A block variational procedure for the iterative diagonalization of non-Hermitian random-phase approximation matrices

Dario Rocca, Zhaojun Bai, Ren-Cang Li, and Giulia Galli

Citation: J. Chem. Phys. **136**, 034111 (2012); doi: 10.1063/1.3677667
View online: http://dx.doi.org/10.1063/1.3677667
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v136/i3
Published by the American Institute of Physics.

A block variational procedure for the iterative diagonalization of non-Hermitian random-phase approximation matrices

Dario Rocca,[1,a)] Zhaojun Bai,[2,3,b)] Ren-Cang Li,[4] and Giulia Galli[1,5]
[1]*Department of Chemistry, University of California, Davis, California 95616, USA*
[2]*Department of Computer Science, University of California, Davis, California 95616, USA*
[3]*Department of Mathematics, University of California, Davis, California 95616, USA*
[4]*Department of Mathematics, University of Texas, Arlington, Texas 76019, USA*
[5]*Department of Physics, University of California, Davis, California 95616, USA*

(Received 10 August 2011; accepted 28 December 2011; published online 18 January 2012)

We present a technique for the iterative diagonalization of random-phase approximation (RPA) matrices, which are encountered in the framework of time-dependent density-functional theory (TDDFT) and the Bethe-Salpeter equation. The non-Hermitian character of these matrices does not permit a straightforward application of standard iterative techniques used, i.e., for the diagonalization of ground state Hamiltonians. We first introduce a new block variational principle for RPA matrices. We then develop an algorithm for the simultaneous calculation of multiple eigenvalues and eigenvectors, with convergence and stability properties similar to techniques used to iteratively diagonalize Hermitian matrices. The algorithm is validated for simple systems (Na_2 and Na_4) and then used to compute multiple low-lying TDDFT excitation energies of the benzene molecule. © *2012 American Institute of Physics*. [doi:10.1063/1.3677667]

I. INTRODUCTION

Time-dependent density-functional theory (TDDFT) and the Bethe-Salpeter equation (BSE) are widely used to compute the optical excitations of a wide range of systems.[1] In the adiabatic approximation using local or hybrid exchange-correlation functionals, TDDFT is considered an accurate approach for molecular systems.[2] The BSE has been mostly used for extended periodic systems[3–5] but its utilization is becoming increasingly popular also for the calculation of molecular spectra.[6–8] In the widely used particle-hole formulation[5,9] or in the density matrix perturbation theory formulation,[8,10,11] the calculation of TDDFT, and BSE excitation energies and polarizabilities are formulated in terms of a non-Hermitian eigenproblem with the structure of a random-phase approximation (RPA) matrix:[12]

$$\mathcal{H}_{\mathrm{RPA}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} A & B \\ -B^* & -A^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \omega \begin{pmatrix} x \\ y \end{pmatrix}, \tag{1}$$

where the block matrices A and B are Hermitian and positive definite. In this work we will assume that the matrices A and B are real, which is always the case for molecular systems with time reversal symmetry. Although in electronic structure theory RPA usually defines a specific approximation used for constructing $\mathcal{H}_{\mathrm{RPA}}$, in the following we will indicate a generic eigenvalue problem with the structure of Eq. (1) as RPA eigenvalue problem. The theory presented in this work is general and does not depend on the specific approximation used to build $\mathcal{H}_{\mathrm{RPA}}$, such as TDDFT or BSE.

The non-Hermitian character of $\mathcal{H}_{\mathrm{RPA}}$ does not allow for the application of standard iterative techniques used, i.e., in ground state DFT calculations.[13] For this reason, the Tamm-Dancoff approximation,[14,15] that approximates the RPA operator in terms of a Hermitian matrix by discarding the coupling blocks B and $-B^*$, has been widely used in the solution of both TDDFT and the BSE;[1] while this approximation is considered accurate for bulk systems, its validity is still controversial in the case of molecules.[7,8]

In recent implementations the non-Hermitian RPA problem of TDDFT and BSE has been solved by directly computing the electronic polarizability using a Lanczos algorithm without direct diagonalization of the RPA operator.[7,8,10,11] This approach is computationally efficient and allows one to compute spectra in a wide energy range. However, it does not allow for a direct assignment of excitation energies in terms of transitions between single particle states.

For the case of TDDFT in the electron-hole formulation,[9] if one uses a local exchange correlation functional it is possible to reformulate the RPA problem in terms of a pseudo-Hermitian operator, and standard iterative techniques are then applicable.[13,16] However, this simplification cannot be easily applied to the case of the BSE, TDDFT with hybrid functionals or the density matrix perturbation theory formulation.[8,10,11]

For the case of RPA matrices of the form of $\mathcal{H}_{\mathrm{RPA}}$ in Eq. (1), a minimization principle for the lowest lying eigenvalue has been introduced by Thouless in 1961.[12] This variational principle and its variants have been exploited by the algorithms introduced in Refs. 17–21. However, Thouless' variational principle is valid only for the lowest lying eigenvalue. For this reason only one eigenvalue and eigenvector at a time can be calculated and used to build a constraint for the next eigenvalue calculation. In contrast, efficient Hermitian iterative methods are based on a block form of the minimization principle of Hermitian matrices.

a)Electronic mail: drocca@ucdavis.edu.
b)Electronic mail: bai@cs.ucdavis.edu.

In this work we introduce a new block minimization principle for RPA non-Hermitian matrices $\mathcal{H}_{\rm RPA}$ and then derive a block steepest descent (SD) algorithm to diagonalize matrices $\mathcal{H}_{\rm RPA}$. We implemented this algorithm in the framework of TDDFT with local exchange correlation, within the approach of Refs. 10, 11, and 22. After validating the method on simplified examples (small sodium nanoclusters), we present an application to the low-lying excitation spectrum of the benzene molecule. In this latter case, the dimension of the non-Hermitian matrix associated with the TDDFT problem is up to about 5.6×10^6. The application of this method to the solution of the BSE and the study of preconditioning schemes will be subject of future work.

The rest of this paper is organized as follows. In Sec. II we introduce the equations of the non-Hermitian RPA eigenvalue problem for the case of TDDFT in the local and semi-local adiabatic approximations. In Sec. III we briefly review the Thouless minimization principle for RPA matrices and then introduce the new block form of this principle. In Sec. IV we use the block minimization principle to develop conjugate gradient-like (CG-like) algorithms that allow for the iterative calculation of multiple eigenvalues simultaneously. Applications to the diagonalization of exactly solvable RPA matrices and to the low-lying TDDFT spectrum of benzene are presented in Secs. V and VI, respectively. Section VII contains our conclusions.

II. THEORETICAL BACKGROUND

A detailed review of the theoretical background of the TDDFT and the BSE approaches can be found in Ref. 1. In the frequency domain formulation and in the adiabatic approximation, the TDDFT and BSE excitations energies are obtained by solving a RPA eigenproblem, as defined in Eq. (1). In the usual approximations used in TDDFT and BSE calculations, the matrices A and B in Eq. (1) are Hermitian and positive definite and thus the eigenvalues of $\mathcal{H}_{\rm RPA}$ are real. Furthermore, it is easy to demonstrate that if ω is an eigenvalue of $\mathcal{H}_{\rm RPA}$ corresponding to the right eigenvector $(x^T, y^T)^T$, then $-\omega$ is an eigenvalue corresponding to the eigenvector $(y^T, x^T)^T$. If in Eq. (1) the coupling matrices B and $-B^*$ are set to zero, the so-called Tamm-Dancoff approximation is obtained.[14,15]

In the following we consider the explicit TDDFT equations for the case of molecular systems with time-reversal symmetry (no magnetic fields are applied to the system). In this case the molecular orbitals can be chosen to be real and, as a consequence, the matrices A and B are real and symmetric. We focus on the TDDFT formalism in the (semi-)local adiabatic approximation[2] and we present examples at this level of theory. A detailed presentation of the BSE or the hybrid functional formalism can be found, e.g., in Refs. 2, 5, and 8. Since the BSE may be cast into the form of Eq. (1), in principle the following discussion and methodology are applicable also to the solution of this equation and this will be subject of future work.

Most practical approaches to solve Eq. (1) for TDDFT and BSE make use of an electron-hole (e-h) basis set;[5,9] in the case of TDDFT this leads to the Casida's equations. Within this approach, in the spin-restricted case, Eq. (1) is

$$A_{vc,v'c'} = (\varepsilon_c^\circ - \varepsilon_v^\circ)\delta_{vv'}\delta_{cc'} + K_{vc,v'c'}, \tag{2}$$

$$B_{vc,v'c'} = K_{vc,c'v'}, \tag{3}$$

where the kernel

$$K_{vc,v'c'} = 2\int \phi_c^\circ(\mathbf{r})\phi_v^\circ(\mathbf{r})\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta V_{xc}(\mathbf{r})}{\delta n(\mathbf{r}')}\delta(\mathbf{r}-\mathbf{r}')\right) \times \phi_{v'}^\circ(\mathbf{r}')\phi_{c'}^\circ(\mathbf{r}')d\mathbf{r}d\mathbf{r}', \tag{4}$$

describes the local field effects and the exchange-correlation (xc) effects; in Eq. (4), $\delta V_{xc}/\delta n$ is the functional derivative of the exchange and correlation potential V_{xc} with respect to the density n. The indexes v and v' label occupied energy levels ε and orbitals ϕ, while c and c' denote empty states. The dimension of the RPA matrix in the electron-hole representation is $2 \times N_v \times N_c$, where N_v is the number of valence and N_c is the number of conduction states. Atomic units have been used in the previous equations and will be used throughout the paper. Since in this formulation the $(A - B)$ matrix is diagonal, the RPA eigenvalue problem in Eq. (1) can be easily reduced to an Hermitian form

$$(A-B)^{1/2}(A+B)(A-B)^{1/2}|z\rangle = \omega^2|z\rangle, \tag{5}$$

where we have defined $z = (A-B)^{-1/2}|x+y\rangle$. Unfortunately, such simplification is convenient only when the electron-hole basis set is used with a local approximation for V_{xc} in the kernel. In the case of BSE, hybrid functional TDDFT or density matrix perturbation theory, the $(A - B)$ matrix is not diagonal and the evaluation of $(A - B)^{1/2}$ is computationally prohibitive. In order to solve Eq. (5) also in the case of hybrid functionals, in Ref. 16 a method was proposed to systematically approximate $(A - B)^{1/2}$ in a small dimensional iterative subspace.

In this work we do not rely on the Hermitian form of Eq. (5). However, the theory and algorithms proposed here for the RPA eigenvalue problem have close similarities to those for Hermitian matrices.[13]

The electron-hole formulation of TDDFT and BSE requires the explicit calculation of the empty electronic states ϕ_c°. In principle, all the empty states should be included but in most practical implementations only a limited number of them are used, and the convergence with respect to their number has to be carefully tested. The inclusion of a large number of conduction states is particularly important when a large portion of the spectrum is needed or when there is a strong coupling between low and high energy levels.[8,23]

Recently, a method has been introduced to solve the equations of TDDFT and the BSE within density matrix perturbation theory.[8,10,11] This approach enables the inclusion of the full conduction subspace without the explicit diagonalization of the ground state Hamiltonian. This is achieved by generalizing concepts of density functional perturbation theory to the case of time dependent linear response.[24,25] Within this formalism the TDDFT equations can be cast into the form of the RPA matrix Eq. (1), with the following definition of the

operators A and B:

$$A_{v,v'}\,|a_{v'}\rangle = (\hat{H}^{\circ} - \varepsilon_{v'})\,\delta_{vv'}\,|a_{v'}\rangle + 2\hat{Q}\left(\int K^{HXC}(\mathbf{r},\mathbf{r}')\phi^{\circ}_{v'}(\mathbf{r}')a_{v'}(\mathbf{r}')d\mathbf{r}'\right)|\phi^{\circ}_{v}\rangle, \tag{6}$$

$$B_{v,v'}\,|b_{v'}\rangle = 2\hat{Q}\left(\int K^{HXC}(\mathbf{r},\mathbf{r}')b_{v'}(\mathbf{r}')\phi^{\circ}_{v'}(\mathbf{r}')d\mathbf{r}'\right)|\phi^{\circ}_{v}\rangle, \tag{7}$$

where

$$K^{HXC}(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta V_{xc}(\mathbf{r})}{\delta n(\mathbf{r}')}\delta(\mathbf{r}-\mathbf{r}'), \tag{8}$$

and $\hat{H}^{\circ}$ is the ground state Hamiltonian; $a_{v'}(\mathbf{r})$ and $b_{v'}(\mathbf{r})$ denote two generic sets of N_v orbitals orthogonal to the occupied ground state orbitals ϕ°_{v}. The projector onto the conduction state subspace $\hat{Q}$ can be computed as $\hat{I} - \hat{P}$, where $\hat{I}$ is the identity operator and $\hat{P}$ is the projector onto the occupied state subspace. It is important to note that the total number of orbitals involved (ϕ°_{v}, a_v, b_v) in Eqs. (6) and (7) is equal to N_v, and the operations required to solve this equation are similar to those required by a ground state calculation: for example, one needs to evaluate the application of the ground state Hamiltonian $\hat{H}^{\circ}$ to N_v orbitals and to compute a Hartree-exchange-correlation term in the kernel. Furthermore, by using well established techniques for ground state calculations, the operators A and B in Eqs. (6) and (7) can be applied to the set of orbitals $a_{v'}$ and $b_{v'}$ without explicitly building the corresponding matrices.[26]

In this formulation the dimension of the explicit RPA matrix $\mathcal{H}_{\mathrm{RPA}}$ is $2 \times N_v \times N_{basis}$, where N_{basis} is the dimension of the basis set used to expand the orbitals. Since in general $N_c \gg N_v$, we have $N_{basis} = N_c + N_v \approx N_c$, namely, the dimension of the matrix in Eq. (1) is approximately the same both in the e-h hole formalism including all the empty states, and in the density matrix perturbation theory formalism. As already mentioned, by using Eqs. (6) and (7), it is not necessary to build the full matrix Eq. (1), and all the empty states are automatically included without diagonalizing the ground state Hamiltonian. This formalism can be extended to the case of the BSE and TDDFT using hybrid functionals.[8]

In Secs. III and IV a new algorithm to iteratively diagonalize the RPA eigenvalue problem is presented.

III. MINIMIZATION PRINCIPLES FOR THE RPA EIGENVALUE PROBLEM

In this section we discuss minimization principles for the RPA eigenvalue problem Eq. (1). We consider the case in which the $m \times m$ A and B matrices in Eq. (1) are real and symmetric positive definite. The spectrum of the matrix in Eq. (1) is characterized by $2m$ real eigenvalues, symmetric with respect to 0, i.e., $\pm\omega_i$ for $i = 1, 2, \ldots, m$. We are interested in the k smallest positive eigenvalues $0 < \omega_1 \le \omega_2 \le \ldots \le \omega_k$, where the number k depends on the specific problem of interest but in general is limited to a relatively small number compared to m.

In 1961, a variational principle was introduced by Thouless to determine the lowest eigenvalue of the RPA eigenvalue problem.[12] Thouless showed that the lowest eigenvalue ω_1 of $\mathcal{H}_{\mathrm{RPA}}$ can be obtained by minimizing the functional

$$\varrho(x,y) = \frac{\begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}}{|x^T x - y^T y|},$$

among all vectors x, y such that $x^T x - y^T y \neq 0$, namely,

$$\omega_1 = \min \varrho(x,y), \tag{9}$$

where the superscript $(\cdot)^T$ transposes a matrix or vector. Introducing the symmetric orthogonal matrix

$$J = \frac{1}{\sqrt{2}} \begin{pmatrix} I_m & I_m \\ I_m & -I_m \end{pmatrix},$$

where I_m is a $m \times m$ identity matrix, we can convert the eigenvalue problem (1) into

$$\mathcal{H}'_{\mathrm{RPA}} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & K \\ M & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \omega \begin{pmatrix} p \\ q \end{pmatrix}, \tag{10}$$

where $K = A - B$ and $M = A + B$, $p = 1/\sqrt{2}(x+y)$ and $q = 1/\sqrt{2}(x-y)$. From the definition of A and B in Eqs. (2) and (3) and Eqs. (6) and (7), we have that both K and M are symmetric positive definite, and the two eigenvalue problems (1) and (10) are equivalent.

As shown by Tsiper,[17] the equivalent of the Thouless minimization principle for the matrix $\mathcal{H}'_{\mathrm{RPA}}$ in Eq. (10) is

$$\omega_1 = \min \rho(p,q), \tag{11}$$

where

$$\rho(p,q) = \frac{1}{2} \cdot \frac{q^T K q + p^T M p}{|q^T p|}.$$

The minimization principles, Eqs. (9) and (11), have been exploited to compute the smallest (positive) eigenpair by using the nonlinear conjugate gradient (CG) method.[18–21] On the other hand, a Lanczos-like algorithm has been developed based on the variational form Eq. (11).[17] In the CG-like approach, only one eigenvalue at a time was computed. In order to compute higher energy eigenvalues, in Refs. 19 and 20 the so-called Wilkinson shift (deflation) was used, while in Ref. 17 the Lanczos vectors were kept orthogonal to the already converged eigenvectors. It is well known that such explicit deflation procedures to compute multiple eigenvalues are numerically unstable and computationally inefficient. For the Lanczos-like method, severe limitations were experienced for large scale RPA eigenvalue problems due to the orthogonality constraints.[20]

In order to develop efficient numerical methods for the simultaneous calculation of multiple low lying eigenvalues of $\mathcal{H}'_{\mathrm{RPA}}$ (and equivalently $\mathcal{H}_{\mathrm{RPA}}$), the Thouless-Tspier minimization principles, Eqs. (9) and (11), have been recently generalized to a block form to include the first few lowest lying excitations.[27] This theory generalizes the well-known trace-minimization principle for the Hermitian eigenvalue problem,

which is the theoretical foundation of block conjugate gradient and Lanczos type methods. Specifically, for the RPA eigenvalue problem one has

$$\sum_{i=1}^{k} \omega_i = \frac{1}{2} \min_{U^T V = I_k} \mathrm{Tr}(V^T K V + U^T M U), \tag{12}$$

where ω_i indicates the ith positive eigenvalue of $\mathcal{H}'_{\mathrm{RPA}}$, U and V are $m \times k$ matrices and Tr is the trace operation. Furthermore, for U_* and V_* that attain the minimum, $(U_*^T, V_*^T)^T$ is a basis matrix of an invariant subspace (eigenvector subspace) of $\mathcal{H}'_{\mathrm{RPA}}$ corresponding to the eigenvalues $\omega_1, \omega_2, \ldots, \omega_k$.

As a consequence of this newly established minimization principle in Eq. (12), it is natural to seek best approximations to $\omega_1, \omega_2, \ldots, \omega_k$ of $\mathcal{H}'_{\mathrm{RPA}}$ by solving the optimization problem on the right hand side of Eq. (12). In Sec. IV, we will develop a block steepest-descent algorithm to iteratively construct the pair of matrices $\hat{U}$ and $\hat{V}$ such that $(\hat{U}^T, \hat{V}^T)^T$ spans an approximate eigenvector subspaces corresponding to the smallest positive eigenvalues of $\mathcal{H}'_{\mathrm{RPA}}$.

For now, let us assume we already have computed such a pair of matrices $\widehat{U}$ and $\widehat{V}$. To find the approximations $\widehat{\omega}_1, \widehat{\omega}_2, \ldots, \widehat{\omega}_k$ of the k smallest positive eigenvalues of $\mathcal{H}'_{\mathrm{RPA}}$, let us define a *structure-preserving projection* of $\mathcal{H}'_{\mathrm{RPA}}$:

$$H_{\mathrm{SR}} = \begin{pmatrix} V^T & 0 \\ 0 & U^T \end{pmatrix} \mathcal{H}'_{\mathrm{RPA}} \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} 0 & V^T K V \\ U^T M U & 0 \end{pmatrix}, \tag{13}$$

where $U = \widehat{U} W_1^{-1}$ and $V = \widehat{V} W_2^{-1}$, and $W = \widehat{U}^T \widehat{V}$ is assumed to be nonsingular and factorized as $W = W_1^T W_2$ (Ref. 33). Then by solving the reduced RPA eigenvalue problem

$$H_{\mathrm{SR}} \begin{pmatrix} \widehat{p}_j \\ \widehat{q}_j \end{pmatrix} = \widehat{\omega}_j \begin{pmatrix} \widehat{p}_j \\ \widehat{q}_j \end{pmatrix}, \tag{14}$$

for the k smallest positive eigenpairs $\left\{\widehat{\omega}_j, \left(\begin{smallmatrix} \widehat{p}_j \\ \widehat{q}_j \end{smallmatrix}\right)\right\}$ for $j = 1, 2, \ldots, k$, we obtain the approximate eigenpairs for the k smallest positive eigenvalues $\omega_1, \omega_2, \ldots, \omega_k$ of $\mathcal{H}'_{\mathrm{RPA}}$ and the corresponding approximate eigenvectors are given by

$$\begin{pmatrix} \widetilde{p}_j \\ \widetilde{q}_j \end{pmatrix} = \begin{pmatrix} U \widehat{p}_j \\ V \widehat{q}_j \end{pmatrix} = \begin{pmatrix} \widehat{U} W_1^{-1} \widehat{p}_j \\ \widehat{V} W_2^{-1} \widehat{q}_j \end{pmatrix}. \tag{15}$$

The use of the minimization principle of Eq. (12) provides a quantitative justification of the fact that $\widehat{\omega}_1, \widehat{\omega}_2, \ldots, \widehat{\omega}_k$ are the best approximation to the k smallest positive eigenvalues $\omega_1, \omega_2, \ldots, \omega_k$ of $\mathcal{H}'_{\mathrm{RPA}}$ in the subspace spanned by the columns of $\widehat{U}$ and $\widehat{V}$.[27]

IV. CONJUGATE GRADIENT-LIKE ALGORITHMS

In this section, we construct the pair of matrices $\hat{U}$ and $\hat{V}$ in such a way that $(\hat{U}^T, \hat{V}^T)^T$ spans an approximate eigenvector subspace corresponding to the smallest positive eigenvalues of $\mathcal{H}'_{\mathrm{RPA}}$. Then by combining with the structure-preserving subspace projection approximation in Eqs. (13)–(15), we are able to develop conjugate gradient-like algorithms that can compute several low-lying positive eigenvalues of $\mathcal{H}'_{\mathrm{RPA}}$ (and therefore $\mathcal{H}_{\mathrm{RPA}}$) simultaneously.

Let us start with computing the smallest positive eigenvalue ω_1. Let $(p^T, q^T)^T$ be the current approximation of the eigenvector associated with ω_1 and $\rho(p, q)$ be the corresponding Thouless functional. In a conjugate gradient-like method, one usually performs a *line search* along the gradient of the objective function $\rho(p, q)$, namely, one looks for the best possible scalar argument t along the line

$$\left\{ \begin{pmatrix} p \\ q \end{pmatrix} + t \begin{pmatrix} r_p \\ r_q \end{pmatrix} : t \in \mathbb{R} \right\} \tag{16}$$

to minimize the Thouless functional $\rho(p, q)$, where r_p and r_q are properly chosen search directions. Such a line search approach has been developed in Refs. 18 and 19. To improve convergence rates, in Ref. 21, a quasi-independent Rayleigh quotient iteration (QUIRQI) scheme has recently been introduced to generalize the line search by solving the minimization problem

$$\min_{s,t} \rho(p + s\, r_p, q + t\, r_q), \tag{17}$$

where the search directions r_p and r_q are selected as $r_p = \nabla_p \rho(p, q)$ and $r_q = \nabla_q \rho(p, q)$, the partial gradients of ρ with respect to p and q:

$$\nabla_p \rho(p, q) = \frac{1}{q^T p} [M p - \rho(p, q)\, q],$$

$$\nabla_q \rho(p, q) = \frac{1}{q^T p} [K q - \rho(p, q)\, p].$$

The QUIRQI scheme is a dual channel optimization scheme with channels coupled only weakly through the line search procedure. The minimization problem in Eq. (17) is solved iteratively by freezing either s or t and minimizing the functional ρ with respect to the other variable, in an alternative manner until convergence is reached. With initial s and t chosen sufficiently near the optimal parameters, convergence should be attained.

To develop an efficient computational technique which is able to compute simultaneously a set of low-lying positive eigenvalues and eigenvectors of $\mathcal{H}'_{\mathrm{RPA}}$, we propose to look for four arguments α, β, s, t that minimize

$$\min_{\alpha,\beta,s,t} \rho(\alpha p + s\, r_p, \beta q + t\, r_q). \tag{18}$$

We call this technique a *4-D search* since it involves a 4-dimensional subspace search. Now let us show that the optimization problem Eq. (18) can be easily solved by solving a reduced 4×4 RPA eigenvalue problem for its smallest positive eigenvalue. Specifically, let $\widehat{U} = (p, r_p)$, $\widehat{V} = (q, r_q)$, $u = \begin{pmatrix} \alpha \\ s \end{pmatrix}$, and $v = \begin{pmatrix} \beta \\ t \end{pmatrix}$, then the objective function in Eq. (18) can be written as

$$\begin{aligned} &\rho(\alpha p + s\, r_p, \beta q + t\, r_q) \\ &\quad = \rho(\widehat{U} u, \widehat{V} v) \\ &\quad = \frac{v^T \widehat{V}^T K \widehat{V} v + u^T \widehat{U}^T M \widehat{U} u}{2|u^T \widehat{U}^T \widehat{V} v|} \\ &\quad = \frac{\widehat{y}^T W_2^{-T} \widehat{V}^T K \widehat{V} W_2^{-1} \widehat{y} + \widehat{x}^T W_1^{-T} \widehat{U}^T M \widehat{U} W_1^{-1} \widehat{x}}{2|\widehat{x}^T \widehat{y}|} \\ &\quad = \frac{\widehat{y}^T V^T K V \widehat{y} + \widehat{x}^T U^T M U \widehat{x}}{2|\widehat{x}^T \widehat{y}|}, \end{aligned} \tag{19}$$

where $\widehat{x} = W_1 u$ and $\widehat{y} = W_2 v$, $\widehat{U}^T \widehat{V} = W_1^T W_2$ is assumed to be nonsingular and both W_i are of dimension 2×2. Then the minimization problem (18) becomes

$$\min_{\alpha,\beta,s,t} \rho(\alpha p + s\, r_p, \beta q + t\, r_q) = \min_{\widehat{x},\widehat{y}} \frac{\widehat{y}^T V^T K V \widehat{y} + \widehat{x}^T U^T M U \widehat{x}}{2|\widehat{x}^T \widehat{y}|} = \widehat{\omega}_1(H_{\mathrm{SR}}), \quad (20)$$

where the last equality derives from the minimization principle Eq. (11), and H_{SR} is a 4×4 structure-preserving projection of the RPA matrix $\mathcal{H}'_{\mathrm{RPA}}$ on the search subspace $U \oplus V$ defined by Eq. (13). Subsequently, if $\begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix}$ is the corresponding eigenvectors of $\widehat{\omega}_1$ of H_{SR}, namely,

$$H_{\mathrm{SR}} \begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix} = \widehat{\omega}_1 \begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix},$$

then the corresponding approximate eigenvector of the original RPA matrix $\mathcal{H}'_{\mathrm{RPA}}$ is given by

$$\begin{pmatrix} \widetilde{p}_1 \\ \widetilde{q}_1 \end{pmatrix} = \begin{pmatrix} U\widehat{x} \\ V\widehat{y} \end{pmatrix}.$$

This gives rise to the *4-D SD algorithm*. Since it involves a larger search space, such an algorithm is obviously faster than the search schemes based on the minimization problems Eqs. (16) and (17).

By applying the trace minimization principle Eq. (12) and the structure-preserving subspace projection approximation discussed in Eqs. (13)–(15), we can immediately extend this 4-D search scheme to the block case and derive the following algorithm to compute simultaneously a set of k smallest positive eigenvalues and corresponding eigenvectors of $\mathcal{H}'_{\mathrm{RPA}}$.

Block 4-D SD algorithm

1 Select initial approximations $P^0 = [p_1, \ldots, p_k]$ and $Q^0 = [q_1, \ldots, q_k]$
2 for $\ell = 0, 1, \ldots$ until convergence:
3 if $\ell = 0, f_j = \rho(p_j, q_j)$ else $f_j = \lambda_j$ for $1 \le j \le k$;
4 $R_K = KQ^\ell - P^\ell \,\mathrm{diag}(f)$; $R_M = MP^\ell - Q^\ell \,\mathrm{diag}(f)$;
5 convergence test;
6 compute the factorization $W = \widehat{U}^T \widehat{V} = W_1^T W_2$, where $\widehat{U} = \begin{pmatrix} P^\ell & R_K \end{pmatrix}$ and $\widehat{V} = \begin{pmatrix} Q^\ell & R_M \end{pmatrix}$;
7 compute the k smallest positive eigenvalues $\Lambda^\ell = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ and the associated eigenvectors $[\widehat{p}_j^T, \widehat{q}_j^T]^T$ of H_{SR} defined in Eq. (13);
8 $P^{\ell+1} = \widehat{U} W_1^{-1} [\widehat{p}_1, \cdots, \widehat{p}_k]$;
$Q^{\ell+1} = \widehat{V} W_2^{-1} [\widehat{q}_1, \cdots, \widehat{q}_k]$;
9 normalize $p_j := p_j/\alpha_j$; $q_j := q_j/\alpha_j$; $\alpha_j = \|[p_j^T, q_j^T]^T\|$ for $1 \le j \le k$;
10 end
11 return $\{\Lambda^\ell, P^\ell, Q^\ell\}$

A few remarks are in order:

Line 4: R_K and R_M are residual vectors of the current approximate eigenvalues and eigenvectors. It is easy to see that each residual vector is proportional to the corresponding partial gradient of the objective function ρ.

Line 5: The convergence is tested by the condition on the normalized residual:

$$\frac{\|\mathcal{H}'_{\mathrm{RPA}} z_j^\ell - \lambda_j z_j^\ell\|}{\|r_j^{(0)}\|} \le tol, \quad \text{where} \quad z_j^\ell = \begin{pmatrix} p_j^\ell \\ q_j^\ell \end{pmatrix}, \quad (21)$$

p_j^ℓ and q_j^ℓ are the jth columns of P^ℓ and Q^ℓ, respectively. $r_j^0 = \mathcal{H}'_{\mathrm{RPA}} z_j^0 - \lambda_j^0 z_j^0$ is the initial residual, and tol is a desired reduction of residual norms, say $tol = 10^{-4}$. The converged residuals (corresponding to the columns of R_K and R_M) are not included in the next steps, Lines 6 and 7. Hence, the number of columns in R_K and R_M is i, with $i \le k$.

Line 6: A simple choice of the decomposition is $W_1 = W^T$ and $W_2 = I$. For a robust implementation, one should also consider the case when W is singular. We note that, unlike the Lanczos method, the iterative subspace used in the 4-D SD algorithm contains a fixed number $2(k + i)$ of vectors; since the dimension of the iterative subspace does not increase as a function of the number of iterations, the 4-D algorithm does not suffer from the numerical instabilities typical of the Lanczos algorithm, such as the loss of (bi-)orthogonality.

Line 7: This is an eigenvalue problem with the same matrix structure as the original RPA problem, but of $2(k + i) \times 2(k + i)$ dimension, which is much smaller than $2m \times 2m$. This small RPA eigenvalue problem can be treated as a dense eigenvalue problem and solved by using LAPACK routines.[28]

Finally, we note that one can incorporate a preconditioning scheme in the block 4-D algorithm for faster convergence rate. In this case, at each iteration, we seek to pre-condition the search direction

$$\begin{pmatrix} R_K \\ R_M \end{pmatrix} := \Phi \begin{pmatrix} R_K \\ R_M \end{pmatrix} \quad (22)$$

between Lines 5 and 6, where Φ is a properly chosen preconditioner. It is an important subject of future work. In Secs. V and VI, we show that even without a preconditioner, the block 4-D algorithm shows satisfactory convergence property.

V. VALIDATION OF THE ALGORITHM

The block 4-D steepest descent algorithm has been implemented in the plane wave-pseudopotential turboTDDFT code,[22] that is part of the QUANTUM ESPRESSO (QE) package.[26] The turboTDDFT code provides an implementation of the density matrix perturbation theory formulation of TDDFT given in Eqs. (6) and (7). Using fast Fourier transform techniques, the multiplication of the TDDFT matrix with a vector can be efficiently performed without building explicitly or storing the full matrix.

To test the accuracy of the new algorithm, simplified TDDFT matrices for the optical spectra of the Na_2 and Na_4 sodium clusters were built explicitly; the results of the direct diagonalization using LAPACK libraries[28] were compared with the results of our iterative algorithm. We denote such matrices as simplified since we used a plane-wave energy

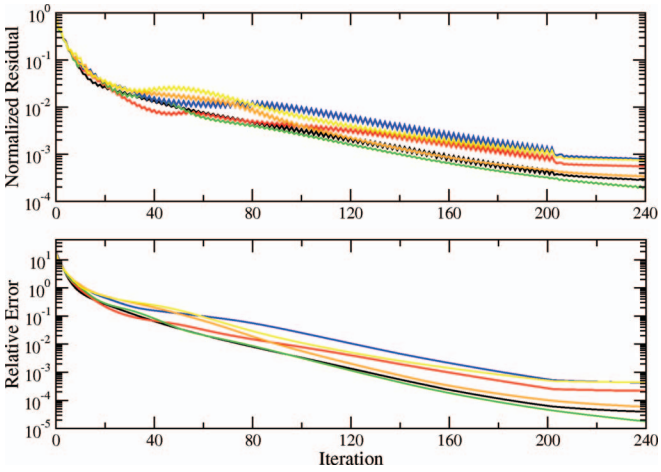

FIG. 1. Iterative diagonalization of a simplified TDLDA eigenproblem (Na_2 molecule; see text): Normalized residual (defined by Eq. (21)) as a function of the number of iterations for the six lowest positive eigenvalues (top panel); relative error of the eigenvalues computed iteratively with respect to the exact diagonalization as a function of the number of iterations (bottom panel).

cutoff that does not correspond to a fully converged basis set and relatively small unit cells. We had to resort to these simplifications to be able to explicitly build and diagonalize the matrices. The Na_2 calculation was performed using a cubic supercell of side 17 a_0 and a cutoff of 8 Ry to expand the wavefunctions (32 Ry for the charge density); in the Na_4 calculation we used a cubic supercell of side 22 a_0 and a 4 Ry cutoff. In Sec. VI we will consider a fully converged calculation for the benzene molecule; in this case the very large dimension of the explicit matrix do not enable storage and direct diagonalization using LAPACK routines. For the sodium cluster calculations the local density approximation (LDA) in the Perdew-Zunger parametrization[29] was used and the norm-conserving pseudopotentials were taken from the QE library.[30] For Na_2 the dimension of the TDLDA matrix is

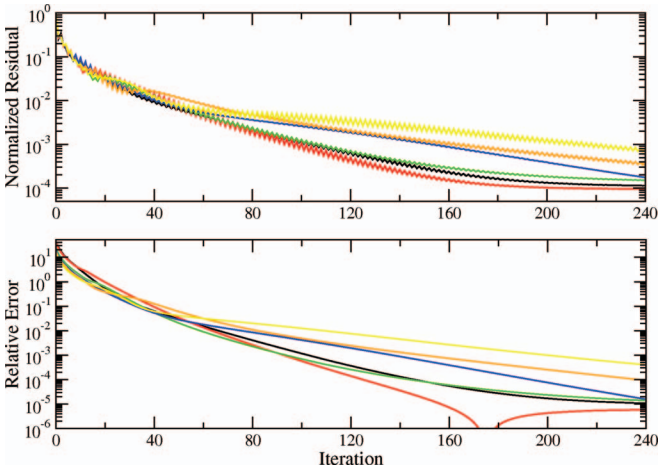

FIG. 2. Iterative diagonalization of a simplified TDLDA eigenproblem (Na_4 molecule; see text): Normalized residual (defined by Eq. (21)) as a function of the number of iterations for the six lowest positive eigenvalues (top panel); relative error of the eigenvalues computed iteratively with respect to the exact diagonalization as a function of the number of iterations (bottom panel).[31]

TABLE I. First six eigenvalues (eV) of a simplified TDLDA eigenproblem (Na_2 molecule; see text) computed by the 4-D steepest descent algorithm described in this work. The calculation was stopped after 240 iterations.

Exact diagonalization	Iterative diagonalization	Relative error	Absolute error	Normalized residual
2.663634	2.663740	4.0×10^{-5}	-1.1×10^{-4}	2.8×10^{-4}
2.894810	2.895445	2.2×10^{-4}	-6.4×10^{-4}	5.5×10^{-4}
2.983893	2.983947	1.8×10^{-5}	-5.5×10^{-5}	1.9×10^{-4}
2.983893	2.985178	4.3×10^{-4}	-1.3×10^{-3}	7.8×10^{-4}
3.173206	3.173396	6.0×10^{-5}	-1.9×10^{-4}	3.3×10^{-4}
3.497192	3.498715	4.4×10^{-4}	-1.5×10^{-3}	7.4×10^{-4}

1864 × 1864. The iterative calculations were performed with a threshold *tol* of 10^{-3} in the convergence test Eq. (21). In the top panel of Fig. 1 we show the behavior of the residual defined in Eq. (21) as a function of the number of iterations. Even without the help of a preconditioner the residual steadily decreases. A similar behavior is shown by the relative error in the bottom panel of Fig. 1. The relative error of the eigenvalues was computed by comparing the results of the iterative calculation with those of the exact diagonalization with LAPACK libraries. Below the convergence threshold of 10^{-3} we found that the residual can be used as an upper bound to the relative error. In Table I we show explicitly the calculated eigenvalues together with the relative and absolute errors after 240 iterations. The absolute errors in eV are definitely well below the accuracy required by this kind of calculations (a numerical accuracy in the diagonalization within 0.01 eV can be considered satisfactory). In Fig. 2 and Table II we show the same quantities for Na_4. In this case the dimension of the explicit matrix is 2840 × 2840. Considerations similar to those of the Na_2 case hold here, with the residual and the relative error steadily decreasing as a function of the number of iterations.

VI. APPLICATION TO BENZENE

In this section we present an application of the 4-D SD algorithm introduced in Sec. IV to a more challenging example, namely, a fully converged calculation of the low-lying TDLDA excitation spectrum of the benzene molecule. We also provide an assignment of the transitions and we compare our results with previous calculations.[6,16,32] We considered the benzene molecule in a tetragonal cell of dimension

TABLE II. First six eigenvalues (eV) of a simplified TDLDA eigenproblem (Na_4 molecule; see text) computed by the 4-D steepest descent algorithm described in this work. The calculation was stopped after 240 iterations.

Exact diagonalization	Iterative diagonalization	Relative error	Absolute error	Normalized residual
0.6722296	0.6722369	1.1×10^{-5}	-7.3×10^{-6}	1.1×10^{-4}
0.7585424	0.7585379	5.9×10^{-6}	4.4×10^{-6}	9.5×10^{-5}
1.282119	1.282137	1.4×10^{-5}	-1.8×10^{-5}	1.5×10^{-4}
1.613599	1.613626	1.6×10^{-5}	-2.6×10^{-5}	1.8×10^{-4}
1.811909	1.812080	9.4×10^{-5}	-1.7×10^{-4}	3.4×10^{-4}
2.147626	2.148505	4.1×10^{-4}	-8.8×10^{-4}	6.8×10^{-4}

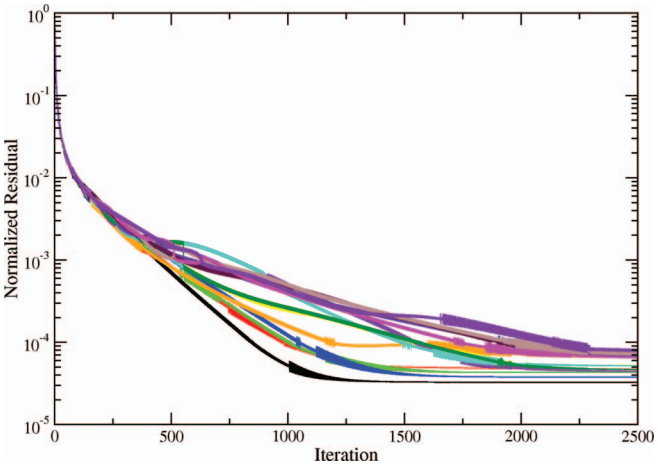

FIG. 3. Iterative calculation of the low-lying TDLDA excitation energies of benzene: Normalized residual (defined by Eq. (21)) as a function of the number of iterations.

$30 \times 30 \times 20$ a_0^3 and we used a 60 Ry cutoff to expand the wavefunctions, corresponding to 70597 plane-waves (PWs). The LDA in the Perdew-Zunger parametrization[29] was used. The dimensions of the explicit TDLDA matrix are 2 117 910 $\times$ 2 117 910 and the direct diagonalization using LAPACK libraries or even the memory storage of the full matrix are prohibitive. In order to test the convergence of our results with respect to the basis set dimension we have performed a calculation using a 70 Ry cutoff for the wavefunctions (corresponding to a 2 670 150 $\times$ 2 670 150 explicit RPA matrix); the results were not significantly different (differences smaller than 0.01 eV) with respect to those of the 60 Ry calculation and they will not be explicitly reported in this paper. Furthermore, we carefully tested the convergence of the excitation energies with respect to the supercell size, considering a $40 \times 40 \times 30$ a_0^3 supercell (which corresponds to a 5 650 410 $\times$ 5 650 410 explicit RPA matrix). The differences in the excitations computed for different cell sizes were within 0.1 eV. The 4-D SD algorithm has been applied with a threshold *tol* of 10^{-4}. In Fig. 3 we show the behavior of the normalized residual Eq. (21) for the first 14 eigenvalues, where the degenerate 13th and 14th eigenvalues correspond to the E_{1u} excitation in Table III ($30 \times 30 \times 20$ a_0^3 cell and 60 Ry cutoff). Also in this case, even without using a preconditioner, the residual steadily decreases below the threshold. The calculations on the larger matrices, corresponding to a larger wavefunction cutoff (70 Ry) or a larger supercell ($40 \times 40 \times 30$ a_0^3), have shown a convergence rate of the normalized residual similar to that shown in Fig. 3. This means that the convergence rate, rather than depending on the matrix size, depends on the matrix condition number.

In Table III we present our results for some low-lying transitions and we compare them with other calculations in the literature; for the sake of completeness the experimental results are also reported in the last column. In order to assign these transitions we have used a scheme similar to the one proposed by Casida.[9] The results of Refs. 16 and 32 were obtained using Gaussian-type localized basis sets with added diffuse functions, such as 6-31+G*, AUG-cc-pVTZ, and pVTZ+. In Ref. 6 a real space grid implementation was used. The low energy spectrum of benzene is characterized by a few valence and Rydberg excitations. We note that the Rydberg excitations involve very delocalized orbitals and can be strongly affected by the local basis set used.[16,32] For the valence excitations with $\pi \rightarrow \pi^*$ character our results are in satisfactory agreement with the literature, with differences of at most 0.13 eV. More challenging is the comparison for the Rydberg excitations. In this case we find differences up to 0.3–0.4 eV. This discrepancy is likely due to the limited accuracy of the basis sets used in Refs. 16 and 32. As shown in Table III, the differences between the results obtained with 6-31+G* and the AUG-cc-pVTZ basis sets can be as large as 0.17 eV for the $\pi \rightarrow 3s$ transition and as large as about 0.2 eV for the $\pi \rightarrow 3p$ excitations (results from Ref. 16). As discussed in Ref. 32, the use of a basis set without diffuse functions such as pVTZ completely misses the description of the $\pi \rightarrow 3p$ excitations. In our plane-wave implementation the basis set is large enough and particularly suitable to describe delocalized states; its accuracy can be systematically tested by just increasing the wavefunction cutoff. By increasing the wavefunction cutoff from 60 Ry to 70 Ry the differences are smaller than 0.01 eV for all the excitation energies considered here. The effect of the supercell approach on the accuracy of our calculation has also been considered. In Table III we compare the results obtained using a $30 \times 30 \times 20$ a_0^3 cell and a $40 \times 40 \times 20$ a_0^3 cell. Small differences are found but always within 0.1 eV. In conclusion, taking into account the quality of convergence of our and other calculations, we consider our results in satisfactory agreement with previously published data.

TABLE III. Comparison between TDLDA excitation energies (eV) of the benzene molecule as computed in this work and published results. The second row specifies the cell sizes used in our work, the basis set used in Refs. 16 and 32 and the technique used in Ref. 6.

Transitions	This work	This work	Ref. 16	Ref. 16	Ref. 32	Ref. 6	Expt. (Ref. 32)
	$30 \times 30 \times 20$ a_0^3	$40 \times 40 \times 20$ a_0^3	6-31+G*	AUG-cc-pVTZ	pVTZ+	real space mesh	
B_{2u} ($\pi \rightarrow \pi^*$)	5.39	5.39	5.31	5.26	5.28	5.40	4.90
E_{1g} ($\pi \rightarrow 3s$)	5.95	6.03	6.36	6.19	5.99	n.a.	6.33
B_{1u} ($\pi \rightarrow \pi^*$)	6.10	6.10	6.10	6.02	6.10	6.23	6.20
E_{2u} ($\pi \rightarrow 3p$)	6.58	6.56	6.98	6.80	6.45	n.a.	6.95
A_{2u} ($\pi \rightarrow 3p$)	6.58	6.56	6.99	6.80	6.44	n.a.	6.93
E_{1u} ($\pi \rightarrow \pi^*$)	6.95	6.86	6.94	n.a.	6.92	6.9-7.2	6.94

VII. CONCLUSIONS

In this paper we have established a block minimization principle for the non-Hermitian RPA eigenvalue problem, as defined by Eq. (1). This problem appears in the solution of the TDDFT equations and the BSE. Within the proposed formalism, we have developed a four dimensional steepest descent-like algorithm that can compute simultaneously several low-lying positive eigenvalues. We have first tested the accuracy and stability of this approach on some simplified TDDFT calculations of the excitation spectra of sodium clusters. The small size of the matrices considered in these cases allowed us to draw a systematic comparison between the results of our iterative technique and those of exact diagonalization. The agreement was found to be excellent. Then we have computed the low-lying TDDFT spectrum of benzene, using the LDA; we found good agreement with previously published data. In all the examples considered here, our SD-like algorithm has shown a steady variational convergence similar to that of Hermitian matrix techniques. The principle and algorithm presented here enable the assignment of excitation peaks to specific transitions between single particle states, when using TDDFT without the Tamm-Damcoff approximation, and when using techniques based on density functional perturbation theory which do not require the explicit calculation of empty electronic states. The study of suitable preconditioning schemes and the application of this method to the BSE will be subject of future work.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation (NSF) (Grant Nos. OCI-0749217, DMR-1035468, CHEM-68D-1086057, DMS-0810506, and DMS-1115817/1115834), and through TeraGrid resources provided by TACC and NICS (Grant Nos. TG-ASC090004 and TG-MCA06N063).

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[2] R. Bauernschmitt and R. Ahlrichs, Chem. Phys. Lett. **256**, 454 (1996).
[3] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. B **57**, R9385 (1998).
[4] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998).
[5] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[6] M. L. Tiago and J. R. Chelikowsky, Phys. Rev. B **73**, 205334 (2006).
[7] M. Gruning, A. Marini, and X. Gonze, Nano Lett. **9**, 2820 (2009).
[8] D. Rocca, D. Lu, and G. Galli, J. Chem. Phys. **133**, 164109 (2010).
[9] M. E. Casida, in *Recent Advances in Density Functional Methods, Part I*, edited by D. P. Chong (World Scientific, Singapore, 1995), p. 155.
[10] B. Walker, A. M. Saitta, R. Gebauer, and S. Baroni, Phys. Rev. Lett. **96**, 113001 (2006).
[11] D. Rocca, R. Gebauer, Y. Saad, and S. Baroni, J. Chem. Phys. **128**, 154105 (2008).
[12] D. J. Thouless, Nucl. Phys. **22**, 78 (1961).
[13] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[14] I. Tamm, J. Phys. (Moscow) **9**, 449 (1945).
[15] S. M. Dancoff, Phys. Rev. **78**, 382 (1950).
[16] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **109**, 8218 (1998).
[17] E. V. Tsiper, J. Phys. B **34**, L401 (2001).
[18] A. Muta, J. Iwata, Y. Hashimoto, and K. Yabana, Prog. Theor. Phys. **108**, 1065 (2002).
[19] M. J. Lucero, A. M. N. Niklasson, S. Tretiak, and M. Challacombe, J. Chem. Phys. **129**, 064114 (2008).
[20] S. Tretiak, C. M. Isborn, A. M. N. Niklasson, and M. Challacombe, J. Chem. Phys. **130**, 054111 (2009).
[21] M. Challacombe, e-print arXiv:1001.2586v2 [quant-ph].
[22] O. B. Malcıoğlu, R. Gebauer, D. Rocca, and S. Baroni, Comput. Phys. Commun. **182**, 1744 (2011).
[23] P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **72**, 245425 (2005).
[24] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).
[25] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[26] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009).
[27] Z. Bai and R.-C. Li, CS-Technical Report, UC Davis, 2011.
[28] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. D. Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, LAPACK Users' Guide, 3rd ed. (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999).
[29] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[30] The QUANTUM ESPRESSO pseudopotential library is available online at http://www.quantum-espresso.org/pseudo.php. In this work the pseudopotential file `Na.pz-n-vbc.UPF` was used for sodium, `C.pz-vbc.UPF` for carbon, and `H.pz-vbc.UPF` for hydrogen.
[31] Instabilities might be present (red curve) when the matrix W in the Line 6 of the 4-D SD algorithm becomes nearly singular.
[32] H. H. Heinze, A. Görling, and N. Rösch, J. Chem. Phys. **113**, 2088 (2000).
[33] For example, we can take $W_1 = W^T$ and $W_2 = I$ for simplicity.