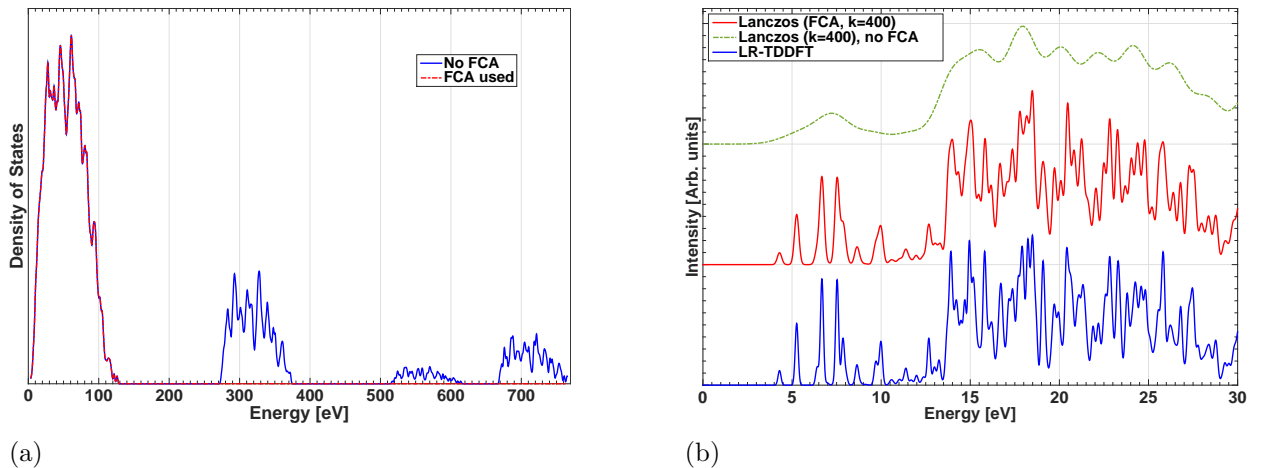
Figure 3: (a) Density of states of H with and without the FCA (b) The absorption spectrum obtained from 400 Lanczos iterations with and without the FCA.

6.4 Resolution and the number of Lanczos steps

In Figure 4, we show that the resolution of the absorption spectrum obtained from the Lanczos algorithm clearly improves as we take more Lanczos iterations for TFBA and HBDMI. In this set of experiments, we use FCA for all test runs.

When $k = 400$ iterations are performed, the Lanczos approximation is nearly indistinguishable from traditional diagonalization. When k is as small as 100, the general features of the absorption spectrum are clearly captured.
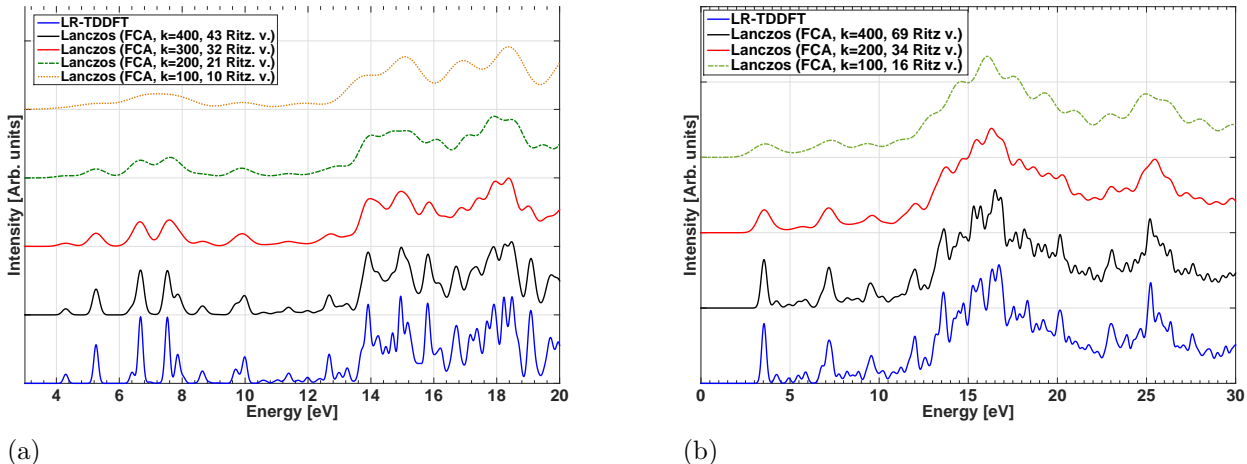

(a)   (b)

Figure 4: Resolution improvement with respect to the number of Lanczos iterations k for (a) TFBA (b) HBDMI.

6.5 Approximation for Larger Molecules

In Figure 5, we compare the absorption spectra obtained from the Lanczos-based LR-TDDFT calculation with those obtained from RT-TDDFT simulations for both the F2N12S and P3B2 molecules. We can see that without the FCA, the Lanczos algorithm can capture the basic features of the F2N12S absorption spectrum exhibited by the RT-TDDFT simulation after $k = 1200$ iterations, but does not clearly reveal all the peaks. However, when the FCA is used, the result matches extremely well with that produced by RT-TDDFT. For the P3B2 molecule, we need to run 1200 Lanczos iterations to obtain the result that matches well

with that produced by RT-TDDFT. When the FCA is used, only 400 Lanczos iterations are needed to achieve essentially full resolution in the computed absorption spectrum.

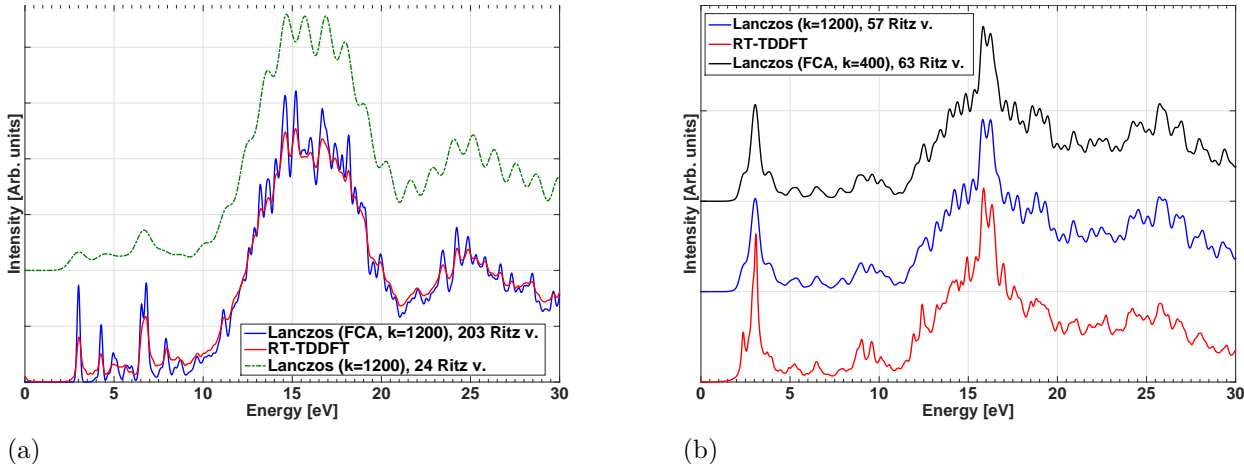(a)   (b)

Figure 5: Comparison of simulated absorption spectra of F2N12S calculated by RT- and Lanczos-LR-TDDFT with and without FCA. Comparison of simulated absorption spectra of P3B2 calculated by RT and Lanczos-LR-TDDFT. The number of Ritz values in the displayed interval is 57 for $k = 1200$ or 63 for $k = 400$ while FCA is used.

6.6 Computational efficiency

Clearly, the most expensive method for estimating the absorption spectrum is the full diagonalization approach, which constructs a full Casida matrix of size $O(n_o^2 n_v^2)$ and performs a diagonalization that requires $O((n_o n_v)^3)$ floating point operations (flops). The Davidson, Lanczos and KPM algorithms are iterative methods that require one multiplication of vector(s) by matrices K and M per iteration. These multiplications are the most expensive operations relative to other linear algebra operations. The efficiency of these methods depends on the number of iterations required to reach convergence, and the number of eigenvalues to be computed in the case of the Davidson algorithm since the number of matrix vector multiplications performed in each Davidson iteration is proportional to the number of eigenvalues to be computed. For the Davidson method, the error of each eigenvector can be estimated by computing the residual norm of each eigenpair. For the Lanczos and KPM algorithm, we do not yet have an efficient estimator to estimate the error of the absorption

23

spectrum directly. In the results presented in this paper, we use a visual inspection of the approximate absorption spectrum to determine whether sufficient resolution is reached after k steps. A more systematic approach to terminate these algorithms is to compare approximations produced in consecutive iterations, and terminate when the difference between these approximations is sufficiently small.

In the case of the P3B2 system, we set the residual norm threshold to 10^{-4}, and compute 200 lowest excited states, which cover the $[0, 4.5]$ eV energy range. In each iteration, up to 400 matrix-vector multiplications is performed, which results in 1926 multiplications in total. In order to obtain the valence-level absorption spectrum with good resolution (roughly 0.5 eV) by the Lanczos algorithm, we need to perform $k = 400$ Lanczos iterations. A similar resolution is reached in KPM when approximately $k = 600$ steps are taken. Because in Lanczos as well as in KPM we need to compute the xx, yy, and zz components of the dynamic polarizability, the total number of Lanczos matrix-vector multiplications is $3k = 1200$ Lanczos and $3k = 1800$ for KPM. We remark that when the absorption spectrum over a wider range of interval is needed, the cost of Davidson method increases with respect to the number of eigenvectors, while the cost of Lanczos and KPM methods stays approximately the same.

To illustrate the computational efficiency of the implementation of Lanczos algorithm, we compare the wall clock time taken by 400 Lanczos steps with that used by the Davidson algorithm to compute the lowest 200 eigenvalues of the matrix KM, as well with that used by the RT-TDDFT to run a 24.2-femtosecond trajectory with a 4.8 attosecond time step (5000 time steps). For these settings the RT-TDDFT gives similar resolution of spectra as these obtained by Lanczos. All calculations were performed with a development version of the NWChem[30] program on the Cascade system, which is equipped with 1440 Xeon E5-2670 8C 2.6 GHz 16-core CPUs plus Xeon Phi "MIC" accelerators, 128 GB memory per compute

node and a Infiniband FDR network, and maintained at the EMSL user facility located at the Pacific Northwest National Laboratory. The Xeon Phi accelerators were not utilized in this work.

Each calculation was performed using 1536 cores. In both the Lanczos and RT-TDDFT methods, the xx, yy, and zz components of the dynamic polarizability can be computed simultaneously. We use 512 cores for each component. The Lanczos calculation required 2.5 hours, while Davidson required 4 hours and RT-TDDFT 15 hours.

6.7   Lanczos vs. KPM and DGL

In Figure 6 (a), we compare the absorption spectrum obtained from the Lanczos algorithm, KPM and DGL for the TFBA dye. We ran 400 Lanczos steps, which produces a relatively high resolution approximation to the absorption spectrum produced by exact diagonalization. To make the computational cost of the KPM and DGL method comparable to that of the Lanczos algorithm, we set the degree of the polynomial approximation to 400 also in this test. We can see from Figure 6 (a) that the absorption spectra produced by the KPM and DGL algorithm match with "exact" solution reasonably well in the [0,10] eV energy range. However, they miss some of the peaks beyond 10 eV. We can see that the approximate absorption spectra produced by the KPM and the DGL algorithm are not strictly non-negative, which is an undesirable feature. Furthermore, KPM tends to produce some artificial oscillations and peaks in the approximate absorption spectrum, which may be misinterpreted as real excited states. These artificial oscillations and peaks are the result of the Gibbs oscillation that are present when a discontinuous function, such as the sum of a number of Dirac-δ distributions is approximated by a high degree polynomial. This observation is consistent with that reported in Ref.[38] Because the absorption spectrum associated with molecules contain well isolated peaks, the KPM, which is based on polynomial approximation in a continuous measure, often exhibit Gibbs oscillation when degree of the polynomial is high. The Gibbs oscillations are clearly visible on Figure 6 between 0–5 eV and also create

pseudopeak around 5.5 eV (Figure 6 (b)). This problem is reduced in the DGL algorithm.

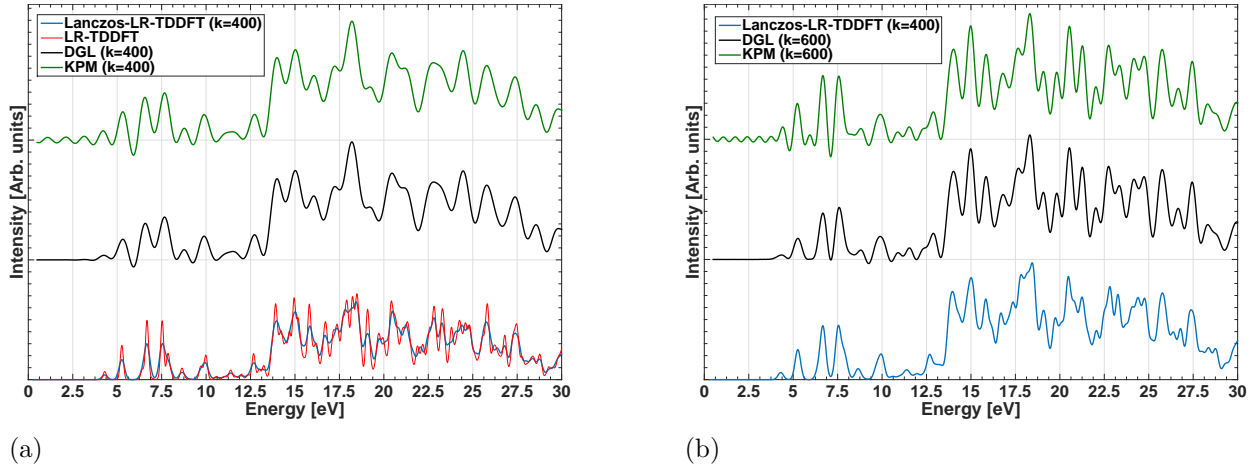(a)   (b)

Figure 6: Comparison of simulated absorption spectra of TFBA calculated by LR-TDDFT, Lanczos-LR-TDDFT, KPM and DGL. The FCA was used. (a) $k = 400$ (b) $k = 400$ for the Lanczos-LR-TDDFT and $k = 600$ for KPM and DGL.

Figure 6 (b) shows that as the degree of the kernel polynomial increases, for example, to 600, all major peaks of the absorption spectrum in $[0, 30]$ eV are correctly resolved.

6.8   DOS by Lanczos

In Figure 7, we plot the DOS approximation for the FTBA dye obtained from the Lanczos algorithm and compare it with that obtained from a full diagonalization of the KM matrix. The FCA is used to show the DOS in the energy range $[0, 125]$ eV. We use $n_{\text{vec}} = 10$ randomly generated vectors to run the Lanczos algorithm. When 200 Lanczos steps are used in each run, the DOS approximation obtained from (29) is nearly indistinguishable from that constructed from the eigenvalues of the KM matrix. The total number of matrix vector multiplications used in this case is $10 \times 200 = 2000$. If fewer Lanczos steps are taken in each run, the resulting DOS approximation becomes less well resolved. However, the general features of the DOS can still be seen clearly when only 50 Lanczos steps are taken in each run.
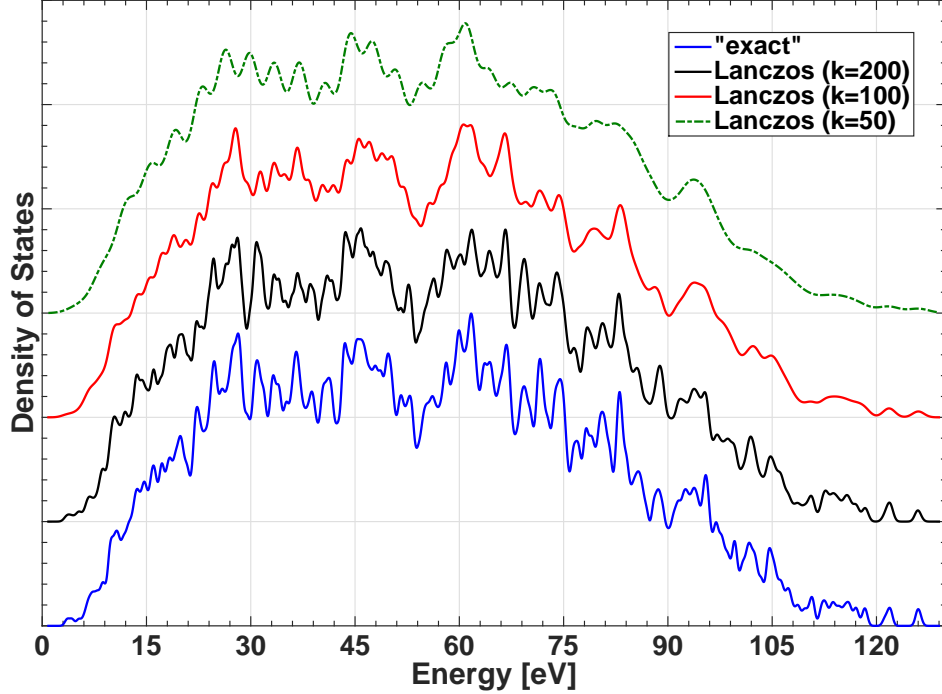
Figure 7: Resolution improvement of the density of states for FTBA, the FCA approximation is used. The number of random initial vectors for Lanczos iteration is 10.

7   Conclusion

We have described two iterative algorithms for approximating the absorption spectrum of finite systems within LR-TDDFT. We used the fact that the product eigenvalue problem is self-adjoint with respect to an appropriately chosen inner product, which allowed us to propose the symmetric Lanczos algorithm and corresponding KPM algorithm as an low-memory option. Our computational examples show that these methods can be much more efficient than traditional methods. In addition, the Lanczos algorithm generally gives more accurate approximation to the absorption spectrum than that produced by the KPM or DGL, when the same number of matrix vector multiplications are used. In particular, the approximate absorption spectrum produced by the Lanczos algorithm is strictly non-negative. This is not necessarily the case for KPM or DGL. Furthermore, KPM tends to produce additional artificial oscillations, that may be misinterpreted as fictitious excited states. However, the Lanczos algorithm requires storing more vectors in order to overcome potential numerical

instability. In contrast, both the KPM and DGL methods can be stably implemented using a three-term recurrence, leading to a minimal storage requirement. While we have demonstrated the efficiency of our approaches to estimate spectra features, more work needs to be done to obtain information about the composition of the excited states and state-specific properties like forces and Hessians where traditional methods based on partial diagonalization of the Casida matrix are still very useful.

Acknowledgement

Supporting Information Available

Geometries and energies of all testing systems can be found in Supplementary information. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

Appendix

In this section, we provide detailed derivations and explanations for some of the expressions presented in the main text.

A   Spectral function

The *spectral function* associated with a function of the form

$$f(\omega) = \sum_i \frac{a_i}{\omega - \lambda_i},$$

is defined to be

$$s(\omega) = \sum_i a_i \delta(\omega - \lambda_i).$$

It is an elegant way to describe the numerator (or weighting factor) associated with each pole of $f(\omega)$.

An alternative way to express $s(\omega)$ is through the expression

$$s(\omega) = -\frac{1}{\pi} \lim_{\eta \to 0^+} \mathrm{Im} f(\omega + i\eta) = \frac{1}{\pi} \lim_{\eta \to 0^+} \sum_i a_i \frac{\eta}{(\omega - \lambda_i)^2 + \eta^2}.$$

This is the expression we used in (1) to define the dynamic polarizability tensor α.

B   Derivation for χ

The expression of $\chi(\omega)$ given in (2) can be derived in a number of ways, (e.g. through the Liouville super-operator presented in Ref.[29]) We give a simple derivation of this expression here using basic linear algebra.

We start from a matrix representation of (2)

$$\chi(\omega) = \varepsilon^{-1}(\omega)\chi_0(\omega) = [I - \chi_0(\omega)f_{Hxc}]^{-1}\chi_0(\omega),$$

and substitute χ_0 with (5) to obtain

$$\chi(\omega) = \left[I - \hat{\Phi}\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T f_{Hxc}\right]^{-1}\hat{\Phi}\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T.$$

Using the Sherman–Morrison–Woodbury formula for the following matrix inverse

$$
\begin{aligned}
\left[I - \hat{\Phi}\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T f_{Hxc}\right]^{-1} &= I + \hat{\Phi}\left[\big((\omega + i\eta)C - D\big) - \hat{\Phi}^T f_{Hxc}\hat{\Phi}\right]^{-1}\hat{\Phi}^T f_{Hxc} \\
&= I + \hat{\Phi}\big((\omega + i\eta)C - \Omega\big)^{-1}\hat{\Phi}^T f_{Hxc},
\end{aligned}
$$

where Ω is defined by (7). We obtain

$$
\begin{aligned}
\chi(\omega) &= \left[I + \hat{\Phi}\big((\omega + i\eta)C - \Omega\big)^{-1}\hat{\Phi}^T f_{Hxc}\right]\hat{\Phi}\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T \\
&= \hat{\Phi}\left[I + \big((\omega + i\eta)C - \Omega\big)^{-1}(\Omega - D)\right]\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T \\
&= \hat{\Phi}\left[I + \big((\omega + i\eta)C - \Omega\big)^{-1}\big(\Omega - (\omega + i\eta)C + (\omega + i\eta)C - D\big)\right]\big((\omega + i\eta)C - D\big)^{-1}\hat{\Phi}^T \\
&= \hat{\Phi}\big((\omega + i\eta)C - \Omega\big)^{-1}\hat{\Phi}^T.
\end{aligned}
$$

$$(30)$$

C   Eigendecomposition of the Casida Matrix

It is known (see, e.g., Ref.[47]) that when both $M \equiv A + B$ and $K \equiv A - B$ are positive definite, the Casida matrix, H, defined in (11) admits an eigendecomposition of the form

$$
\begin{bmatrix} A & B \\ -B & -A \end{bmatrix} = \begin{bmatrix} U & V \\ V & U \end{bmatrix}\begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}\begin{bmatrix} U & V \\ V & U \end{bmatrix}^{-1} = \begin{bmatrix} U & V \\ V & U \end{bmatrix}\begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}\begin{bmatrix} U & -V \\ -V & U \end{bmatrix}^T,
$$

where U, V are real $n_o n_v \times n_o n_v$ matrices satisfying $U^T U - V^T V = I$, $U^T V - V^T U = 0$, and $\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_{n_o n_v}\}$. Let $U = [u_1, \ldots, u_{n_o n_v}]$, $V = [v_1, \ldots, v_{n_o n_v}]$. Then the right and left eigenvectors of H associated with λ_i are $[u_i^T, v_i^T]^T$ and $[u_i^T, -v_i^T]^T$, respectively. In addition, we have

$$
\begin{aligned}
\langle x|\chi(\omega)|x\rangle &= \hat{x}^T \big[(\omega + i\eta)I - H\big]^{-1} C\hat{x}^T \\
&= [\hat{x}_1^T, \hat{x}_1^T] \left((\omega + i\eta)I - \begin{bmatrix} A & B \\ -B & -A \end{bmatrix}\right)^{-1} \begin{bmatrix} \hat{x}_1 \\ -\hat{x}_1 \end{bmatrix} \\
&= [\hat{x}_1^T, \hat{x}_1^T] \begin{bmatrix} U & V \\ V & U \end{bmatrix} \left((\omega + i\eta)I - \begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}\right)^{-1} \begin{bmatrix} U & -V \\ -V & U \end{bmatrix}^T \begin{bmatrix} \hat{x}_1 \\ -\hat{x}_1 \end{bmatrix} \\
&= [\hat{x}_1^T(U+V), \hat{x}_1^T(U+V)] \left((\omega + i\eta)I - \begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}\right)^{-1} \begin{bmatrix} (U+V)^T \hat{x}_1 \\ -(U+V)^T \hat{x}_1 \end{bmatrix} \\
&= \sum_{i=1}^{n_o n_v} \left(\frac{[\hat{x}_1^T(u_i + v_i)]^2}{\omega - \lambda_i + i\eta} - \frac{[\hat{x}_1^T(u_i + v_i)]^2}{\omega + \lambda_i + i\eta} \right).
\end{aligned}
\tag{31}
$$

In the limit of $\eta \to 0^+$, the imaginary part of (31) becomes

$$
\lim_{\eta \to 0^+} \text{Im}\langle x|\chi(\omega)|x\rangle = -\pi \sum_{i=1}^{n_o n_v} \left([\hat{x}_1^T(u_i + v_i)]^2 \delta(\omega - \lambda_i) - [\hat{x}_1^T(u_i + v_i)]^2 \delta(\omega + \lambda_i) \right).
$$

$$
= -2\pi \, \text{sign}(\omega) \sum_{i=1}^{n_o n_v} \lambda_i [\hat{x}_1^T(u_i + v_i)]^2 \delta(\omega^2 - \lambda_i^2).
$$

Here we make use of the identity

$$
\delta(\omega \mp \lambda_i) = 2\lambda_i H(\pm\omega)\delta(\omega^2 - \lambda_i^2),
$$

where $H(x) = \left[1 + \text{sign}(x)\right]/2$ is the Heaviside step function. This identity can be proved by showing

$$\xi(\lambda_i) = \int_0^{+\infty} 2\lambda_i \delta(\omega^2 - \lambda_i^2)\xi(\omega)d\omega = \int_{-\infty}^{+\infty} 2\lambda_i H(\omega)\delta(\omega^2 - \lambda_i^2)\xi(\omega)d\omega,$$

$$\xi(-\lambda_i) = \int_{-\infty}^0 2\lambda_i \delta(\omega^2 - \lambda_i^2)\xi(\omega)d\omega = \int_{-\infty}^{+\infty} 2\lambda_i H(-\omega)\delta(\omega^2 - \lambda_i^2)\xi(\omega)d\omega,$$

for any continuous function $\xi(\omega)$, through a change of variable $\varpi = \omega^2$, $\omega = \pm\sqrt{\varpi}$, and $d\omega = \pm\frac{d\varpi}{2\sqrt{\varpi}}$.

It can be verified that the normalization condition $U^T U - V^T V = I$ implies that $(U + V)^T(U - V) = I$ and $(U + V)^T M (U + V) = \Lambda$. Consequently, we have

$$M = (U - V)\Lambda(U - V)^T, \qquad K = (U + V)\Lambda(U + V)^T.$$

The eigendecompositions of MK and KM are thus given by

$$MK = (U - V)\Lambda^2(U - V)^{-1} = (U - V)\Lambda^2(U + V)^T,$$

$$KM = (U + V)\Lambda^2(U + V)^{-1} = (U + V)\Lambda^2(U - V)^T.$$

Therefore, the vectors $u_i + v_i$ $(i = 1, 2, \ldots, n_o n_v)$ in (31) are the right eigenvector of KM, as well as the left eigenvector of MK, both associated with the eigenvalues λ_i^2. These results lead to

$$\alpha_{x,x}(\omega) = 2\,\text{sign}(\omega)\sum_{i=1}^{n_o n_v} \lambda_i \left[\hat{x}_1^T(u_i + v_i)\right]^2 \delta(\omega^2 - \lambda_i^2)$$

$$= 2\,\text{sign}(\omega)\hat{x}_1^T(U + V)\Lambda\delta(\omega^2 I - \Lambda^2)(U + V)^T\hat{x}_1$$

$$= 2\,\text{sign}(\omega)\hat{x}_1^T(U + V)\Lambda(U + V)^T(U - V)\delta(\omega^2 I - \Lambda^2)(U + V)^T\hat{x}_1$$

$$= 2\,\text{sign}(\omega)\hat{x}_1^T K\delta(\omega^2 I - MK)\hat{x}_1.$$

Thus we have proved (14), which is an expression for $\alpha_{x,x}(\omega)$ that involves matrices of dimension $n_o n_v \times n_o n_v$ (instead of $2n_o n_v \times 2n_o n_v$ in (8)).

Notes and References

(1) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

(2) Onida, G.; Reining, L.; Rubio, A. *Rev. Mod. Phys.* **2002**, *74*, 601–659.

(3) Casida, M. E. In *Recent Advances in Density Functional Methods*; Chong, D. E., Ed.; World Scientific, Singapore, 1995; pp 155–192.

(4) Casida, M. E. In *Recent Developments and Application of Modern Density Functional Theory*; Seminario, J. M., Ed.; Elsevier, Amsterdam, 1996; pp 391–439.

(5) Tretiak, S.; Isborn, C. M.; Niklasson, A. M.; Challacombe, M. *J. Chem. Phys.* **2009**, *130*, 054111.

(6) Joergensen, P. *Second Quantization-Based Methods in Quantum Chemistry*; Elsevier Science, 2012.

(7) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 8218–8224.

(8) Liang, W.; Fischer, S. A.; Frisch, M. J.; Li, X. *J. Chem. Theory Comput.* **2011**, *7*, 3540–3547.

(9) Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.

(10) Kauczor, J.; Jørgensen, P.; Norman, P. *J. Chem. Theory Comput.* **2011**, *7*, 1610–1630.

(11) Ullrich, C. A. *Time-dependent Density Functional Theory*; Oxford University, 2012.

(12) Olsen, J.; Jensen, H. J. A.; Jørgensen, P. *J. Comput. Phys.* **1988**, *74*, 265–282.

(13) Vasiliev, I.; Öğüt, S.; Chelikowsky, J. R. *Phys. Rev. Lett.* **2001**, *86*, 1813.

(14) Wang, Y.; Lopata, K.; Chambers, S. A.; Govind, N.; Sushko, P. V. *J. Phys. Chem. C* **2013**, *117*, 25504–25512.

(15) Tussupbayev, S.; Govind, N.; Lopata, K.; Cramer, C. J. *J. Chem. Theor. Comput.* **2015**, *11*, 1102–1109.

(16) Kauczor, J.; Norman, P. *J. Chem. Theory Comput.* **2014**, *10*, 2449–2455.

(17) Jensen, L.; Autschbach, J.; Schatz, G. C. *J. Chem. Phys.* **2005**, *122*, 224115.

(18) Devarajan, A.; Gaenko, A.; Autschbach, J. *J. Chem. Phys.* **2009**, *130*, 194102.

(19) Hübener, H.; Giustino, F. *J. Chem. Phys.* **2014**, *141*, 044117.

(20) Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Phys.* **2012**, *136*, 124107.

(21) Grimme, S. *J. Chem. Phys.* **2013**, *138*, 244104.

(22) Bannwarth, C.; Grimme, S. *Comp. Theor. Chem.* **2014**, *1040*, 45–53.

(23) Neuhauser, D.; Baer, R. *J. Chem. Phys.* **2005**, *123*, 204105:1–204105:6.

(24) Li, X.; Smith, S. M.; Markevitch, A. N.; Romanov, D. A.; Levis, R. J.; Schlegel, H. B. *Phys. Chem. Chem. Phys.* **2005**, *7*, 233–239.

(25) Lopata, K.; Govind, N. *J. Chem. Theory Comput.* **2011**, *7*, 1344–1355.

(26) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

(27) Silver, R. N.; Röder, H.; Voter, A. F.; Kress, J. D. *J. Comput. Phys.* **1996**, *124*, 115–130.

(28) Gordienko, A. B.; Filippov, S. I. *Phys. Status Solidi B* **2014**, *251*, 628–632.

(29) Rocca, D.; Gebauer, R.; Saad, Y.; Baroni, S. *J. Chem. Phys.* **2008**, *128*, 154105.

(30) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.

(31) Bai, Z.; Li, R.-C. *SIAM J. Matrix Anal. Appl.* **2012**, *33*, 1075–1100.

(32) Benner, P.; Mehrmann, V.; Xu, H. *Numer. Math.* **1998**, *78*, 329–357.

(33) Bai, Z.; Li, R.-C. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 392–416.

(34) Challacombe, M. *Computation* **2014**, *2*, 1–11.

(35) Paige, C. C. The computation of eigenvalues and eigenvectors of very large sparse matrices. Ph.D. thesis, University of London, London, England, 1971.

(36) Moro, G.; Freed, J. H. *J. Chem. Phys.* **1975**, *75*, 3157.

(37) Haydock, R.; Heine, V.; Kelly, M. J. *J. Phys. C: Solid State Phys.* **1972**, *5*, 2845.

(38) Lin, L.; Saad, Y.; Yang, C. *to appear in SIAM Rev.* **2015**, arXiv:1308.5467.

(39) Jackson, D. *The theory of approximation*; American Mathematical Society, New York, 1930.

(40) Cohen, M.; Kelly, P. S. *Can. J. Phys.* **1966**, *44*, 3227–3240.

(41) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(42) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(43) Hariharan, P.; Pople, J. *Theor. Chim. Acta.* **1973**, *28*, 213–222.

(44) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(45) See Supplementary Material Document No. for geometries.

(46) Shao, M.; Jornada, F. H.; Yang, C.; Deslippe, J.; Louie, S. G. *submitted to Linear Algebra Appl.* **2015**, arXiv:1501.03830.
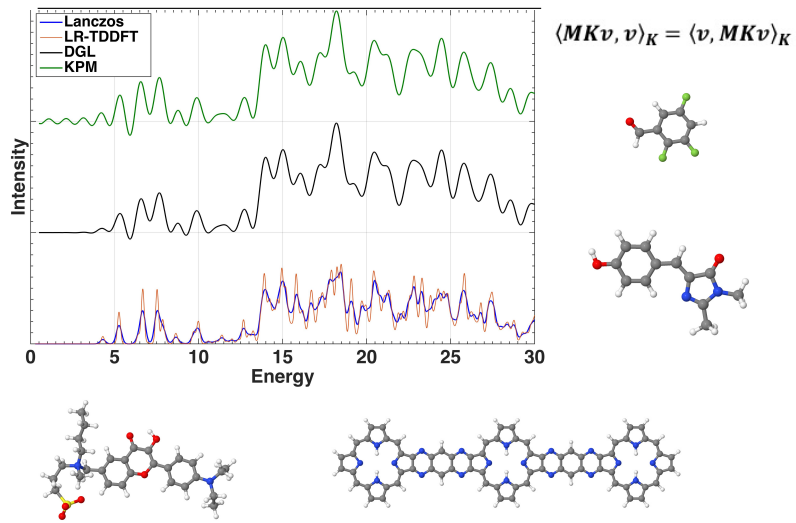
(47) Ref. 46, Theorem 3.

For Table of Contents use only

Figure 8: Table of Contents