JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

On the Efficiency of Algorithms for Solving Hartree−Fock and Kohn−Sham Response Equations

Joanna Kauczor,*[,†] Poul Jørgensen,[†] and Patrick Norman[‡]

[†]Lundbeck Foundation Center for Theoretical Chemistry, Department of Chemistry, Aarhus University, Langelandsgade 140, Dk-8000 Aarhus C, Denmark

[‡]Department of Physics, Chemistry, and Biology, Linköping University, Se-581 83 Linköping, Sweden

ABSTRACT: The response equations as occurring in the Hartree−Fock, multiconfigurational self-consistent field, and Kohn−Sham density functional theory have identical matrix structures. The algorithms that are used for solving these equations are discussed, and new algorithms are proposed where trial vectors are split into symmetric and antisymmetric components. Numerical examples are given to compare the performance of the algorithms. The calculations show that the standard response equation for frequencies smaller than the highest occupied molecular orbital−lowest unoccupied molecular orbital gap is best solved using the preconditioned conjugate gradient or conjugate residual algorithms where trial vectors are split into symmetric and antisymmetric components. For larger frequencies in the standard response equation as well as in the damped response equation in general, the preconditioned iterative subspace approach with symmetrized trial vectors should be used. For the response eigenvalue equation, the Davidson algorithm with either paired or symmetrized trial vectors constitutes equally good options.

1. INTRODUCTION

Molecular properties are fundamental quantities underlying the macroscopic behavior of matter and their determination constitutes one of the most fruitful areas of interplay between experiment and theory. From a theoretical point of view, the determination of molecular properties can be achieved by calculations of response functions that express the responses of a molecular system toward weak perturbing fields, such as internal magnetic moments or externally applied electric and magnetic fields.[1] By considering poles and residues of response functions, it is also possible to evaluate energy separations and various couplings between specific states and molecular properties of excited states.[1]

In approximate state electronic structure theory, the major task when calculating response functions and their poles and residues translates into solving sets of response equations, and the key module in programs devoted to the determination of molecular properties therefore is the response equation solver. The large dimensionality of the response equations imposes the use of iterative algorithms, and the two criteria against which such solver routines should be judged are stability and efficiency. We consider the solution of response equations for variational wave functions, such as Hartree−Fock (HF), multiconfiguration self-consistent field (MCSCF), and Kohn−Sham (KS) density functional theory (DFT). The overall structure of the response equations is identical for these wave functions, and our discussion will be relevant to all these cases. We will propose new algorithms based on a splitting of trial vectors into symmetric and antisymmetric components and, by numerical examples, compare them to the conventional algorithms that have been used for solving response equations.

For a real unperturbed reference state, three different types of response equations will be considered:

- Standard response equation: The standard response equation determines the responses in the wave function parameters (collected as a vector $\mathbf{X}_S$) toward an off-resonance time-dependent perturbing field oscillating with frequency ω:

$$(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})\mathbf{X}_S = \mathbf{G} \tag{1}$$

  where $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ are the generalized Hessian and metric matrix, respectively, and $\mathbf{G}$ is a generalized gradient for the operator describing the coupling between the perturbing field and the molecular system—we refer to ref 1 for the definition of the response matrices in the HF and MCSCF theory and to ref 2 in the case of KS theory. We note that eq 1 represents a set of linear equations for a real symmetric matrix.
- Damped response equation: In near-resonance and resonance regions of the spectrum, eq 1 no longer provides a reasonable description of the induced polarization. However, in a situation with fast relaxation channels that depopulates the manifold of excited states, perturbation theory may still be applicable as long as relaxation parameters (here represented by a single parameter γ) are introduced:[3−5]

$$(\mathbf{E}^{[2]} - (\omega + i\gamma)\mathbf{S}^{[2]})\mathbf{X}_D = \mathbf{G} \tag{2}$$

  The damping parameters are associated with the inverse (finite) lifetimes of the excited states.[6,7] In this case, the solution vector $\mathbf{X}_D$ is complex with real and imaginary components describing dispersion and absorption processes, respectively. We note that eq 2 represents a set of linear equations for a complex non-Hermitian matrix.
- Response eigenvalue equation: Excitation energies occur at the poles of the linear response function and can be determined by solving the generalized eigenvalue equation:

$$(\mathbf{E}^{[2]} - \omega_f\mathbf{S}^{[2]})\mathbf{X}_f = 0 \tag{3}$$

Received: December 20, 2010
Published: April 29, 2011

where ω_f is the transition frequency between the reference state $|0\rangle$ and excited state $|f\rangle$, and $\mathbf{X}_f$ is the excitation vector. Since the metric $\mathbf{S}^{[2]}$ is not positive definite, eq 3 represents a non-Hermitian eigenvalue equation.

In the early days of quantum chemistry only small molecular systems were considered, and the response matrices in eqs 1−3 were set up explicitly. The solutions to these equations were obtained by applying standard algorithms of numerical analysis. The standard response equation (eq 1) was for example solved using diagonalization and triangularization methods, e.g., using Gauss−Jordan elimination or LU factorization.[8−10] The response eigenvalue equation (eq 3) was solved using standard algorithms for a non-Hermitian eigenvalue problem. However, since solving a Hermitian eigenvalue equation is simpler and, from an algorithmic point of view, more robust as compared to solving a non-Hermitian eigenvalue equation, eq 3 was expressed in a Hermitian form that involved the diagonalization of two Hermitian matrices of half the dimension of the $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ matrices.[11]

For larger molecular systems it has become a standard practice in quantum chemistry to use iterative subspace algorithms. For example, Pople et al.[12] solved the coupled HF equations (eq 1) and Purvis and Bartlett[13] the coupled cluster singles and doubles amplitude equations using iterative subspace algorithms. Wormer et al.[14] recognized that both these algorithms lead to an iteration sequence identical to the one obtained using the conjugate gradient (CG) algorithm.[15] The advantage of using a CG formulation,[15−18] and similar for the conjugate residual (CR) algorithm,[19] is that only the last three trial vectors are necessary to store the information content of all previous trial vectors. In a subspace approach the last three trial vectors together with their linear transformed vectors need to be stored on disk. In a unidirectional formulation of the CG and CR algorithms, the number of vectors that needs to be stored on disk in each iteration is reduced to three and four, respectively. The handling and storage of trial vectors therefore become simplified in the CG algorithm compared to an iterative subspace formulation. In Appendix A, the connection between the CG and subspace algorithms will be discussed. Our discussion will be referring to system sizes that prohibit the explicit formation of the Hessian matrix but yet small enough to allow the Fock/KS matrix transformations needed to achieve a preconditioning in the canonical self-consistent field (SCF) orbital basis. Using current standard mathematical library routines, diagonalization of Fock matrices of dimensions up to ca. 5×10^4 can straightforwardly be performed, but we acknowledge that in order to reach significantly larger system sizes the orbital transformation becomes a bottleneck, and "orbital-free" approaches have to be used, as has been demonstrated for time-independent[20] as well as time-dependent response theory.[21]

The Davidson algorithm[22] is an iterative subspace algorithm for solving a standard Hermitian eigenvalue equation. It has been generalized to the non-Hermitian response eigenvalue equation (eq 3).[23] However, this generalization turned out unsuccessful because the non-Hermitian character of the response eigenvalue equation may introduce complex eigenvalues in the reduced subspace eigenvalue equations, which, in turn, made it difficult to generate new trial vectors and converge the response eigenvalue equation.[24] A remedy to this problem was proposed by Olsen et al.,[25] who for ground-state calculations recognized that the response eigenvalue equation could be viewed as a Hermitian eigenvalue equation because the generalized Hessian $\mathbf{E}^{[2]}$ is positive definite and therefore could be viewed as a metric with the inverse of the excitation energies as eigenvalues. Olsen et al. further recognized that the response eigenvalue equation has paired eigenvalues and that it is important to keep the paired structure of the eigenvalues in the reduced space eigenvalue equations. Paired eigenvalues can be obtained by adding paired trial vectors to the subspace in the iterative procedure. With these advances the Davidson algorithm was successfully applied.[25] We analyze the theoretical background for adding paired trial vectors. We also discuss the Olsen algorithm[26] in comparison with the Davidson algorithm, for the generation of new trial vectors in the iterative subspace algorithm.

From the linear transformation of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ on a trial vector, the linear transformation on its paired counterpart can be generated at no additional cost. This suggests that it may be advantageous to solve also the standard response equation using the iterative subspace algorithm where trial vectors are added in pairs to the reduced subspace. In doing so, however, the attractive feature of the CG algorithm to keep the information from the entire trial vector space in the last three vectors is lost. But, as we shall see in the present work, by invoking an algorithm with symmetrized trial vectors, this feature can be retained while keeping the improved convergence rate seen when paired trial vectors are used.

In 1990, Casida suggested to solve the standard response and response eigenvalue equations by means of a block diagonalization of the Hessian (see ref 2) using an approach similar to the one described in ref 11. This algorithm is briefly discussed in Appendix D.

Saue and co-workers[27−29] introduced a symmetry division of trial vectors for solving the standard response and eigenvalue equations in the four-component HF and DFT approximations. In their approach, trial vectors are split according to their Hermiticity and time-reversal symmetry, which are the two fundamental operator symmetries in a relativistic framework. The trial vectors are complex due to the complex nature of the four-component wave function, however, the reduced subspace equations remain real, due to the symmetry properties of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$. Recently, Villaume et al.[30] have generalized this approach to solve the damped response equation, including a presentation of a highly efficient preconditioner. This development represented a significant improvement of the original algorithm for the solution of damped response equations presented by Norman et al.[3,4] (with computational improvements by Kristensen et al.),[5] in which the complex response equation was considered as two coupled real equations.

In the spirit of the work by Villaume et al.,[30] we introduce an algorithm for solving response equations (eqs 1−3) in a nonrelativistic framework (in a basis of real orbitals) where the solution vectors are expressed in terms of symmetric and antisymmetric components constituting the real and imaginary parts of Hermitian and anti-Hermitian vectors. We show that the standard response equation within this framework for frequencies smaller than the highest occupied molecular orbital−lowest unoccupied molecular orbital (HOMO−LUMO) gap may be solved using the preconditioned CG/CR algorithm, while for larger frequencies the preconditioned iterative subspace algorithm with symmetrized trial vectors should be used. We also show that the damped response equation (eq 2) may be expressed in terms of a set of linear equations for a symmetric but not positive definite matrix. These algorithms may be supplemented

with an efficient preconditioner, analogous to the one proposed by Villaume et al.[30] However, since this preconditioner is not positive definite, the preconditioned CR algorithm cannot be safely used, as discussed in Appendix B, and the damped response equation is therefore best solved using the preconditioned iterative subspace approach with symmetrized trial vectors.

In the next section, the structure of the response matrices in a basis of nonrelativistic real orbitals will be discussed. In Section 3, the standard iterative subspace algorithms are discussed for solving the three categories of response equations. In Section 4 iterative subspace methods with paired trial vectors are considered, and in Section 5 the advantages of using symmetrized trial vectors are discussed. In Section 6, we provide numerical examples to illustrate the efficiency of the various algorithms, and Section 7 contains the concluding remarks.

2. STRUCTURES OF RESPONSE MATRICES

The response equations are often written in a 2×2 matrix blocked form referencing the excitation and deexcitation spaces.[1] In this form, $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ have the structure:

$$\mathbf{E}^{[2]} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}, \quad \mathbf{S}^{[2]} = \begin{pmatrix} \Sigma & \Delta \\ -\Delta & -\Sigma \end{pmatrix} \tag{4}$$

where

$$\begin{aligned} A_{pq} &= \langle 0|[[Q_p, \hat{H}_0], Q_q^\dagger]|0\rangle \\ B_{pq} &= \langle 0|[[Q_p, \hat{H}_0], Q_q]|0\rangle \\ \Sigma_{pq} &= \langle 0|[Q_p, Q_q^\dagger]|0\rangle \\ \Delta_{pq} &= \langle 0|[Q_p, Q_q]|0\rangle \end{aligned} \tag{5}$$

where $|0\rangle$ is the reference state, $\hat{H}_0$ is the nonrelativistic Hamiltonian, and $Q_p^\dagger$ and Q_p are excitation and deexcitation operators, respectively. Both $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ are symmetric matrices. $\mathbf{A}$, $\mathbf{B}$, and Σ are symmetric, and Δ is antisymmetric.[1]

For a closed shell system and a HF reference state $|0\rangle = |\mathrm{HF}\rangle$, the excitation operator $Q_p^\dagger$ becomes

$$Q_p^\dagger = Q_{AI}^\dagger = \frac{1}{\sqrt{2}}(a_{A\alpha}^\dagger a_{I\alpha} + a_{A\beta}^\dagger a_{I\beta}) \tag{6}$$

where I and A refer to occupied and unoccupied molecular orbital (MO) indices, respectively. In the canonical SCF basis, the elements of $\mathbf{A}$, $\mathbf{B}$, Σ, and Δ become[11]

$$\begin{aligned} A_{pq} = A_{AI,BJ} &= \langle \mathrm{HF}|[[Q_p, \hat{H}_0], Q_q^\dagger]|\mathrm{HF}\rangle \\ &= \delta_{AB}\delta_{IJ}(\varepsilon_A - \varepsilon_I) + (AI|JB) - (AB|IJ) \end{aligned} \tag{7}$$

$$\begin{aligned} B_{pq} = B_{AI,BJ} &= \langle \mathrm{HF}|[[Q_p, \hat{H}_0], Q_q]|\mathrm{HF}\rangle \\ &= (AI|BJ) - (AJ|BI) \end{aligned} \tag{8}$$

$$\Sigma_{pq} = \langle \mathrm{HF}|[Q_p, Q_q^\dagger]|\mathrm{HF}\rangle = \delta_{AB}\delta_{IJ} \tag{9}$$

$$\Delta_{pq} = \langle \mathrm{HF}|[Q_p, Q_q]|\mathrm{HF}\rangle = 0 \tag{10}$$

respectively. ε_A and ε_I refer to orbital energies, and the Mulliken notation is used for the two-electron integrals. In KS theory, the two-electron integrals in eqs 7 and 8 have to be modified, and an exchange−correlation contribution added.[2] In MCSCF wave function theory, the explicit expressions for the matrix elements in the electronic Hessian and overlap matrices are significantly more complicated, but the form is identical to that in single determinant theory.[1] Our present considerations regarding response equation solvers will be valid also in the case of MCSCF theory but with use of MCSCF specific approximations of the Hessians and overlap matrices for the construction of preconditioners.

A Møller−Plesset perturbation analysis shows that the $\mathbf{E}^{[2]}$ matrix can be split into the zeroth and the first order contribution:

$$\mathbf{E}^{[2]} = \mathbf{E}_0^{[2]} + \mathbf{E}_1^{[2]} \tag{11}$$

where $\mathbf{E}_0^{[2]}$ is a diagonal matrix containing the orbital energy differences:

$$\mathbf{E}_0^{[2]} = \begin{pmatrix} \Delta\varepsilon & 0 \\ 0 & \Delta\varepsilon \end{pmatrix} \tag{12}$$

where the elements of $\Delta\varepsilon$ are given by

$$\Delta\varepsilon_{AI,BJ} = \delta_{AB}\delta_{IJ}(\varepsilon_A - \varepsilon_I) \tag{13}$$

$\mathbf{E}_1^{[2]}$ contains the electron−electron repulsion contributions to $\mathbf{E}^{[2]}$, as can be seen from eqs 7 and 8. In KS theory, an exchange−correlation contribution has to be added to $\mathbf{E}_1^{[2]}$. The metric matrix $\mathbf{S}^{[2]}$ in the canonical SCF representation becomes

$$\mathbf{S}^{[2]} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \tag{14}$$

The solution vectors $\mathbf{X}$ in eqs 1−3 can also be written in the blocked form as

$$\mathbf{X} = \begin{pmatrix} \boldsymbol{x}_{AI} \\ \boldsymbol{x}_{JB} \end{pmatrix} \tag{15}$$

where $\boldsymbol{x}_{AI}$ and $\boldsymbol{x}_{JB}$ are the excitation and deexcitation components,[1] respectively.

Using the above notation, the response eigenvalue equation in the canonical SCF representation may be expressed as

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} x_{AI,f} \\ x_{JB,f} \end{pmatrix} = \omega_f \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{AI,f} \\ \boldsymbol{x}_{JB,f} \end{pmatrix} \tag{16}$$

where the eigenvector satisfies the positive normalization condition

$$(\boldsymbol{x}_{AI,f}^T \quad \boldsymbol{x}_{JB,f}^T) \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{AI,f} \\ \boldsymbol{x}_{JB,f} \end{pmatrix} = 1 \tag{17}$$

Together with the positive $\mathbf{S}^{[2]}$-normed solution $\mathbf{X}_{f+} = \begin{pmatrix} \boldsymbol{x}_{AI,f} \\ \boldsymbol{x}_{JB,f} \end{pmatrix}$, the response eigenvalue equation has a paired solution $\mathbf{X}_{f-} = \begin{pmatrix} \boldsymbol{x}_{JB,f} \\ \boldsymbol{x}_{AI,f} \end{pmatrix}$, with the eigenvalue $-\omega_f$.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{JB,f} \\ \boldsymbol{x}_{AI,f} \end{pmatrix} = -\omega_f \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{JB,f} \\ \boldsymbol{x}_{AI,f} \end{pmatrix} \tag{18}$$

where the eigenvector satisfies the negative normalization condition:

$$\begin{pmatrix} \boldsymbol{x}_{JB,f}^T & \boldsymbol{x}_{AI,f}^T \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{JB,f} \\ \boldsymbol{x}_{AI,f} \end{pmatrix} = -1 \qquad (19)$$

The excitation vectors may be collected as the columns of an eigenvector matrix $\mathbf{X}$ with the positive $\mathbf{S}^{[2]}$-normed eigenvectors collected first $\mathbf{X}_+ = \{\mathbf{X}_{1+}, \mathbf{X}_{2+}, ..., \mathbf{X}_{f+}, ...\}$, followed by the negative $\mathbf{S}^{[2]}$-normed vectors $\mathbf{X}_- = \{\mathbf{X}_{1-}, \mathbf{X}_{2-}, ..., \mathbf{X}_{f-}, ...\}$.[31]

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_+ & \mathbf{X}_- \end{pmatrix} \qquad (20)$$

In matrix form the response matrices may be expressed as[1]

$$\mathbf{X}^\dagger \mathbf{E}^{[2]} \mathbf{X} = \begin{pmatrix} \boldsymbol{\omega}_{\mathrm{exc}} & 0 \\ 0 & \boldsymbol{\omega}_{\mathrm{exc}} \end{pmatrix}; \quad \mathbf{X}^\dagger \mathbf{S}^{[2]} \mathbf{X} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \qquad (21)$$

where $\boldsymbol{\omega}_{\mathrm{exc}}$ is a diagonal matrix containing the excitation energies $\{\omega_1, \omega_2, ..., \omega_f, ...\}$.

For the ground state, $\mathbf{E}^{[2]}$ is a positive definite matrix. To see this, $\mathbf{E}^{[2]}$ may be transformed to block diagonal form using the unitary matrix $\mathbf{U}$:

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad (22)$$

giving

$$\mathbf{U}^\dagger \mathbf{E}^{[2]} \mathbf{U} = \begin{pmatrix} \mathbf{A}+\mathbf{B} & 0 \\ 0 & \mathbf{A}-\mathbf{B} \end{pmatrix} \qquad (23)$$

Submatrices $\mathbf{A}+\mathbf{B}$ and $\mathbf{A}-\mathbf{B}$ represent the stability conditions[32] in the wave function $|0\rangle$ with respect to imaginary and real variations, respectively. For a ground state, both $\mathbf{A}-\mathbf{B}$ and $\mathbf{A}+\mathbf{B}$ have to be positive definite, implying that $\mathbf{E}^{[2]}$ is positive definite. For a ground state, $\mathbf{E}^{[2]}$ may thus be viewed as a positive definite metric matrix, and the response eigenvalue equation, eq 3, may be expressed as

$$\left(\frac{1}{\omega_f} \mathbf{E}^{[2]} - \mathbf{S}^{[2]} \right) \mathbf{X}_f = 0 \qquad (24)$$

which is a standard symmetric eigenvalue equation where $(\omega_f)^{-1}$ are the eigenvalues.

In the following discussion, we will focus on HF and KS theory and assume that we are using a canonical SCF representation, where the generalized metric $\mathbf{S}^{[2]}$ has the structure in eq 14, and the generalized Hessian $\mathbf{E}^{[2]}$ can be split into the zeroth and the first order contribution as in eq 11. However, in the practical calculations, we evaluate the linear transformations of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ on trial vectors in the atomic orbital (AO) representation, as discussed in detail in refs 21 and 33. The detailed structure of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ in the AO representation is described in ref 31. We note that the preconditioning of the response equations is always most efficiently performed in the canonical SCF representation where $\mathbf{S}^{[2]}$ has a particular structure (eq 14), which allows an exact treatment of this matrix and where the access to orbital energy differences enables a representation of the $\mathbf{E}^{[2]}$ matrix that is superior to available approximations in the AO basis.

3. SOLVING RESPONSE EQUATIONS USING STANDARD ITERATIVE METHODS

In electronic structure theory, the response equations (eqs 1−3) are conventionally solved using subspace iterative algorithms. In the subspace iterative algorithms, it is assumed that the linear transformations of the generalized Hessian $\mathbf{E}^{[2]}$ and metric $\mathbf{S}^{[2]}$ matrices on a trial vector $\mathbf{b}$ can be carried out

$$\boldsymbol{\sigma} = \mathbf{E}^{[2]}\mathbf{b}, \quad \boldsymbol{\rho} = \mathbf{S}^{[2]}\mathbf{b} \qquad (25)$$

The solution to the response equations is obtained from a sequence of linear transformations on trial vectors giving gradually improved solutions to the response equations.

3.1. General Subspace Iterative Algorithm. The solution to the response equations may be obtained using an iterative subspace algorithm. After iteration n of an iterative subspace algorithm we have n trial vectors:

$$\mathbf{b}^n = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\} \qquad (26)$$

and the linear transformed vectors

$$\boldsymbol{\sigma}^n = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n\}, \quad \boldsymbol{\rho}^n = \{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, ..., \boldsymbol{\rho}_n\} \qquad (27)$$

Reduced response equations are then set up in the subspace $\mathbf{b}^n$ giving

$$(\mathbf{E}_{\mathrm{R}}^{[2]} - \omega \mathbf{S}_{\mathrm{R}}^{[2]})(\mathbf{X}_{\mathrm{S}})_{\mathrm{R}} = \mathbf{G}_{\mathrm{R}} \qquad (28)$$

$$[\mathbf{E}_{\mathrm{R}}^{[2]} - (\omega + i\gamma)\mathbf{S}_{\mathrm{R}}^{[2]}](\mathbf{X}_{\mathrm{D}})_{\mathrm{R}} = \mathbf{G}_{\mathrm{R}} \qquad (29)$$

$$(\mathbf{E}_{\mathrm{R}}^{[2]} - \omega_f^{\mathrm{R}} \mathbf{S}_{\mathrm{R}}^{[2]})(\mathbf{X}_f)_{\mathrm{R}} = 0 \qquad (30)$$

for the standard, damped, and eigenvalue equation, respectively, where

$$(\mathbf{E}_{\mathrm{R}}^{[2]})_{ij} = \mathbf{b}_i^\dagger \boldsymbol{\sigma}_j, \quad (\mathbf{S}_{\mathrm{R}}^{[2]})_{ij} = \mathbf{b}_i^\dagger \boldsymbol{\rho}_j \qquad (31)$$

and

$$(\mathbf{G}_{\mathrm{R}})_i = \mathbf{b}_i^\dagger \mathbf{G} \qquad (32)$$

Equations 28−30 determine the optimal solution vector $\mathbf{X}_{n+1}$ in the subspace $\mathbf{b}^n$

$$\mathbf{X}_{n+1} = \sum_{i=1}^{n} (\mathbf{X}_{\mathrm{R}})_i \mathbf{b}_i \qquad (33)$$

The residuals for the response equations are given as

$$\begin{aligned} \boldsymbol{R}_{n+1} &= (\mathbf{E}^{[2]} - \omega \mathbf{S}^{[2]})(\mathbf{X}_{\mathrm{S}})_{n+1} - \mathbf{G} \\ &= \sum_{i=1}^{n} [(\mathbf{X}_{\mathrm{S}})_{\mathrm{R}}]_i (\boldsymbol{\sigma}_i - \omega \boldsymbol{\rho}_i) - \mathbf{G} \end{aligned} \qquad (34)$$

$$\begin{aligned} \boldsymbol{R}_{n+1} &= \mathbf{E}^{[2]} - (\omega + i\gamma)\mathbf{S}^{[2]}](\mathbf{X}_{\mathrm{D}})_{n+1} - \mathbf{G} \\ &= \sum_{i=1}^{n} [(\mathbf{X}_{\mathrm{D}})_{\mathrm{R}}]_i [\boldsymbol{\sigma}_i - (\omega + i\gamma)\boldsymbol{\rho}_i] - \mathbf{G} \end{aligned} \qquad (35)$$

$$\begin{aligned} \boldsymbol{R}_{n+1} &= (\mathbf{E}^{[2]} - \omega_f^{\mathrm{R}} \mathbf{S}^{[2]})(\mathbf{X}_f)_{n+1} \\ &= \sum_{i=1}^{n} [(\mathbf{X}_f)_{\mathrm{R}}]_i (\boldsymbol{\sigma}_i - \omega_f^{\mathrm{R}} \boldsymbol{\rho}_i) \end{aligned} \qquad (36)$$

for the standard, damped, and eigenvalue equation, respectively. To improve the convergence of the standard and damped response equations they may be preconditioned.[34] The new trial vector is then obtained from the preconditioned residual:

$$\mathbf{b}_{n+1} = \mathbf{L}^{-1}\boldsymbol{R}_{n+1} \tag{37}$$

where the preconditioner $\mathbf{L}$ is an approximation to $(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})$ for the standard response equation and $[\mathbf{E}^{[2]} - (\omega + i\gamma)\mathbf{S}^{[2]}]$ for the damped response equation. Due to the fact that $\mathbf{S}^{[2]}$ is a diagonal matrix (see eq 14) and $\mathbf{E}^{[2]}$ is a diagonally dominant matrix (see eq 11), $(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})$ and $[\mathbf{E}_0^{[2]} - (\omega + i\gamma)\mathbf{S}^{[2]}]$ may be used as preconditioners for the standard and damped response equations, respectively. The improvement in convergence obtained when preconditioning is introduced will be discussed in more detail in Section 3.3 and the subsequent sections.

When the eigenvalue equation (eq 3) is solved, new trial vectors may be obtained either using the Davidson[22] algorithm or the Olsen[26] algorithm, see details in the subsequent section. The new trial vector $\mathbf{b}_{n+1}$ is added to the subspace in eq 26, and the iteration procedure is continued until convergence is obtained. The iterative scheme is converged when the residual Euclidean norm $\|\boldsymbol{R}\|$ is smaller than a preset threshold.

3.2. Response Eigenvalue Equation. In this section we describe how new trial vectors are generated when solving the response eigenvalue equation, using the Davidson[22] and the Olsen[26] algorithms. Both these algorithms were originally designed for solving Hermitian eigenvalue equations. We discuss the problems that may occur when the Davidson and the Olsen algorithms are applied to the non-Hermitian response eigenvalue equation.

3.2.1. Davidson Algorithm. In the Davidson algorithm a new trial vector is obtained as for the preconditioned standard and damped response equations, according to eq 37, where the matrix $\boldsymbol{L}$ is a diagonal approximation to $(\boldsymbol{E}^{[2]} - \omega_f^R\mathbf{S}^{[2]})$

$$\mathbf{b}_{n+1} = (\mathbf{E}_0^{[2]} - \omega_f^R\mathbf{S}^{[2]})^{-1}\boldsymbol{R}_{n+1} \tag{38}$$

For the linear equation improved convergence is obtained when $\mathbf{E}_0^{[2]}$ becomes an improved approximation to the $\mathbf{E}^{[2]}$ matrix. In fact, when $\mathbf{E}_0^{[2]}$ is replaced by $\mathbf{E}^{[2]}$, the converged solution is obtained right away for a set of linear equations. However, when the eigenvalue equation is solved using the Davidson algorithm, a problem arises in the limit where $\mathbf{E}_0^{[2]}$ is approaching $\mathbf{E}^{[2]}$ as no new direction is generated. To see this, we replace $\mathbf{E}_0^{[2]}$ with $\mathbf{E}^{[2]}$ in eq 38 and introduce the definition of the residual in eq 36, giving

$$\begin{aligned}\mathbf{b}_{n+1} &= (\mathbf{E}^{[2]} - \omega_f^{\mathrm{R}}\mathbf{S}^{[2]})^{-1}(\mathbf{E}^{[2]} - \omega_f^{\mathrm{R}}\mathbf{S}^{[2]})(\mathbf{X}_f)_{n+1} \\ &= (\mathbf{X}_f)_{n+1}\end{aligned} \tag{39}$$

The proposed new trial vector is therefore equal to the subspace solution to eq 30, so in effect, no new trial vector is obtained using the Davidson algorithm.[35] This may be cumbersome when improved preconditioners are used, and a remedy to this problem is provided in the Olsen algorithm.[26]

3.2.2. Olsen Algorithm. In the Olsen algorithm, it is assumed that we know an approximate solution to the response eigenvalue equation. This may be the optimal solution vector $(\mathbf{X}_f)_{n+1}$ of eq 33, for convenience denoted as $\mathbf{X}^0$, which satisfies the normalization condition in eq 17. The zeroth order eigenvalue ω_f^0 associated with $\mathbf{X}^0$ is obtained by projecting the eigenvalue equation in eq 3 with $(\mathbf{X}^0)^T$

$$(\mathbf{X}^0)^T(\mathbf{E}^{[2]} - \omega_f^0\mathbf{S}^{[2]})\mathbf{X}^0 = 0 \tag{40}$$

Note that ω_f^0 obtained from eq 40 is identical to ω_f^{R} in eq 30.

To get an improved solution vector in the Olsen algorithm, we express eq 3 in terms of the zeroth order and the correction components:

$$\omega_f = \omega_f^0 + \omega_f^1 \tag{41}$$

$$\mathbf{X}_f = \mathbf{X}^0 + \mathbf{X}^1 \tag{42}$$

where ω_f^1 and $\mathbf{X}^1$ are correction terms to the eigenvalue ω_f^0 and the eigenvector $\mathbf{X}^0$, respectively. The $\mathbf{E}^{[2]}$ matrix may also be written in terms of a zeroth order and a correction component as in eq 11. Inserting eqs 11, 41, and 42 into eq 3 gives

$$[(\mathbf{E}_0^{[2]} + \mathbf{E}_1^{[2]}) - (\omega_f^0 + \omega_f^1)\mathbf{S}^{[2]}](\mathbf{X}^0 + \mathbf{X}^1) = 0 \tag{43}$$

Neglecting terms that are quadratic in the corrections, we get

$$(\mathbf{E}_0^{[2]} - \omega_f^0\mathbf{S}^{[2]})\mathbf{X}^1 = -(\mathbf{E}^{[2]} - \omega_f^0\mathbf{S}^{[2]})\mathbf{X}^0 + \omega_f^1\mathbf{S}^{[2]}\mathbf{X}^0 \tag{44}$$

and $\mathbf{X}^1$ can be determined as

$$\mathbf{X}^1 = -(\mathbf{E}_0^{[2]} - \omega_f^0\mathbf{S}^{[2]})^{-1}[(\mathbf{E}^{[2]} - \omega_f^0\mathbf{S}^{[2]})\mathbf{X}^0 - \omega_f^1\mathbf{S}^{[2]}\mathbf{X}^0] \tag{45}$$

By requiring that the eigenvector correction $\mathbf{X}^1$ is orthogonal to $\mathbf{X}^0$ in the generalized metric $\mathbf{S}^{[2]}$

$$(\mathbf{X}^1)^T\mathbf{S}^{[2]}\mathbf{X}^0 = 0 \tag{46}$$

we may determine ω_f^1 by multiplying eq 45 with $(\mathbf{X}^0)^T\mathbf{S}^{[2]}$

$$\omega_f^1 = \frac{(\mathbf{X}^0)^T\mathbf{S}^{[2]}(\mathbf{E}_0^{[2]} - \omega_f^0\mathbf{S}^{[2]})^{-1}(\mathbf{E}^{[2]} - \omega_f^0\mathbf{S}^{[2]})\mathbf{X}^0}{(\mathbf{X}^0)^T\mathbf{S}^{[2]}(\mathbf{E}_0^{[2]} - \omega_f^0\mathbf{S}^{[2]})^{-1}\mathbf{S}^{[2]}\mathbf{X}^0} \tag{47}$$

The inverse matrix $(\mathbf{E}_0^{[2]} - \omega_f^0\mathbf{S}^{[2]})^{-1}$ is readily constructed as both $\mathbf{E}_0^{[2]}$ and $\mathbf{S}^{[2]}$, which are diagonal matrices. Once the linear transformations $\mathbf{E}^{[2]}\mathbf{X}^0$ and $\mathbf{S}^{[2]}\mathbf{X}^0$ are known, we may determine ω_f^1 from eq 47, and then the correction vector $\mathbf{X}^1$ may be obtained from eq 45. The new trial vector $\mathbf{b}_{n+1} = \mathbf{X}^1$ is added to the trial vector subspace in eq 26, and the iteration procedure is continued until convergence.

Note that the first term in eq 45 gives the Davidson correction in eq 38. The second term in eq 45 thus ensures that a new improved trial vector is obtained also when $\mathbf{E}_0^{[2]}$ is approaching $\mathbf{E}^{[2]}$. In fact, when $\mathbf{E}_0^{[2]}$ in the Olsen algorithm is replaced by $\mathbf{E}^{[2]}$, we establish the inverse-iteration method with the Rayleigh quotient.[16,35,36] Note that new trial vectors are obtained using the Olsen algorithm without carrying out new linear transformations compared to the Davidson algorithm.

3.2.3. Problems When Solving the Response Eigenvalue Equation. Due to the fact that $\mathbf{S}^{[2]}$ is an easily invertible matrix $[(\mathbf{S}^{[2]})^{-1} = \mathbf{S}^{[2]}]$, the response eigenvalue equation may be expressed as

$$\mathbf{S}^{[2]}\mathbf{E}^{[2]}\mathbf{X} = \omega\mathbf{X} \tag{48}$$

eq 48 was used in early iterative subspace algorithms to obtain the solution to the response eigenvalue equation.[24] However, $(\mathbf{S}^{[2]}\mathbf{E}^{[2]})$ is a non-Hermitian matrix, and eq 48 therefore represents a non-Hermitian eigenvalue equation. When a non-Hermitian rather than

a standard Hermitian eigenvalue equation is solved using an iterative subspace algorithm, some difficulties may occur. For a Hermitian eigenvalue equation, the subspace eigenvalue equation is guaranteed to have real eigenvalues. Also monotonic convergence is obtained to the lowest eigenvalues due to MacDonald's theorem.[37] In contrast, when a non-Hermitian eigenvalue equation is solved the eigenvalues of the reduced generalized eigenvalue equation may be complex. Then new trial vectors cannot be obtained neither using the Davidson nor the Olsen algorithms because, in both cases, it is assumed that ω_f^{R} is real. Bouman et al.[24] encountered this problem when the Davidson algorithm was applied to eq 48. Also the monotonic convergence to the lowest eigenvalues is lost.

In eq 24, it is shown that for the ground state, the non-Hermitian response eigenvalue equation may be transformed into a Hermitian eigenvalue equation. In Section 4, we discuss how this Hermitian eigenvalue equation may be solved using a generalization of the Davidson iterative subspace algorithm, by imposing the paired structure of the response matrices $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ on their subspace counterparts.

3.3. Standard Response Equation. The standard response equation in eq 1 represents a set of linear equations with a symmetric matrix. For angular frequencies ω that are smaller than the lowest excitation energy, the response matrix $(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})$ is also positive definite. In Appendix A, it is described how, for a positive definite symmetric matrix, the subspace iterative algorithm described in Section 3.1 and the CG algorithm[15−18] lead to the same iteration sequence.[14] This is obtained because the residuals in the subspace algorithm and the optimal directions in the CG method span the same space. The CG algorithm is designed such that the set of vectors stored on disk may be truncated to the three vectors, so the storing and the manipulation of the large amount of directions (trial vectors) that is required to set up the reduced space equations in the subspace algorithm can be avoided. For angular frequencies ω larger than the lowest excitation energy, $(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})$ is not positive definite but still symmetric, and the CR algorithm may, in principle, be applied. To improve convergence the CG and CR algorithms are used with the diagonal preconditioner $(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})$. However, this preconditioner is positive definite only for ω smaller than the HOMO−LUMO gap, and the preconditioned CR algorithm can therefore be safely used to solve the standard response equation for ω smaller than the HOMO−LUMO gap (see Appendix B). For larger ω the preconditioned iterative subspace algorithm is the best option.

In the absence of a preconditioner, the convergence rate for an iterative solution of the standard response equation is determined by the condition number of the response matrix $(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})$, i.e., the ratio between the smallest and largest eigenvalues of the response matrix $\omega_f^{\max}/\omega_f^{\min}$. This condition number is typically large since there are very large eigenvalues in the Hessian referring to, e.g., core excitations as well as nonphysical excitations into the continuum energy spectrum that appear as discrete levels due to the use of a localized atomic orbital basis. However, by using $(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})$ as preconditioner, the convergence rate is instead determined by the condition number of the matrix:

$$\begin{aligned}(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})^{-1}(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]}) &= (\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})^{-1}[(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]}) + \mathbf{E}_1^{[2]}] \\ &= 1 + (\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})^{-1}\mathbf{E}_1^{[2]} \end{aligned} \quad (49)$$

and the condition number is therefore greatly reduced since all the large eigenvalues discussed above (e.g., as due to core excitations) are scaled more or less to unity by the preconditioner. We note, however, that, when the optical frequency ω is in close resonance with excitation energies of the system or rather with orbital energy differences, the matrix inverse $(\mathbf{E}_0^{[2]} - \omega\mathbf{S}^{[2]})^{-1}$ becomes nearly singular, and the preconditioning will give rise to an increase in the condition number for these near resonance excitation energies. This will result in a slower convergence rate as compared to the nonresonant case. It is also clear that the higher the density of states in the region of the optical frequency, the worse the convergence rate we expect, simply since there are a larger number of eigenvalues that will be scaled poorly by the preconditioner.

3.4. Damped Response Equation. The damped response equation of the form in eq 2 represents a set of linear equations for a non-Hermitian complex matrix. We will now describe how eq 2 may be transformed to a set of linear equations for a real symmetric matrix, thereby avoiding complex algebra.

The solution to eq 2 may be expressed in terms of the solution for the real and the imaginary component of eq 2. This results in two coupled real equations

$$(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})\mathbf{X}_{\mathrm{D}}^{\mathrm{R}} = \mathbf{G}^{\mathrm{R}} - \gamma\mathbf{S}^{[2]}\mathbf{X}_{\mathrm{D}}^{\mathrm{I}} \quad (50)$$

$$(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})\mathbf{X}_{\mathrm{D}}^{\mathrm{I}} = \mathbf{G}^{\mathrm{I}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_{\mathrm{D}}^{\mathrm{R}} \quad (51)$$

where $\mathbf{G}^{\mathrm{R}}$ and $\mathbf{G}^{\mathrm{I}}$ is the real and the imaginary component of the gradient vector $\mathbf{G}$, respectively. When eqs 50 and 51 are solved separately without explicit coupling between the two equations, this may lead to divergence in the resonance regions where the solution vector has a large eigenvector component in both $\mathbf{X}_{\mathrm{D}}^{\mathrm{R}}$ and $\mathbf{X}_{\mathrm{D}}^{\mathrm{I}}$. The coupling between $\mathbf{X}_{\mathrm{D}}^{\mathrm{R}}$ and $\mathbf{X}_{\mathrm{D}}^{\mathrm{I}}$ can be considered explicitly by expressing eqs 50 and 51 in the matrix form:

$$\begin{pmatrix} \mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]} & \gamma\mathbf{S}^{[2]} \\ -\gamma\mathbf{S}^{[2]} & \mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\mathrm{D}}^{\mathrm{R}} \\ \mathbf{X}_{\mathrm{D}}^{\mathrm{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{G}^{\mathrm{R}} \\ \mathbf{G}^{\mathrm{I}} \end{pmatrix} \quad (52)$$

that represents a set of linear equations for a nonsymmetric matrix. However, since it is advantageous to solve the set of linear equation for a symmetric matrix, by reversing the sign of the second row, we express eq 52 as

$$\begin{pmatrix} \mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]} & \gamma\mathbf{S}^{[2]} \\ \gamma\mathbf{S}^{[2]} & -(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]}) \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\mathrm{D}}^{\mathrm{R}} \\ \mathbf{X}_{\mathrm{D}}^{\mathrm{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{G}^{\mathrm{R}} \\ -\mathbf{G}^{\mathrm{I}} \end{pmatrix} \quad (53)$$

which represents a standard set of linear equations for a symmetric, indefinite, and real matrix.

The residuals of eq 53 are given by

$$\boldsymbol{R}_{n+1}^{\mathrm{R}} = (\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})(\mathbf{X}_{\mathrm{D}}^{\mathrm{R}})_{n+1} - \mathbf{G}^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}(\mathbf{X}_{\mathrm{D}}^{\mathrm{I}})_{n+1} \quad (54)$$

$$\boldsymbol{R}_{n+1}^{\mathrm{I}} = -(\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})(\mathbf{X}_{\mathrm{D}}^{\mathrm{I}})_{n+1} + \mathbf{G}^{\mathrm{I}} + \gamma\mathbf{S}^{[2]}(\mathbf{X}_{\mathrm{D}}^{\mathrm{R}})_{n+1} \quad (55)$$

and may be used to check for convergence. In the preconditioned iterative subspace algorithm, new trial vectors are obtained by

preconditioning the residuals:

$$\begin{pmatrix} \mathbf{b}^{\mathrm{R}}_{n+1} \\ \mathbf{b}^{\mathrm{I}}_{n+1} \end{pmatrix} = [(\mathbf{E}^{[2]}_0 - \omega\mathbf{S}^{[2]})^2 + \gamma^2 1]^{-1} \otimes \begin{pmatrix} \mathbf{E}^{[2]}_0 - \omega\mathbf{S}^{[2]} & \gamma\mathbf{S}^{[2]} \\ \gamma\mathbf{S}^{[2]} & -(\mathbf{E}^{[2]}_0 - \omega\mathbf{S}^{[2]}) \end{pmatrix} \begin{pmatrix} \boldsymbol{R}^{\mathrm{R}}_{n+1} \\ \boldsymbol{R}^{\mathrm{I}}_{n+1} \end{pmatrix} \tag{56}$$

where we have introduced a direct matrix product according to

$$\mathbf{c} = \mathbf{a} \otimes \begin{pmatrix} \mathbf{b}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{21} & \mathbf{b}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{ab}_{11} & \mathbf{ab}_{12} \\ \mathbf{ab}_{21} & \mathbf{ab}_{22} \end{pmatrix} \tag{57}$$

It can be seen by inspection that the inverse preconditioner in eq 56 is the inverse matrix to the matrix in eq 53, where $\mathbf{E}^{[2]}$ is replaced by $\mathbf{E}^{[2]}_0$.

Since eq 53 represents a set of linear equations for a symmetric, indefinite matrix, it cannot be solved using the CG algorithm, but the CR algorithm[19,38] may be applied. However, the preconditioner in eq 56 is not positive definite, and therefore the preconditioned CR algorithm cannot be safely used (see Appendix B). When solving eq 53 using preconditioning, the iterative subspace algorithm therefore becomes the best option.

The condition number is significantly reduced by preconditioning for similar reasons as presented in Section 3.3 for the standard response equation, as the dominant contributions to the large eigenvalues are removed. The condition number will increase when $[(\mathbf{E}^{[2]}_0 - \omega\mathbf{S}^{[2]})^2 + \gamma^2 1]^{-1}$ is approaching a singularity, and a slower convergence will be observed. Since the γ parameter is small, the condition number of the preconditioned standard response equation and the preconditioned damped response equation will be similar, and similar convergence will be observed for the preconditioned standard and damped response equations.

The preconditioned iterative subspace approach presented in this section, represents a simple and straightforward scheme for solving the damped response equation, using a real trial vector space. In Section 5.3, we show that introducing symmetrized trial vectors leads to an improved efficiency with respect to solving the damped response equation.

4. ITERATIVE METHODS WITH PAIRED TRIAL VECTORS

In the previous section, we discussed how iterative algorithms may be used to solve the response equations, where the response matrices $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ were considered to be standard symmetric matrices. As discussed in Section 2, the response matrices have a 2×2 matrix block structure (see eq 4), which imposes that the response eigenvalue equation has paired eigensolutions (see eqs 17−21). When the response equations are solved using the general subspace iterative approach described in Section 3.1, this paired structure is lost in the reduced space equations. However, preserving the 2×2 block structure in the reduced space equations in eqs 28−30, leads to improvement in convergence.[25] This paired structure may be recovered in the reduced space by adding trial vectors in the iterative subspace algorithm in pairs[25] as will be described below.

In paired iterative algorithms, together with a trial vector $\mathbf{b}$

$$\mathbf{b} = \begin{pmatrix} \mathbf{b}_{AI} \\ \mathbf{b}_{JB} \end{pmatrix} \tag{58}$$

the paired vector $\mathbf{b}^{\mathrm{P}}$

$$\mathbf{b}^{\mathrm{P}} = \begin{pmatrix} \mathbf{b}_{JB} \\ \mathbf{b}_{AI} \end{pmatrix} \tag{59}$$

is added to the subspace in each iteration of the iterative procedure. From the linear transformations of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ on $\mathbf{b}$ in eq 25, the linear transformations on the corresponding paired trial vector can be obtained without significant computational cost, as

$$\mathbf{E}^{[2]}\mathbf{b}^{\mathrm{P}}_i = \boldsymbol{\sigma}^{\mathrm{P}}_i, \quad \mathbf{S}^{[2]}\mathbf{b}^{\mathrm{P}}_i = -\boldsymbol{\rho}^{\mathrm{P}}_i \tag{60}$$

where we have used eqs 4, 58, and 59.

After n iterations in the paired iterative subspace algorithm, the reduced space $\mathbf{b}^{2n}$ consists of $2n$ trial vectors:

$$\mathbf{b}^{2n} = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n, \mathbf{b}^{\mathrm{P}}_1, \mathbf{b}^{\mathrm{P}}_2, ..., \mathbf{b}^{\mathrm{P}}_n\} \tag{61}$$

and the associated spaces of linear transformed vectors $\boldsymbol{\sigma}^{2n}$ and $\boldsymbol{\rho}^{2n}$ are also known. The reduced matrix equations of the form depicted in eq 28−30 may now be set up, and the matrix elements of the reduced space generalized Hessian and metric matrices are given as

$$(\mathbf{E}^{[2]}_{\mathrm{R}})_{ij} = (\mathbf{b}^{2n}_i)^T\boldsymbol{\sigma}^{2n}_j, \quad (\mathbf{S}^{[2]}_{\mathrm{R}})_{ij} = (\mathbf{b}^{2n}_i)^T\boldsymbol{\rho}^{2n}_j \tag{62}$$

respectively. In matrix block form $\mathbf{E}^{[2]}_{\mathrm{R}}$ and $\mathbf{S}^{[2]}_{\mathrm{R}}$ may be expressed as

$$\mathbf{E}^{[2]}_{\mathrm{R}} = \begin{pmatrix} \mathbf{A}_{\mathrm{R}} & \mathbf{B}_{\mathrm{R}} \\ \mathbf{B}_{\mathrm{R}} & \mathbf{A}_{\mathrm{R}} \end{pmatrix}, \quad \mathbf{S}^{[2]}_{\mathrm{R}} = \begin{pmatrix} \Sigma_{\mathrm{R}} & \Delta_{\mathrm{R}} \\ -\Delta_{\mathrm{R}} & -\Sigma_{\mathrm{R}} \end{pmatrix} \tag{63}$$

where $\mathbf{A}_{\mathrm{R}}$, $\mathbf{B}_{\mathrm{R}}$, and Σ_{R} are symmetric, and Δ_{R} is an antisymmetric matrix of the dimension n. The reduced subspace matrices thus have the same paired structure as their full matrix counterparts in eq 4. The reduced space right-hand sides in eqs 28 and 29 are given by

$$(\mathbf{G}_{\mathrm{R}})_i = (\mathbf{b}^{2n}_i)^T\mathbf{G} \tag{64}$$

Solving the reduced response equations (eqs 28−30) yields a solution vector $(\mathbf{X})_{n+1}$ in the subspace according to

$$(\mathbf{X})_{n+1} = \sum_{i=1}^{2n} (\mathbf{X}_{\mathrm{R}})_i \mathbf{b}^{2n}_i \tag{65}$$

The residuals for $(\mathbf{X})_{n+1}$ become equal to

$$\begin{aligned} \boldsymbol{R}_{n+1} &= (\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})(\mathbf{X}_{\mathrm{S}})_{n+1} - \mathbf{G} \\ &= \sum_{i=1}^{2n} [(\mathbf{X}_{\mathrm{S}})_{\mathrm{R}}]_i(\boldsymbol{\sigma}^{2n}_i - \omega\boldsymbol{\rho}^{2n}_i) - \mathbf{G} \end{aligned} \tag{66}$$

$$\begin{aligned} \boldsymbol{R}_{n+1} &= (\mathbf{E}^{[2]} - \omega^{\mathrm{R}}_f\mathbf{S}^{[2]})(\mathbf{X}_f)_{n+1} \\ &= \sum_{i=1}^{2n} [(\mathbf{X}_f)_{\mathrm{R}}]_i(\boldsymbol{\sigma}^{2n}_i - \omega^{\mathrm{R}}_f\boldsymbol{\rho}^{2n}_i) \end{aligned} \tag{67}$$

for the standard and eigenvalue response equation, respectively. For the damped response equations the real and imaginary

components of the residual are given as

$$\boldsymbol{R}^{\mathrm{R}}_{n+1} = (\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})(\mathbf{X}^{\mathrm{R}}_{\mathrm{D}})_{n+1} - \mathbf{G}^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}(\mathbf{X}^{\mathrm{I}}_{\mathrm{D}})_{n+1} = \sum_{i=1}^{4n} [(\mathbf{X}^{\mathrm{R}}_{\mathrm{D}})_{\mathrm{R}}]_i(\boldsymbol{\sigma}^{4n}_i - \omega\boldsymbol{\rho}^{4n}_i) - \mathbf{G}^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}(\mathbf{X}^{\mathrm{I}}_{\mathrm{D}})_{n+1} \quad (68)$$

and

$$\boldsymbol{R}^{\mathrm{I}}_{n+1} = (\mathbf{E}^{[2]} - \omega\mathbf{S}^{[2]})(\mathbf{X}^{\mathrm{I}}_{\mathrm{D}})_{n+1} - \mathbf{G}^{\mathrm{I}} - \gamma\mathbf{S}^{[2]}(\mathbf{X}^{\mathrm{R}}_{\mathrm{D}})_{n+1} = \sum_{i=1}^{4n} [(\mathbf{X}^{\mathrm{I}}_{\mathrm{D}})_{\mathrm{R}}]_i(\boldsymbol{\sigma}^{4n}_i - \omega\boldsymbol{\rho}^{4n}_i) - \mathbf{G}^{\mathrm{I}} - \gamma\mathbf{S}^{[2]}(\mathbf{X}^{\mathrm{R}}_{\mathrm{D}})_{n+1} \quad (69)$$

respectively. When the damped response equation is solved, the reduced subspace has a dimension $4n$, as it contains trial vectors for both the real and imaginary components of the solution vector together with their paired counterparts.[3,5]

The residual norm is calculated to check for convergence. A new trial vector $\mathbf{b}_{n+1}$ is obtained from the preconditioned residuals according to eq 37. The new trial vector, together with its paired counterpart $\mathbf{b}^{\mathrm{P}}_{n+1}$, are added to the reduced subspace in eq 61, and the iterative procedure is continued until convergence.

When the response eigenvalue equation in eq 24 is solved using the algorithm with paired trial vectors, the reduced subspace matrices thus have the 2×2 block structure similar to their exact counterparts. For a ground-state response calculation, $\mathbf{E}^{[2]}_{\mathrm{R}}$ is also positive definite, and the reduced eigenvalue equations will therefore have paired real eigenvalues with positive eigenvalues for the positive $\mathbf{S}^{[2]}$-normed eigensolutions. It was stated in ref 25 that the lowest positive eigenvalue will converge monotonically toward the full space lowest positive eigenvalue,[37] e.g., the lowest positive eigenvalue of iterations n and $n+1$, satisfies $(\omega^{\mathrm{R}}_1)_n \geq (\omega^{\mathrm{R}}_1)_{n+1}$. This statement is proven in the present work in Appendix C. When the response eigenvalue equation is solved using the subspace algorithm with paired trial vectors, the Davidson algorithm may straightforwardly be applied as for a Hermitian eigenvalue equation, and monotonic convergence will be obtained toward the full space lowest eigenvalue.

For the standard response equation, Olsen et al.[25] demonstrated an improved convergence using an iterative subspace algorithm with paired trial vectors. When the standard response equation (eq 1) is solved without pairing, an identical iteration sequence is obtained for ω smaller than the HOMO−LUMO gap regardless of whether the iterative subspace or the CG algorithm is used (see Appendix A). But when paired trial vectors are employed, the reduction in the number of subspace vectors is not obtained for the CG algorithm. We shall see in Section 5.2, however, that this reduction is retained in the CG algorithm when symmetrized trial vectors are used.

5. ITERATIVE METHODS WITH SYMMETRIZED TRIAL VECTORS

In the previous section, we discussed advantages of using paired trial vectors to solve response equations. In this section, we consider the advantages that may be obtained by splitting trial vectors into symmetric and antisymmetric components.

A real symmetric (g) and antisymmetric (u) trial vector may be written as

$$\mathbf{b}_g = \begin{pmatrix} \mathbf{b}_{IA} \\ \mathbf{b}_{IA} \end{pmatrix} \quad (70)$$

and

$$\mathbf{b}_u = \begin{pmatrix} \mathbf{b}_{IA} \\ -\mathbf{b}_{IA} \end{pmatrix} \quad (71)$$

respectively.

A set of paired trial vectors, as given in eqs 58 and 59, may be represented by the symmetric and antisymmetric components of the vector $\mathbf{b}$ according to

$$\mathbf{b}_g = \frac{1}{2}(\mathbf{b} + \mathbf{b}^{\mathrm{P}}) = \frac{1}{2}\begin{pmatrix} \mathbf{b}_{AI} + \mathbf{b}_{JB} \\ \mathbf{b}_{JB} + \mathbf{b}_{AI} \end{pmatrix} \quad (72)$$

$$\mathbf{b}_u = \frac{1}{2}(\mathbf{b} - \mathbf{b}^{\mathrm{P}}) = \frac{1}{2}\begin{pmatrix} \mathbf{b}_{AI} - \mathbf{b}_{JB} \\ \mathbf{b}_{JB} - \mathbf{b}_{AI} \end{pmatrix} \quad (73)$$

Adding always one symmetric and one antisymmetric trial vector in a subspace algorithm is thus equivalent to adding a set of paired vectors, and it will ensure an implicit paired structure of the reduced space equations.

The symmetry of a trial vector is conserved for the linear transformation with respect to $\mathbf{E}^{[2]}$:

$$\boldsymbol{\sigma}_g = \mathbf{E}^{[2]}\mathbf{b}_g, \quad \boldsymbol{\sigma}_u = \mathbf{E}^{[2]}\mathbf{b}_u \quad (74)$$

and reversed for the linear transformation with respect to $\mathbf{S}^{[2]}$:

$$\boldsymbol{\rho}_g = \mathbf{S}^{[2]}\mathbf{b}_u, \quad \boldsymbol{\rho}_u = \mathbf{S}^{[2]}\mathbf{b}_g \quad (75)$$

From the linear transformation of $\mathbf{E}^{[2]}$ on the sum of one symmetric and one antisymmetric trial vector ($\boldsymbol{b}_g$ and $\boldsymbol{b}_u$ of eqs 70 and 71), we may determine the linear transformations on the individual components in eq 74 as the linear transformation conserves the symmetry of a trial vector. Similar arguments hold for linear transformation involving the $\mathbf{S}^{[2]}$ matrix. The computational cost of the linear transformation on a general trial vector and on its symmetric and antisymmetric components is thus identical.

A subspace representation of the $\mathbf{E}^{[2]}$ matrix for a set of symmetric vectors $\mathbf{b}_g$ gives a subspace representation for the $\mathbf{A}+\mathbf{B}$ matrix

$$\mathbf{b}^T_{ig}\mathbf{E}^{[2]}\mathbf{b}_{jg} = \mathbf{b}^T_{ig}\boldsymbol{\sigma}_{jg} = 2(\mathbf{A}+\mathbf{B})^{\mathrm{R}}_{ij} \quad (76)$$

where $\mathbf{b}_{ig}$ is the i'th element of the vector subspace

$$\mathbf{b}^n_g = \{\mathbf{b}_{1g}, \mathbf{b}_{2g}, ..., \mathbf{b}_{ng}\} \quad (77)$$

and $\boldsymbol{\sigma}_{jg}$ is the j'th element of the $\mathbf{E}^{[2]}$ linear transformed vector subspace

$$\boldsymbol{\sigma}^n_g = \{\boldsymbol{\sigma}_{1g}, \boldsymbol{\sigma}_{2g}, ..., \boldsymbol{\sigma}_{ng}\} \quad (78)$$

A similar subspace representation of the $\mathbf{E}^{[2]}$ matrix for the antisymmetric $\mathbf{b}_u$ vectors gives a subspace representation of the $\mathbf{A}-\mathbf{B}$ matrix.

The symmetries of the $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ matrices in eqs 74 and 75 make it advantageous to split the solution to the response

equations into symmetric and antisymmetric components. For the standard response equation and the eigenvalue equation, we express the solution vectors as

$$\mathbf{X}_S = \mathbf{X}_g + \mathbf{X}_u \tag{79}$$

$$\mathbf{X}_f = \mathbf{X}_{g,f} + \mathbf{X}_{u,f} \tag{80}$$

respectively. The solution to the damped response equation is written in terms of the Hermitian and anti-Hermitian components

$$\mathbf{X}_D = \mathbf{X}^H + \mathbf{X}^A \tag{81}$$

where

$$\mathbf{X}^H = \mathbf{X}_g^R + i\mathbf{X}_u^I \tag{82}$$

$$\mathbf{X}^A = \mathbf{X}_u^R + i\mathbf{X}_g^I \tag{83}$$

and $\mathbf{X}_g^R$, $\mathbf{X}_g^I$, $\mathbf{X}_u^R$, and $\mathbf{X}_u^I$ are real. Introducing $\mathbf{X}_S$, $\mathbf{X}_f$, and $\mathbf{X}_D$ of eqs 79−81 in the response equations (in eqs 1−3) simplifies the solving of these equations, as will be demonstrated below.

5.1. Response Eigenvalue Equation. Employing the symmetries of the $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ matrices and expressing the solution vector as in eq 80, the response eigenvalue equation in eq 3 may be written in matrix blocked form, according to

$$\begin{pmatrix} \mathbf{E}^{[2]} & -\omega_f \mathbf{S}^{[2]} \\ -\omega_f \mathbf{S}^{[2]} & \mathbf{E}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{g,f} \\ \mathbf{X}_{u,f} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{84}$$

where the coupling that occurs between the two components $\mathbf{X}_{g,f}$ and $\mathbf{X}_{u,f}$ is introduced explicitly. Equation 84 may be written in the form of a generalized, non-Hermitian, eigenvalue equation:

$$\begin{pmatrix} \mathbf{E}^{[2]} & 0 \\ 0 & \mathbf{E}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{g,f} \\ \mathbf{X}_{u,f} \end{pmatrix} = \omega_f \begin{pmatrix} 0 & \mathbf{S}^{[2]} \\ \mathbf{S}^{[2]} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X}_{g,f} \\ \mathbf{X}_{u,f} \end{pmatrix} \tag{85}$$

where the generalized indefinite overlap matrix is viewed as metric. If instead the generalized, positive definite Hessian is viewed as metric, then eq 85 represents a standard eigenvalue equation for a Hermitian matrix, with eigenvalues $(\omega_f)^{-1}$, analogous to eq 24.

Equation 84 may be solved using an iterative subspace algorithm where in each iteration a symmetric and an antisymmetric trial vector is added. In iteration n, we therefore have two sets of trial vectors:

$$\mathbf{b}_g^n = \{\mathbf{b}_{1g}, \mathbf{b}_{2g}, ..., \mathbf{b}_{ng}\} \tag{86}$$

$$\mathbf{b}_u^n = \{\mathbf{b}_{1u}, \mathbf{b}_{2u}, ..., \mathbf{b}_{nu}\} \tag{87}$$

and the corresponding sets of linear transformed vectors $\boldsymbol{\sigma}_g^n$, $\boldsymbol{\sigma}_u^n$ $\boldsymbol{\rho}_u^n$, and $\boldsymbol{\rho}_g^n$; the i^{th} vector in these latter sets is referred to as $\boldsymbol{\sigma}_{ig}$, $\boldsymbol{\sigma}_{iu}$, $\boldsymbol{\rho}_{iu}$, and $\boldsymbol{\rho}_{ig}$, respectively. Equation 84 is solved in the subspace given by eqs 86 and 87 yielding a reduced space equation:

$$\begin{pmatrix} \mathbf{E}_{R,gg}^{[2]} & -\omega_f^R \mathbf{S}_{R,gu}^{[2]} \\ -\omega_f^R \mathbf{S}_{R,ug}^{[2]} & \mathbf{E}_{R,uu}^{[2]} \end{pmatrix} \begin{pmatrix} (\mathbf{X}_{g,f})_R \\ (\mathbf{X}_{u,f})_R \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{88}$$

where

$$\begin{aligned} \mathbf{E}_{R,gg}^{[2]} &= \mathbf{b}_g^T \mathbf{E}^{[2]} \mathbf{b}_g, \quad \mathbf{E}_{R,uu}^{[2]} = \mathbf{b}_u^T \mathbf{E}^{[2]} \mathbf{b}_u, \\ \mathbf{S}_{R,gu}^{[2]} &= \mathbf{b}_g^T \mathbf{S}^{[2]} \mathbf{b}_u, \quad \mathbf{S}_{R,ug}^{[2]} = \mathbf{b}_u^T \mathbf{S}^{[2]} \mathbf{b}_g \end{aligned} \tag{89}$$

The optimal vectors in the subspace in eqs 86 and 87 have a form

$$\mathbf{X}_{n+1,g,f} = \sum_{i=1}^{n} [(\mathbf{X}_{g,f})_R]_i \mathbf{b}_{ig} \tag{90}$$

$$\mathbf{X}_{n+1,u,f} = \sum_{i=1}^{n} [(\mathbf{X}_{u,f})_R]_i \mathbf{b}_{iu} \tag{91}$$

From the solution vectors, the residuals are obtained

$$\begin{aligned} \boldsymbol{R}_{n+1,g} &= \mathbf{E}^{[2]} \mathbf{X}_{n+1,g,f} - \omega_f^R \mathbf{S}^{[2]} \mathbf{X}_{n+1,u,f} \\ &= \sum_{i=1}^{n} [(\mathbf{X}_{g,f})_R]_i \boldsymbol{\sigma}_{ig} - \omega_f^R \sum_{i=1}^{n} [(\mathbf{X}_{u,f})_R]_i \boldsymbol{\rho}_{ig} \end{aligned} \tag{92}$$

$$\begin{aligned} \boldsymbol{R}_{n+1,u} &= \mathbf{E}^{[2]} \mathbf{X}_{n+1,u,f} - \omega_f^R \mathbf{S}^{[2]} \mathbf{X}_{n+1,g,f} \\ &= \sum_{i=1}^{n} [(X_{u,f})_R]_i \boldsymbol{\sigma}_{iu} - \omega_f^R \sum_{i=1}^{n} [(X_{g,f})_R]_i \boldsymbol{\rho}_{iu} \end{aligned} \tag{93}$$

that are used to check for convergence.

New trial vectors $\mathbf{b}_{n+1,g}$ and $\mathbf{b}_{n+1,u}$ may be obtained from the optimal solution vectors $\mathbf{X}_{n+1,g}$ and $\mathbf{X}_{n+1,u}$ using either the Davidson or the Olsen algorithm, discussed in Sections 3.2.1 and 3.2.2, respectively. As discussed in Section 3.2.2, the Olsen algorithm is used to resolve the problem that arises when $\mathbf{E}_0^{[2]}$ is approaching $\mathbf{E}^{[2]}$. When the response eigenvalue equations are solved, $\mathbf{E}_0^{[2]}$ is about equally good approximation to $\mathbf{E}^{[2]}$ in the whole frequency range, and the Davidson algorithm can therefore safely be used to generate new trial vectors. In the Davidson algorithm, new trial vectors are constructed from eq 38, which by writing the residuals and trial vectors in terms of a symmetric and an antisymmetric component may be expressed as

$$(\mathbf{b}_{n+1,g} + \mathbf{b}_{n+1,u}) = (\mathbf{E}_0^{[2]} - \omega_f^R \mathbf{S}^{[2]})^{-1} (\boldsymbol{R}_{n+1,g} + \boldsymbol{R}_{n+1,u}) \tag{94}$$

To obtain trial vectors that are symmetric and antisymmetric, we use that

$$\begin{aligned} (\mathbf{E}_0^{[2]} - \omega_f^R \mathbf{S}^{[2]})^{-1} &= (\mathbf{E}_0^{[2]} - \omega_f^R \mathbf{S}^{[2]})^{-1} (\mathbf{E}_0^{[2]} + \omega_f^R \mathbf{S}^{[2]})^{-1} (\mathbf{E}_0^{[2]} + \omega_f^R \mathbf{S}^{[2]}) \\ &= [(\mathbf{E}_0^{[2]})^2 - (\omega_f^R)^2 1]^{-1} (\mathbf{E}_0^{[2]} + \omega_f^R \mathbf{S}^{[2]}) \end{aligned} \tag{95}$$

Equation 94 may now be expressed in the matrix blocked form according to

$$\begin{pmatrix} \mathbf{b}_{n+1,g} \\ \mathbf{b}_{n+1,u} \end{pmatrix} = [(\mathbf{E}_0^{[2]})^2 - (\omega_f^R)^2 1]^{-1} \otimes \begin{pmatrix} \mathbf{E}_0^{[2]} & \omega_f^R \mathbf{S}^{[2]} \\ \omega_f^R \mathbf{S}^{[2]} & \mathbf{E}_0^{[2]} \end{pmatrix} \begin{pmatrix} \boldsymbol{R}_{n+1,g} \\ \boldsymbol{R}_{n+1,u} \end{pmatrix} \tag{96}$$

giving directly new trial vectors that are symmetric and antisymmetric, respectively. The new trial vectors are added to the reduced subspaces in eqs 86 and 87, and the iterative procedure is repeated until convergence.

5.2. Standard Response Equation. Inserting eq 79 into eq 1, we may write eq 1 in the matrix blocked form

$$\begin{pmatrix} \mathbf{E}^{[2]} & -\omega\mathbf{S}^{[2]} \\ -\omega\mathbf{S}^{[2]} & \mathbf{E}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}_g \\ \mathbf{X}_u \end{pmatrix} = \begin{pmatrix} \mathbf{G}_g \\ \mathbf{G}_u \end{pmatrix} \tag{97}$$

where $\mathbf{G}_g$ and $\mathbf{G}_u$ are the symmetric and antisymmetric components, respectively, of the gradient vector $\mathbf{G}$. Equation 97 may be solved using a subspace algorithm, as for the response eigenvalue equation. Assuming that after iteration n, the trial vector subspaces of the forms given in eqs 86 and 87 are known, we obtain a reduced standard response equation according to

$$\begin{pmatrix} \mathbf{E}^{[2]}_{\mathrm{R},gg} & -\omega\mathbf{S}^{[2]}_{\mathrm{R},gu} \\ -\omega\mathbf{S}^{[2]}_{\mathrm{R},ug} & \mathbf{E}^{[2]}_{\mathrm{R},uu} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\mathrm{R},g} \\ \mathbf{X}_{\mathrm{R},u} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{\mathrm{R},g} \\ \mathbf{G}_{\mathrm{R},u} \end{pmatrix} \tag{98}$$

where the reduced Hessian and metric matrices are given in eq 89, and the reduced right-hand side has the form

$$\mathbf{G}_{\mathrm{R},g} = \mathbf{b}_g^T\mathbf{G}_g, \quad \mathbf{G}_{\mathrm{R},u} = \mathbf{b}_u^T\mathbf{G}_u \tag{99}$$

Solving eq 98 leads to the optimal solution vectors of the form given in eqs 90 and 91, and the residuals $\boldsymbol{R}_{n+1,g}$ and $\boldsymbol{R}_{n+1,u}$ may then be obtained as

$$\begin{aligned} \boldsymbol{R}_{n+1,g} &= \mathbf{E}^{[2]}\mathbf{X}_{n+1,g} - \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,u} - \mathbf{G}_g \\ &= \sum_{i=1}^{n} (\mathbf{X}_{\mathrm{R},g})_i\boldsymbol{\sigma}_{ig} - \omega\sum_{i=1}^{n} (\mathbf{X}_{\mathrm{R},u})_i\boldsymbol{\rho}_{ig} - \mathbf{G}_g \end{aligned} \tag{100}$$

$$\begin{aligned} \boldsymbol{R}_{n+1,u} &= \mathbf{E}^{[2]}X_{n+1,u} - \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,g} - \mathbf{G}_u \\ &= \sum_{i=1}^{n} (\mathbf{X}_{\mathrm{R},u})_i\boldsymbol{\sigma}_{iu} - \omega\sum_{i=1}^{n} (\mathbf{X}_{\mathrm{R},g})_i\boldsymbol{\rho}_{iu} - \mathbf{G}_u \end{aligned} \tag{101}$$

From the residuals the new trial vectors may be constructed by using a preconditioner similar to the one in eq 95

$$\begin{pmatrix} \mathbf{b}_{n+1,g} \\ \mathbf{b}_{n+1,u} \end{pmatrix} = [(\mathbf{E}_0^{[2]})^2 - \omega^2 1]^{-1} \otimes \begin{pmatrix} \mathbf{E}_0^{[2]} & \omega\mathbf{S}^{[2]} \\ \omega\mathbf{S}^{[2]} & \mathbf{E}_0^{[2]} \end{pmatrix} \begin{pmatrix} \boldsymbol{R}_{n+1,g} \\ \boldsymbol{R}_{n+1,u} \end{pmatrix} \tag{102}$$

The iterative sequence is continued until convergence.

For frequencies smaller than the HOMO−LUMO gap, both the matrix in eq 97 and the preconditioner in eq 102 are positive definite, and the preconditioned CG/CR algorithms may therefore be used. For larger frequencies neither of these matrices are positive definite, and therefore the CG/CR algorithms cannot safely be applied. Equation 97 then has to be solved using the preconditioned subspace algorithm.

For small ω the condition number is significantly reduced by the preconditioning in eq 102, as the dominant contribution to the large eigenvalues is removed. However, when ω becomes larger and approaches an orbital energy difference, the preconditioning matrix will approach a singularity leading to an increased condition number and poor convergence. Note that the preconditioner in eq 102 is identical to the one used in the algorithm with paired trial vectors described in Section 4, as in eq 102 we have exploited the symmetry properties of $\mathbf{E}^{[2]}$ and $\mathbf{S}^{[2]}$ matrices in eqs 74 and 75.

5.3. Damped Response Equation. *5.3.1. Symmetric and Antisymmetric Trial Vectors.* When the damped response equation in eq 2 is solved, we write the solution vector in the form of eq 81. This gives a set of linear equations for each of the four components

$$\begin{aligned} \mathbf{E}^{[2]}\mathbf{X}_g^{\mathrm{R}} - \omega\mathbf{S}^{[2]}\mathbf{X}_u^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_u^{\mathrm{I}} &= \mathbf{G}_g^{\mathrm{R}} \\ \mathbf{E}^{[2]}\mathbf{X}_u^{\mathrm{R}} - \omega\mathbf{S}^{[2]}\mathbf{X}_g^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_g^{\mathrm{I}} &= \mathbf{G}_u^{\mathrm{R}} \\ \mathbf{E}^{[2]}\mathbf{X}_u^{\mathrm{I}} - \omega\mathbf{S}^{[2]}\mathbf{X}_g^{\mathrm{I}} - \gamma\mathbf{S}^{[2]}\mathbf{X}_g^{\mathrm{R}} &= \mathbf{G}_u^{\mathrm{I}} \\ \mathbf{E}^{[2]}\mathbf{X}_g^{\mathrm{I}} - \omega\mathbf{S}^{[2]}\mathbf{X}_u^{\mathrm{I}} - \gamma\mathbf{S}^{[2]}\mathbf{X}_u^{\mathrm{R}} &= \mathbf{G}_g^{\mathrm{I}} \end{aligned} \tag{103}$$

where $\mathbf{G}_g^{\mathrm{R}}$ and $\mathbf{G}_u^{\mathrm{R}}$ ($\mathbf{G}_g^{\mathrm{I}}$ and $\mathbf{G}_u^{\mathrm{I}}$) are the symmetric and antisymmetric components of the real (imaginary) part of the gradient vector $\mathbf{G}$. When expressed in matrix form, eq 103 is represented by a set of linear equations for a nonsymmetric matrix. Equation 103 may alternatively be expressed in terms of a coupled set of linear equations for a real symmetric matrix (as in Section 3.4):

$$\begin{pmatrix} \mathbf{E}^{[2]} & -\omega\mathbf{S}^{[2]} & \gamma\mathbf{S}^{[2]} & 0 \\ -\omega\mathbf{S}^{[2]} & \mathbf{E}^{[2]} & 0 & \gamma\mathbf{S}^{[2]} \\ \gamma\mathbf{S}^{[2]} & 0 & -\mathbf{E}^{[2]} & \omega\mathbf{S}^{[2]} \\ 0 & \gamma\mathbf{S}^{[2]} & \omega\mathbf{S}^{[2]} & -\mathbf{E}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}_g^{\mathrm{R}} \\ \mathbf{X}_u^{\mathrm{R}} \\ \mathbf{X}_u^{\mathrm{I}} \\ \mathbf{X}_g^{\mathrm{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_g^{\mathrm{R}} \\ \mathbf{G}_u^{\mathrm{R}} \\ -\mathbf{G}_u^{\mathrm{I}} \\ -\mathbf{G}_g^{\mathrm{I}} \end{pmatrix} \tag{104}$$

The complex non-Hermitian damped response equation (eq 2) is thus transformed to a set of linear equations for a real symmetric but indefinite matrix, where the coupling between the different components is considered explicitly. In iteration n, the residuals of eq 104 are given by

$$\begin{aligned} \boldsymbol{R}_{n+1,g}^{\mathrm{R}} &= \mathbf{E}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{R}} - \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{I}} - \mathbf{G}_g^{\mathrm{R}} \\ \boldsymbol{R}_{n+1,u}^{\mathrm{R}} &= \mathbf{E}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{R}} - \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{R}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{I}} - \mathbf{G}_u^{\mathrm{R}} \\ \boldsymbol{R}_{n+1,u}^{\mathrm{I}} &= -\mathbf{E}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{I}} + \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{I}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{R}} + \mathbf{G}_u^{\mathrm{I}} \\ \boldsymbol{R}_{n+1,g}^{\mathrm{I}} &= -\mathbf{E}^{[2]}\mathbf{X}_{n+1,g}^{\mathrm{I}} + \omega\mathbf{S}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{I}} + \gamma\mathbf{S}^{[2]}\mathbf{X}_{n+1,u}^{\mathrm{R}} + \mathbf{G}_g^{\mathrm{I}} \end{aligned} \tag{105}$$

where $\mathbf{X}_{n+1,g}^{\mathrm{R}}$, $\mathbf{X}_{n+1,u}^{\mathrm{R}}$, $\mathbf{X}_{n+1,u}^{\mathrm{I}}$ and $\mathbf{X}_{n+1,g}^{\mathrm{I}}$ are the optimal solution vectors in iteration n. The residual may be used to check for convergence. In the preconditioned iterative subspace approach, new trial vectors are obtained by preconditioning the residuals:

$$\begin{pmatrix} \mathbf{b}_{n+1,g}^{\mathrm{R}} \\ \mathbf{b}_{n+1,u}^{\mathrm{R}} \\ \mathbf{b}_{n+1,u}^{\mathrm{I}} \\ \mathbf{b}_{n+1,g}^{\mathrm{I}} \end{pmatrix} = \mathscr{P} \otimes \begin{pmatrix} \mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D} \\ \mathscr{B} & \mathscr{A} & \mathscr{D} & \mathscr{C} \\ \mathscr{C} & \mathscr{D} & -\mathscr{A} & -\mathscr{B} \\ \mathscr{D} & \mathscr{C} & -\mathscr{B} & -\mathscr{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{R}_{n+1,g}^{\mathrm{R}} \\ \boldsymbol{R}_{n+1,u}^{\mathrm{R}} \\ \boldsymbol{R}_{n+1,u}^{\mathrm{I}} \\ \boldsymbol{R}_{n+1,g}^{\mathrm{I}} \end{pmatrix} \tag{106}$$

where $\mathscr{P}$, $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{D}$ are given as

$$\begin{aligned} \mathscr{P} &= [[(\mathbf{E}_0^{[2]})^2 - (\omega^2 - \gamma^2)1]^2 + 4\omega^2\gamma^2 1]^{-1} \\ \mathscr{A} &= \mathbf{E}_0^{[2]}[(\mathbf{E}_0^{[2]})^2 - (\omega^2 - \gamma^2)1] \\ \mathscr{B} &= \omega\mathbf{S}^{[2]}[(\mathbf{E}_0^{[2]})^2 - (\omega^2 + \gamma^2)1] \\ \mathscr{C} &= \gamma\mathbf{S}^{[2]}[(\mathbf{E}_0^{[2]})^2 + (\omega^2 + \gamma^2)1] \\ \mathscr{D} &= 2\omega\gamma\mathbf{E}_0^{[2]} \end{aligned}$$

The preconditioner in eq 106 is an approximation to the symmetric matrix in eq 104 with $\mathbf{E}^{[2]}$ replaced by $\mathbf{E}_0^{[2]}$.

The condition number of eq 104 is significantly reduced by the preconditioning in eq 106, since the dominant component to the large eigenvalues is removed, as described in Section 3.3, and the convergence is therefore radically improved. However, for small values of γ, the condition number of the preconditioning matrix will increase significantly when the optical frequency approaches an orbital energy difference resulting in a slower rate of convergence. In the resonance region, the extent to which the rate of convergence is reduced depends on the density of states at the optical frequency of interest, since none of the eigenvalues in the resonance region are brought close to unity by the preconditioning matrix. Another issue with calculations of response properties in the resonance regions is concerned with the fact that the preconditioner in eq 106 is not positive definite, and as a consequence, the preconditioned CR algorithm cannot be safely applied (see Appendix B). We therefore conclude that the preconditioned iterative subspace algorithm is the best option for solving eq 104.

5.3.2. Damped Response Equations using Hermitian and Anti-Hermitian Trial Vectors. Villaume et al.[30] have described how the response equations within relativistic theory may be efficiently solved using an algorithm where the solution vectors are split according to their Hermiticity and the time-reversal symmetry. It is beyond the scope of the present work to give a detailed review of the solution of response equations within a relativistic theory, but it is clear that the structure of the relativistic and nonrelativistic equations are similar differing mainly in the fact that electronic wave functions (molecular orbitals) in a relativistic theory are complex and of two- or four-component forms. Matrix representations of operators are consequently complex and characterized by Hermiticity. The introduction of spin-adapted excitation operators in the nonrelativistic realm (see eq 6) finds its correspondence in the exploitation of time-reversal symmetry. The equations presented in the previous section represent a nonrelativistic adaptation of the response equations presented in ref 30 with the requirement of using real trial vectors. We will here give a brief description of how the two representations are connected.

The independent variables in the relativistic formulation of damped response theory were chosen as $\{x^{\mathrm{H}}, i\bar{x}^{\mathrm{A}}, x^{\mathrm{A}}, i\bar{x}^{\mathrm{H}}\}$, representing Hermitian and anti-Hermitian matrices (or vectors) and their complex conjugates. These variables can, however, be divided into symmetric and antisymmetric components according to

$$\begin{aligned} x^{\mathrm{H}} &= x_g^{\mathrm{R}} + ix_u^{\mathrm{I}}, & x^{\mathrm{A}} &= x_u^{\mathrm{R}} + ix_g^{\mathrm{I}} \\ ix^{\mathrm{H}} &= -x_u^{\mathrm{R}} + ix_g^{\mathrm{I}} = -(x^{\mathrm{A}})^*, & ix^{\mathrm{A}} &= -x_g^{\mathrm{R}} + ix_u^{\mathrm{I}}, = -(x^{\mathrm{H}})^* \end{aligned} \tag{107}$$

where the set $\{x_g^{\mathrm{R}}, x_u^{\mathrm{R}}, x_u^{\mathrm{I}}, x_g^{\mathrm{I}}\}$ contains the real variables employed in the previous section. We note that either set may be used to span the solution space as there is a nonsingular transformation connecting them according to

$$\begin{aligned} x_g^{\mathrm{R}} &= \frac{1}{2}(x^{\mathrm{H}} + (x^{\mathrm{H}})^*), & x_u^{\mathrm{I}} &= -\frac{i}{2}(x^{\mathrm{H}} - (x^{\mathrm{H}})^*) \\ x_u^{\mathrm{R}} &= \frac{1}{2}(x^{\mathrm{A}} + (x^{\mathrm{A}})^*), & x_g^{\mathrm{I}} &= -\frac{i}{2}(x^{\mathrm{A}} - (x^{\mathrm{A}})^*). \end{aligned} \tag{108}$$

With use of the set of complex variables, Villaume et al. obtained a damped response equation that can be written in the form

$$\begin{pmatrix} \mathbf{E}^{[2]} & -\omega\mathbf{S}^{[2]} & \gamma\mathbf{S}^{[2]} & 0 \\ -\omega\mathbf{S}^{[2]} & \mathbf{E}^{[2]} & 0 & \gamma\mathbf{S}^{[2]} \\ -\gamma\mathbf{S}^{[2]} & 0 & \mathbf{E}^{[2]} & -\omega\mathbf{S}^{[2]} \\ 0 & -\gamma\mathbf{S}^{[2]} & -\omega\mathbf{S}^{[2]} & \mathbf{E}^{[2]} \end{pmatrix} \begin{pmatrix} \mathbf{X}^{\mathrm{H}} \\ \mathbf{X}^{\mathrm{A}} \\ \bar{\mathbf{X}}^{\mathrm{A}} \\ \bar{\mathbf{X}}^{\mathrm{H}} \end{pmatrix} = \begin{pmatrix} \mathbf{G}^{\mathrm{H}} \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{109}$$

where all components of the solution vector, i.e., $\mathbf{X}^{\mathrm{H}}$, $\mathbf{X}^{\mathrm{A}}$, $\bar{\mathbf{X}}^{\mathrm{A}}$, and $\bar{\mathbf{X}}^{\mathrm{H}}$, are complex. The 4×4 blocked matrix in eq 109 is not symmetric, but it can be made symmetric by reversing signs of rows three and four, in analogy to what we did to arrive at eq 104.

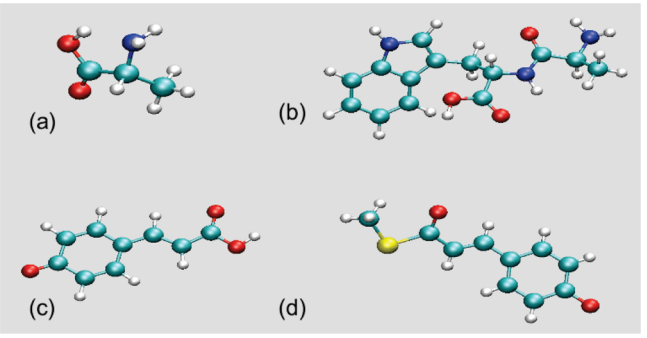

Figure 1. Molecular structures of: (a) alanine, (b) Ala-Trp, (c) pCA^-, and (d) pCT^-.

6. NUMERICAL EXAMPLES

All calculations have been carried out using a local version of DALTON.[39] For all reported results, preconditioning is used in the canonical SCF orbital basis.

6.1. Standard Response Equation. In this section, we compare the convergence of the various algorithms for solving the standard response equation in eq 1 with a property gradient referring to the electric dipole operator and to a residual norm of 10^{-4} au. Our calculation refers to alanine (see Figure 1a) at the HF level of theory with use of the 6-31G basis set,[40] and the performance of the algorithms is compared at the two frequencies of 0.1 and 0.4 au corresponding to the nonresonance and resonance regions of the spectrum, respectively. The adopted molecular geometry is taken from the NIST online database (ref 41).

In Table 1, we compare the convergence of the standard response equation in the nonresonance region, at $\omega = 0.1$ au, using five different iterative algorithms (the transition frequency of the lowest singlet excited state of alanine in the present wave function parametrization is 0.2129 au). In the column order of the table, the presented results refer to the following algorithms: (i) the general subspace approach (discussed in Section 3.3); (ii) the subspace algorithm with paired trial vectors (discussed in Section 4); (iii) the subspace algorithm with symmetrized trial vectors (discussed in Section 5.2); (iv) the CG algorithm with symmetrized trial vectors (discussed in Section 5.2 and Appendix A); and (v) the CR algorithm with symmetrized trial vectors (discussed in Section 5.2 and Appendix B). Residuals $\boldsymbol{R}$ for the general subspace approach and the algorithm with paired trial vectors are given in eqs 34 and 66, respectively. Residuals $\boldsymbol{R}_g$ and $\boldsymbol{R}_u$ in the algorithms with the symmetrized trial vectors are given in eqs 100 and 101, respectively. For all algorithms, one

Table 1. Residual Norms (in au) for Solving the Standard Response Equation with $\omega = 0.1$ au (See eq 1) Using Different Iterative Approaches[a]

	iterative approach with trial vectors of the type:						
			symmetrized				
	general	paired	subspace			CG	CR
iteration number	$\|\|R\|\|$	$\|\|R\|\|$	$\|\|R_g\|\|$	$\|\|R_u\|\|$	$\|\|R_g + R_u\|\|$	$\|\|R_g + R_u\|\|$	$\|\|R_g + R_u\|\|$
1	4.51473	1.81777	0.12568	1.81342	1.81777	0.81189	0.71915
2	1.85994	0.68841	0.04235	0.68710	0.68841	0.23247	0.22109
3	0.47364	0.18677	0.02104	0.18558	0.18677	0.09411	0.08346
4	0.17550	0.08252	0.01044	0.08186	0.08252	0.05206	0.04005
5	0.08012	0.04116	0.00363	0.04100	0.04116	0.03190	0.02242
6	0.04412	0.02368	0.00180	0.02361	0.02368	0.01961	0.01384
7	0.03078	0.00933	0.00076	0.00930	0.00933	0.00698	0.00591
8	0.01389	0.00399	0.00045	0.00396	0.00399	0.00364	0.00301
9	0.00523	0.00178	0.00014	0.00177	0.00178	0.00151	0.00133
10	0.00198	0.00050	0.00004	0.00050	0.00050	0.00047	0.00043
11	0.00084	0.00018	0.00001	0.00018	0.00018	0.00018	0.00017
12	0.00054	0.00004	0.00001	0.00004	0.00004	0.00005	0.00005
13	0.00019						
14	0.00005						

[a] The calculations refer to the alanine molecule at the HF/6-31G level of theory.

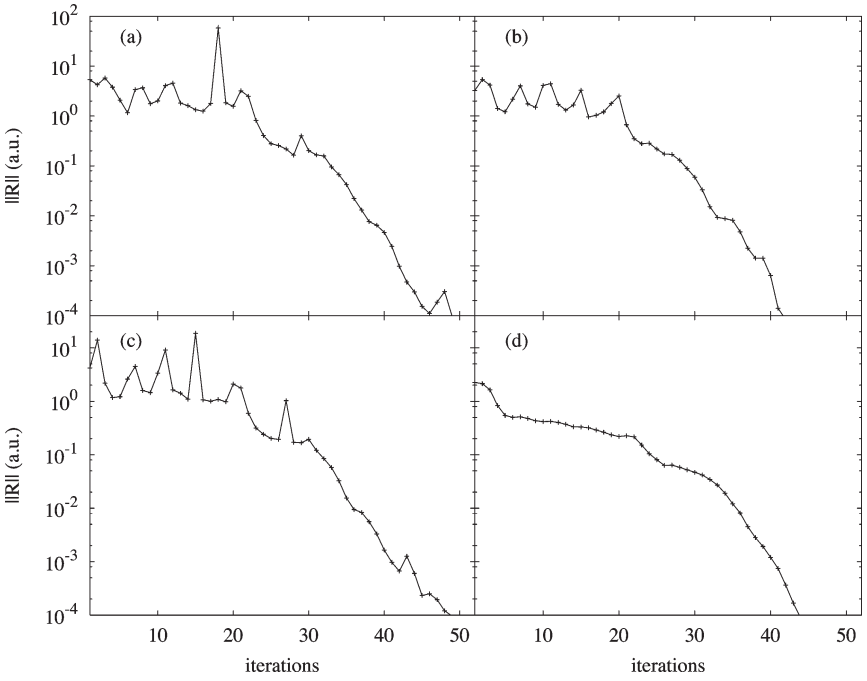

Figure 2. Convergence of the standard response equation for alanine in the resonance region ($\omega = 0.4$ au) using different response solver algorithms: (a) general subspace approach, (b) subspace approach with symmetrized trial vectors, (c) CG approach with symmetrized trial vectors and (d) CR approach with symmetrized trial vectors. Calculations refer to the HF level of theory using basis set 6-31G.

linear transformation of $\mathbf{E}^{[2]}$ (and $\mathbf{S}^{[2]}$) on a trial vector is required in each iteration.

To begin with, we remark that for ω smaller than the HOMO−LUMO gap results obtained using the general subspace approach (second column) are identical to those obtained when the CG algorithm is applied (see Appendix C), and such a comparison is therefore left out in the table. When comparing the results obtained using the general subspace approach (second column) with the ones for the subspace algorithm with paired trial vectors (third column), a small reduction in the number of iterations is observed. This is due to the fact that a larger subspace is spanned when paired trial vectors are used. The obtained

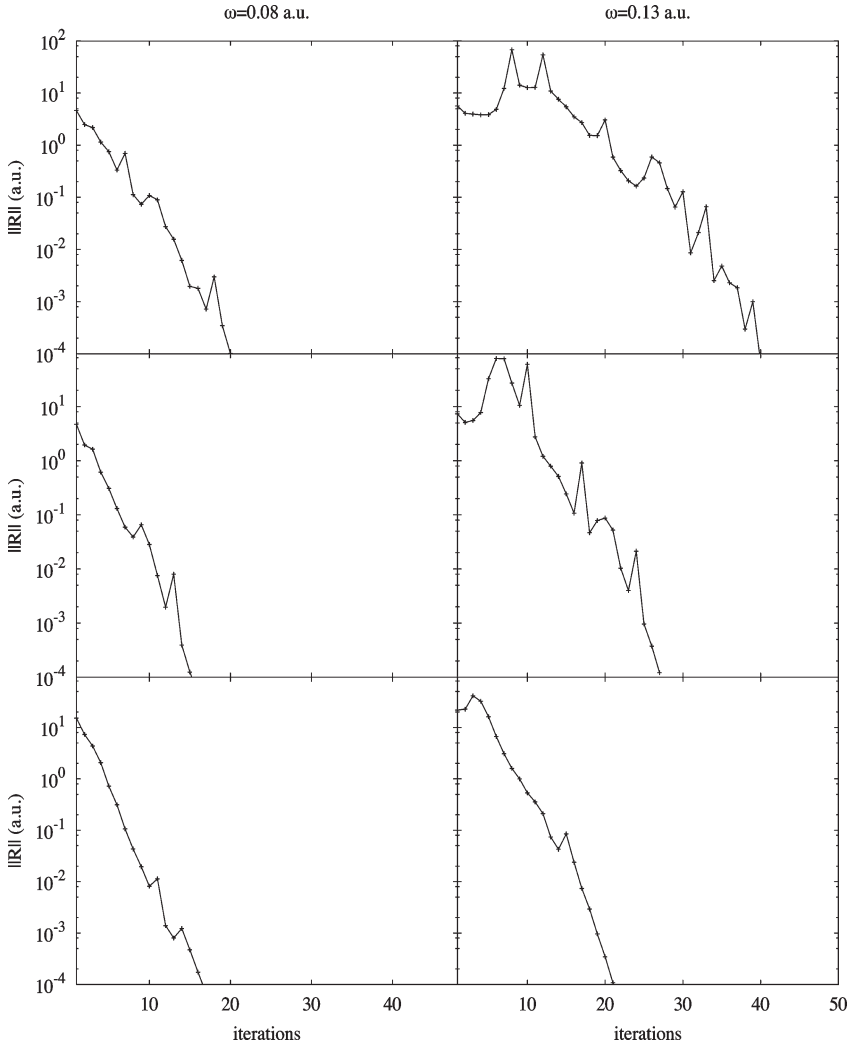

Figure 3. Convergence of damped response equation for pCA^- at frequencies 0.08 (nonresonance region) and 0.13 au (resonance region) employing a damping parameter $\gamma = 0.005$ au. Adopted algorithms are (top) CR with general vectors, (middle) CR with symmetrized trial vectors and (bottom) subspace algorithm with symmetrized trial vectors. Calculations refer to CAM-B3LYP level of theory using basis set cc-pVDZ.

improvement is not dramatic in the present case, but we recall that it is obtained without additional computational cost (see eq 60). The iteration sequences obtained using the algorithm with symmetrized trial vectors (sixth column) and the algorithm with paired trial vectors (third column) are identical. This is related to the fact that the symmetrized trial vectors may be considered as a special set of paired trial vectors, as described in Section 5. When the subspace approach is used, both paired and symmetrized vectors span identical subspaces in any given iteration, and for this reason, their respective iteration sequences become identical, as is seen from Table 1.

Comparing the sixth and the seventh column in Table 1, we see that the iteration sequence for the subspace algorithm with symmetrized trial vectors differs slightly from its CG counterpart. This is due to the fact that, in the CG algorithm, optimal coefficients are determined for vectors of the form $\mathbf{b} = (\mathbf{b}_g, \mathbf{b}_u)^T$ (see eqs 138−140), while in the subspace approach optimal coefficients are obtained for the individual components $\mathbf{b}_g$ and $\mathbf{b}_u$ (see eq 98). However, this difference only has a minor effect on the rate of convergence. By comparing the seventh and the eighth column in Table 1, we see that the convergence rate of the CG and CR algorithms is similar in accordance with theoretical findings.[38] For ω smaller than the HOMO−LUMO gap, we have found that the standard response equation is best solved using the preconditioned CG or CR algorithm, thereby avoiding the storing and the manipulation of a large set of trial vectors needed to set up the reduced space equations.

In Figure 2, we compare the convergence behavior for solving the standard response equation at a near resonant frequency of $\omega = 0.4$ au. In the present wave function parametrization, this frequency is close to the two spin-allowed transition frequencies at 0.4042 and 0.4097 au, respectively. We compare the convergence behavior of four different iterative algorithms: (a) the general subspace approach, (b) the subspace algorithm with symmetrized trial vectors, (c) the CG algorithm with symmetrized trial vectors, and (d) the CR algorithm with symmetrized trial vectors. The results obtained using the algorithm with paired trial vectors are not presented, since, as shown in Table 1, the iteration sequence is identical to the one obtained using the algorithm with symmetrized vectors (plotted in Figure 2b). By comparing results presented in Figure 2 to those listed in Table 1, we see that more iterations are needed to converge the standard

response equation in the resonance region. This is due to the fact that when the frequency ω is close to a transition frequency, the preconditioning of the response equations becomes less efficient as discussed in Section 3.3.

A convergence improvement occurs when the standard response equation is solved using the subspace algorithm with symmetrized trial vectors compared to the general subspace approach (see Figure 2a and b), as discussed above. In Figure 2c and d, the performance of the CG algorithm and the CR algorithm is also displayed. Neither the CG or the CR algorithm can be safely applied for ω = 0.4 au, since neither the matrix in eq 97 nor the preconditioner in eq 102 is positive definite. In the CG algorithm the minimization of the quadratic function (see eqs 111 and 112) is only uniquely defined for an equation with a positive definite matrix. In the CR algorithm the minimization of the preconditioned residual (see eq 156) is only uniquely defined for a positive definite preconditioner. In contrast, in the subspace algorithm the optimal solution is determined by solving a reduced subspace equation and is therefore uniquely determined. In Figure 2, only a small degradation in the convergence is seen when the CG and the CR algorithms are applied in regions where the minimization in the CG and the CR algorithms is not uniquely defined. In other cases the degradation may be much larger, e.g., for ω = 0.6 au the convergence has not been obtained when the CR algorithm has been applied.

For all reported results preconditioning is used. For calculations performed without preconditioning, the convergence has been obtained in 187 and 351 iterations, for ω = 0.1 and 0.4 au, respectively. However, we note that the convergence is rather poor due to the fact that calculations have been performed in the AO representation. In the MO representation (traditionally used), much faster convergence of nonpreconditioned response equations is observed.

The convergence behavior reported above for the general subspace approach and the subspace algorithm with symmetrized trial vectors has been found for calculations on a large variety of molecular systems and also using more extended basis sets as well as at the KS level of theory. The standard response equation for ω larger than the HOMO−LUMO gap is thus best solved using the iterative subspace algorithm with symmetrized trial vectors.

6.2. Damped Response Equation. In this section, we compare the convergence of the various algorithms for solving the damped response equation in eq 2 with a property gradient referring to the electric dipole operator and to a residual norm of 10^{-4} au. In addition to alanine, we will consider three other molecules (see Figure 1), namely: (i) the model dipeptide alanine-tryptophan (Ala-Trp), (ii) deprotonated *trans-p*-coumaric acid (pCA^-), and (iii) deprotonated *trans*-thiomethyl-*p*-coumarate (pCT^-). The adopted molecular structures have different sources. The structure of Ala-Trp is obtained by use of the MAESTRO program[42] without carrying out any additional optimization. The structure of pCA^- is that optimized at the KS level of theory using the B3LYP[43,44] exchange−correlation functional and the aug-cc-pVDZ basis set.[45] The structure of pCT^-, which is a model for the protein chromophore PYP (the photoactive yellow protein), is taken from experiment.[46] The presented response calculations are carried out at the HF and KS levels of theory, in the latter case with employment of the Coulomb attenuated B3LYP exchange−correlation functional (CAM-B3LYP).[47]

In Figure 3, we present the convergence of the damped response equation for calculations on pCA^- at the CAM-B3LYP/

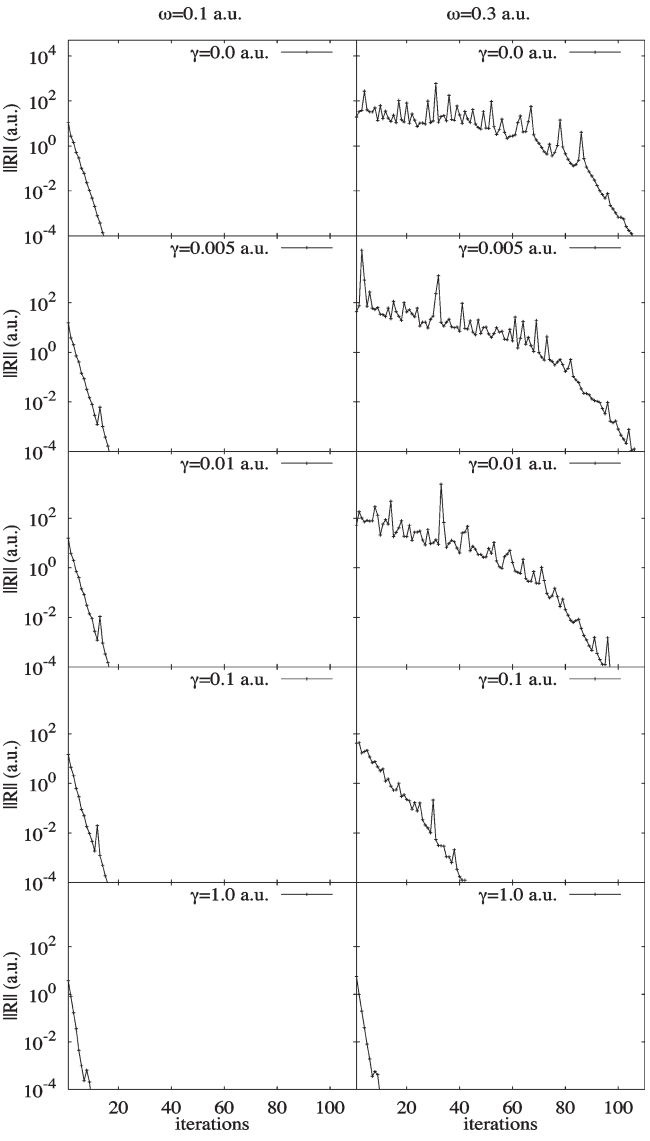

Figure 4. Convergence of the damped response equation for the Ala-Trp molecule at frequencies 0.1 (nonresonance region) and 0.3 au (resonance region) employing damping parameters 0.0, 0.005, 0.01, 0.1, and 1.0 au, respectively. Calculations refer to use of the subspace algorithm with symmetrized trial vectors and are carried out at the CAM-B3LYP level of theory using basis set 6-31G.

cc-pVDZ level of theory. In these calculations, we have adopted a damping parameter of γ = 0.005 au, and we present the norm of the residual as obtained with (i) the CR algorithm with general vectors (discussed in Section 3.4 and Appendix B), (ii) the CR algorithm with symmetrized trial vectors (discussed in Section 5.3 and Appendix B) and (iii) the subspace algorithm with symmetrized trial vectors (discussed in Section 5.3). The left panels correspond to an optical frequency in the nonresonant region (ω = 0.08 au), whereas the right panels correspond to ω = 0.13 au, which is close to the calculated lowest spin-allowed transition frequency ω_1 = 0.1321 au. In the upper and middle panels, where the CR algorithm is used, the preconditioners in eqs 56 and 106 are not positive definite, and the preconditioned CR algorithm is therefore not uniquely defined. In both the upper and middle panels, we therefore observe a significant degradation in the convergence compared to the subspace

approach (lower panel) in particular for the calculations in the resonance region. The convergence obtained with the CR algorithm is also rather erratic because the preconditioner is not positive definite, while a monotonic convergence would be obtained if the preconditioner was positive definite. In the subspace approach with symmetrized trial vectors in the lower panel, we see only a small degradation in the convergence when the frequency is increased.

In Figure 4, we present calculations for Ala-Trp at the CAM-B3LYP/6-31G level of theory using the subspace algorithm with symmetrized trial vectors presented in Section 5.3. The damped response equation was solved at two different frequencies namely ω = 0.1 (left panels) and ω = 0.3 au (right panels) and five different values of the damping parameter namely γ equal to 0.0, 0.005, 0.01, 0.1, and 1.0 au. Note that results obtained with the last two values of γ do not have any physical meaning and are given here only to demonstrate the robustness of our algorithm toward large γ values. We can conclude that, in the nonresonant region, the convergence is independent of the damping parameter γ, owing to the fact that preconditioner given in eq 106 is very efficient. By comparing the left and the right panels of Figure 4, we see that significantly more iterations are needed to converge the damped response equation in the resonance region. This is due to the fact that in the higher frequency region, there is a relatively high density of excited states as seen in Figure 5, and the preconditioning therefore becomes less efficient as these states will interact strongly in this frequency region. For small values of the γ parameter, the preconditioner in eq 106 becomes nearly singular in the higher frequency region, and therefore slow convergence is observed in the top panels in Figure 4. For unphysically large values of γ, on the other hand, the convergence is fast (as can be seen in the bottom panels in Figure 4) due to the fact that the γ contribution becomes the dominant contribution in the preconditioner.

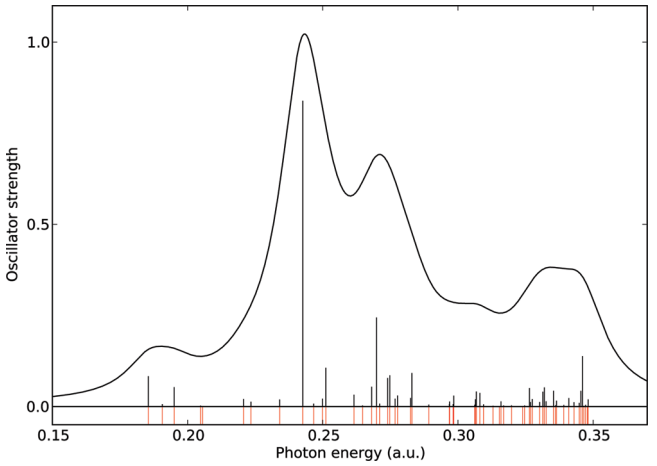

Figure 5. Excitation spectrum for the Ala-Trp molecule at the CAM-B3LYP/6-31G level of theory. Bar heights represent oscillator strengths. All singlet excited states in the spectral region up to 0.3481 au are indicated by red bars below the spectrum.

In Figure 6, we report HF/6-31G calculations of the absorptive part (imaginary part) of the electric dipole polarizability of the pCT^{-} molecule—we here adopt a damping parameter γ equal to 0.005 au. In ref 5, it was demonstrated that, for response theory calculations on pCT^{-}, the preconditioned two-level subspace approach represented a significant improvement in convergence rate with respect to the complex response equation solver described in ref 3. We here compare the convergence of the preconditioned two-level subspace approach with the convergence obtained using the subspace algorithm with symmetrized trial vectors, as described in Section 5.3. The algorithmic efficiency is depicted in terms of the number of iterations that are needed to obtain a solution vector at each spectral point. In each iteration, linear transformations of trial vectors by $\mathbf{E}^{[2]}$ are carried out. The linear transformation requires that a Fock matrix is set up where the computationally expensive two-electron integrals have to be constructed. At the time the calculations presented in ref 5 had been performed, in each iteration of the preconditioned two-level subspace approach, two linear transformations (and therefore two Fock matrices) for the real and the imaginary component of the trial vector were carried out sequentially. However, in an improved implementation, linear transformations are carried simultaneously, and the cost in one iteration is independent of the number of trial vectors, and therefore we compare the

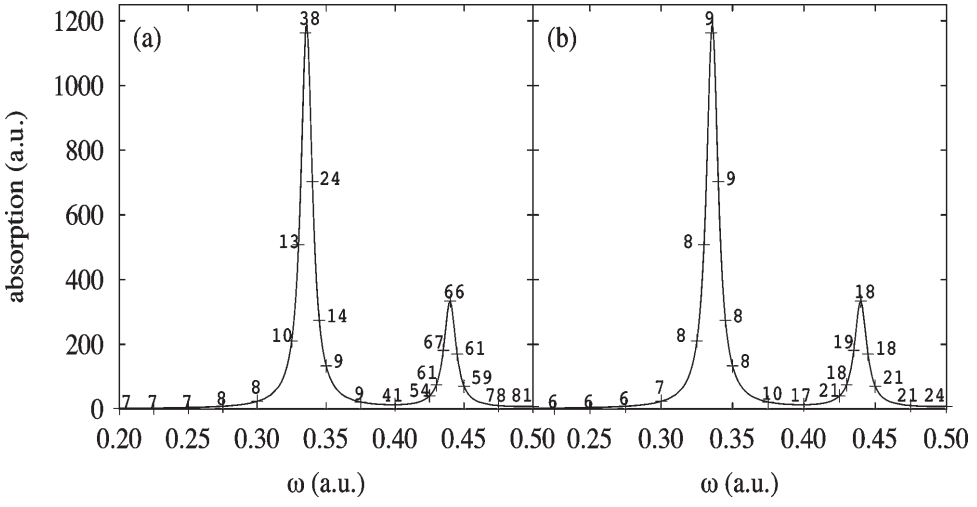

Figure 6. Absorptive part of the electric dipole polarizability determined for pCT^{-} with γ = 0.005 au. Number of iterations required to reach convergence (residual threshold of 10^{-4} au) is given at spectral points. Comparison is made of (a) response solver proposed in ref 5 and (b) response solver with symmetrized trial vectors proposed in the present work. Calculations refer the HF level of theory using basis set 6-31G.

total number of iterations instead of the total number of Fock matrices constructed during the calculations as was done in ref 5. As seen in Figure 6a and as discussed in the previous section, the computational cost depends on the closeness of the optical frequency to resonances and also on the density of excited states. In the nonresonance region, rapid convergence is observed due to a small coupling between the real and the imaginary component of the solution vectors, but in the resonance region, where there are large couplings between the real and the imaginary components, the convergence is considerably slower. In Figure 6a, the computational cost in the region above $\omega = 0.4$ au is therefore also significantly higher than in the region below (even though there is one strong excitation at around $\omega = 0.33$ au).

When comparing the total number of iterations in the calculations using the preconditioned two-level subspace approach (Figure 6a) and the subspace algorithm with the symmetrized trial vectors (Figure 6b), we see that the performance of the latter represents a clear improvement. In the region of the lowest excitation (around $\omega = 0.33$ au), the algorithm with symmetrized trial vectors requires only two to three more iterations to converge eq 2, as compared to the case in the nonresonance region. In the region with a high density of excited states (above $\omega = 0.4$ au), the number of iterations increases but not as drastically as when the preconditioned two-level subspace approach is used. In comparison to employment of the preconditioned two-level subspace approach[5] in spectral resonant regions, the computational cost associated with use of the proposed symmetrized trial vector approach is substantially lower—at times the computational cost is reduced by as much as a factor of 4. This is due to the very efficient way of obtaining new trial vectors in the approach with the symmetrized trial vectors using a preconditioner that includes the coupling between vectors explicitly. It should also be noted that in the preconditioned two-level subspace approach, all the trial vectors (from so-called micro- and macroiterations) are stored on disk; that makes this approach inefficient for treating large molecular systems. When the approach with the symmetrized trial vectors is used, the storage requirements are reduced.

6.3. Response Eigenvalue Equation. In Table 2, the convergence behavior of the response eigenvalue equation (eq 3) is shown. The calculations are carried out for the Ala-Trp molecule at the CAM-B3LYP level of theory using the 6-31G basis set. The response eigenvalue equation was solved for the first eigenvalue to a residual norm of 10^{-4} au, and three different iterative algorithms were used (i) the Davidson algorithm with paired trial vectors (discussed in Section 4); (ii) the Davidson algorithm with symmetrized trial vectors (discussed in Section 5.1); and (iii) the Olsen algorithm (discussed in Section 3.2.2) with symmetrized trial vectors. The residual $\boldsymbol{R}$ for the Davidson algorithm with paired trial vectors is given in eq 67, whereas residuals $\boldsymbol{R}_g$ and $\boldsymbol{R}_u$ in the algorithms with the symmetrized trial vectors are given in eqs 100 and 101, respectively.

As discussed in Sections 5 and 6.1, the symmetrized vectors may be considered as a special set of paired vectors. Therefore the Davidson algorithm, both with paired and with symmetrized trial vectors, will in each iteration span the same subspace, and identical iteration sequences are thus obtained. Since the paired structure of the reduced Hessian and metric matrices ($\mathbf{E}_R^{[2]}$ and $\mathbf{S}_R^{[2]}$) is preserved during the iterative procedure in both algorithms, paired eigenvalues are obtained in the reduced subspace in both cases.

Table 2. Residual Norms (in au) for Determination of the Lowest Root of the Eigenvalue Equation (See eq 3) Using Different Iterative Approaches[a]

	Davidson alg.		Olsen alg.
	paired	symmetrized	symmetrized
iteration number	$\|\|R\|\|$	$\|\|R_g + R_u\|\|$	$\|\|R_g + R_u\|\|$
1	0.11663	0.11663	0.11663
2	0.03819	0.03819	0.03819
3	0.02775	0.02775	0.02782
4	0.01723	0.01723	0.01725
5	0.02148	0.02148	0.02149
⋮	⋮	⋮	⋮
18	0.00146	0.00146	0.00144
19	0.00134	0.00134	0.00132
20	0.00068	0.00068	0.00068
21	0.00045	0.00045	0.00045
22	0.00030	0.00030	0.00030
23	0.00018	0.00018	0.00018
24	0.00008	0.00008	0.00008

[a] The calculations refer to the Ala-Trp molecule at the CAM-B3LYP/6-31G level of theory.

From the third and the fourth column of Table 2, we see that the iteration sequence obtained using the Olsen algorithm is nearly identical to the one obtained using the Davidson algorithms. The Olsen algorithm does not give a significant improvement unless $\mathbf{E}_0^{[2]}$ is approaching $\mathbf{E}^{[2]}$, and this is generally not the case when response eigenvalue equations are solved. Therefore response eigenvalue equations may safely be solved using the Davidson algorithm with paired or symmetrized trial vectors.

7. CONCLUSIONS

The response equations in Hartree—Fock, multiconfigurational self-consistent field, and Kohn—Sham density functional theory have identical structure. We have discussed the algorithms that have been used for solving these equations, and introduced new, efficient, algorithms based on the splitting of trial vectors into symmetric and antisymmetric components. We have used that, upon linear transformation by the generalized Hessian $\mathbf{E}^{[2]}$ (generalized metric $\mathbf{S}^{[2]}$), the symmetry of a trial vector is preserved (reversed). By introducing symmetrized trial vectors, the response equations are transformed to a form which allows them to be solved efficiently, in particular due to the fact that very efficient preconditioners are used with an exact treatment of the $\mathbf{S}^{[2]}$ matrix.

We have demonstrated that the most efficient algorithm for solving the standard and damped response equation is the preconditioned iterative subspace algorithm with symmetrized trial vectors. For the standard response equation at optical frequencies smaller than the HOMO—LUMO gap, the use of symmetrized trial vectors enables the employment of the preconditioned CG or CR algorithms, and one thereby avoids the storing and the handling of a large set of trial vectors to set up the reduced space equations.

For solving the response eigenvalue equations, it is equally efficient to use algorithms with paired[25] as with symmetrized trial vectors, and no improvement is obtained by using the Olsen

algorithm.[26] We have explicitly proven that, when the Davidson algorithm is applied, the lowest eigenvalue of the reduced subspace equation converges monotically toward the exact eigenvalue.

APPENDIX A. CONJUGATE GRADIENT (CG) ALGORITHM

The CG algorithm[15] is an iterative method for solving a set of linear equations:

$$\mathbf{Ax} = \mathbf{b} \tag{110}$$

where $\mathbf{A}$ is a symmetric and positive definite matrix. The minimization of the quadratic function:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Ax} - \mathbf{x}^T\mathbf{b} \tag{111}$$

gives

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{Ax} - \mathbf{b} = 0 \tag{112}$$

that is a solution to eq 110. Note that the residual of eq 111 may be expressed as

$$\mathbf{r} = -\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{b} - \mathbf{Ax} \tag{113}$$

In iteration $n+1$ of the CG algorithm the trial solution $\mathbf{x}_{n+1}$ is parametrized as a linear combination of the residual $\mathbf{r}_n$ of iteration n and the optimal search directions $\{\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_{n-1}\}$ of the previous iterations:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=0}^{n-1} \alpha_i^{(n)}\mathbf{p}_i + \alpha_n^{(n)}\mathbf{r}_n \tag{114}$$

where $\alpha_n^{(n)}$ is determined by applying eq 112. The idea of the CG algorithm is to identify an optimal search direction $\mathbf{p}_n$ and replacing the multiple search direction in eq 114 by a single search direction:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n^{(n)}\mathbf{p}_n \tag{115}$$

such that minimization of the function $f(\mathbf{x}_{n+1})$ calculated from eqs 114 or 115 gives a mathematically identical result. In the CG algorithm, all but the last direction may be discarded without loss of convergence rate.

We will now present the development of the CG algorithm with special emphasis on the developments that allow us to connect the solution that is obtained in the CG algorithm with the one obtained using the subspace iterative algorithm, that is commonly used in quantum chemistry.

A.1. First Iteration. Let us assume that $\mathbf{x}_0$ is our starting guess. In the first iteration we have a single search direction

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0 \tag{116}$$

which therefore is optimal, i.e., $\mathbf{p}_0 = \mathbf{r}_0$. The solution vector at iteration 1 has therefore a form

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0^{(0)}\mathbf{p}_0 \tag{117}$$

Minimization of $f(\mathbf{x}_1)$ in the direction $\mathbf{p}_0$ gives

$$\frac{\partial f(\mathbf{x}_1)}{\partial \alpha_0^{(0)}} = \alpha_0^{(0)}\mathbf{p}_0^T\mathbf{Ap}_0 - \mathbf{p}_0^T\mathbf{r}_0 = 0 \tag{118}$$

that determines the optimal step length in the direction $\mathbf{p}_0$:

$$\alpha_0^{(0)} = \frac{\mathbf{p}_0^T\mathbf{r}_0}{\mathbf{p}_0^T\mathbf{Ap}_0} \tag{119}$$

The residual at $\mathbf{x}_1$:

$$\mathbf{r}_1 = \mathbf{b} - \mathbf{Ax}_1 \tag{120}$$

is orthogonal to the search direction $\mathbf{p}_0$:

$$\mathbf{p}_0^T\mathbf{r}_1 = \mathbf{r}_0^T\mathbf{r}_1 = 0. \tag{121}$$

A.2. Second Iteration. The trial vector now has components along $\mathbf{p}_0$ and $\mathbf{r}_1$

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_0^{(1)}\mathbf{p}_0 + \alpha_1^{(1)}\mathbf{r}_1 \tag{122}$$

Minimization of $f(\mathbf{x}_2)$, with respect to coefficients $\alpha_0^{(1)}$ and $\alpha_1^{(1)}$, gives the subspace equations:

$$\begin{pmatrix} \mathbf{p}_0^T\mathbf{Ap}_0 & \mathbf{p}_0^T\mathbf{Ar}_1 \\ \mathbf{r}_1^T\mathbf{Ap}_0 & \mathbf{r}_1^T\mathbf{Ar}_1 \end{pmatrix} \begin{pmatrix} \alpha_0^{(1)} \\ \alpha_1^{(1)} \end{pmatrix} = \begin{pmatrix} \mathbf{p}_0^T\mathbf{r}_1 \\ \mathbf{r}_1^T\mathbf{r}_1 \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{r}_1^T\mathbf{r}_1 \end{pmatrix} \tag{123}$$

where in the first row of eq 123 the coefficient $\alpha_0^{(1)}$ may be written in terms of $\alpha_1^{(1)}$:

$$\alpha_0^{(1)} = -\alpha_1^{(1)}\frac{\mathbf{p}_0^T\mathbf{Ar}_1}{\mathbf{p}_0^T\mathbf{Ap}_0} \tag{124}$$

The trial vector $\mathbf{x}_2$ may be expressed as a unidirectional search:

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1^{(1)}\mathbf{p}_1 \tag{125}$$

where the direction $\mathbf{p}_1$ is

$$\mathbf{p}_1 = \mathbf{r}_1 - \frac{\mathbf{p}_0^T\mathbf{Ar}_1}{\mathbf{p}_0^T\mathbf{Ap}_0}\mathbf{p}_0 \tag{126}$$

The optimal step length in the direction $\mathbf{p}_1$ becomes

$$\alpha_1^{(1)} = \frac{\mathbf{r}_1^T\mathbf{p}_1}{\mathbf{p}_1^T\mathbf{Ap}_1} \tag{127}$$

The residual at $\mathbf{x}_2$ becomes

$$\mathbf{r}_2 = \mathbf{b} - \mathbf{Ax}_2 \tag{128}$$

or

$$\mathbf{r}_2 = \mathbf{r}_1 - \alpha_1^{(1)}\mathbf{Ap}_1 \tag{129}$$

It can be shown that

$$\mathbf{p}_0^T\mathbf{Ap}_1 = 0, \quad \mathbf{r}_2^T\mathbf{Ap}_0 = 0 \tag{130}$$

$$\mathbf{r}_2^T\mathbf{p}_1 = 0,\ \mathbf{r}_2^T\mathbf{r}_1 = 0,\ \mathbf{r}_2^T\mathbf{p}_0 = 0,\ \mathbf{r}_2^T\mathbf{r}_0 = 0 \tag{131}$$

A.3. n + 1th Iteration. The previous directions and residuals fulfill the relations

$$\mathbf{r}_i^T\mathbf{p}_j = 0,\ \mathbf{r}_i^T\mathbf{r}_j = 0,\ i,j = 0,1\ , ..., n,\ i > j \tag{132}$$

$$\mathbf{r}_i^T\mathbf{Ap}_j = 0,\ i,j = 0,1\ , ..., n,\ i > j+1 \tag{133}$$

$$\mathbf{p}_i^T \mathbf{A} \mathbf{p}_j = \mathbf{p}_i^T \mathbf{A} \mathbf{p}_j \delta_{ij}, \quad i,j = 0,1\ ,\ ...,\ n-1 \tag{134}$$

The new trial vector is initially written as a general vector in the space spanned by the previous search directions $\{\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_{n-1}\}$ and the current residual $\mathbf{r}_n$:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=0}^{n-1} \alpha_i^{(n)} \mathbf{p}_i + \alpha_n^{(n)} \mathbf{r}_n \tag{135}$$

Minimizing $f(\mathbf{x}_{n+1})$ with respect to the $n+1$ free parameters one obtains the subspace equation:

$$\begin{pmatrix} \mathbf{p}_0^T\mathbf{A}\mathbf{p}_0 & \mathbf{p}_0^T\mathbf{A}\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_0^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{p}_0^T\mathbf{A}\mathbf{r}_n \\ \mathbf{p}_1^T\mathbf{A}\mathbf{p}_0 & \mathbf{p}_1^T\mathbf{A}\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_1^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{p}_1^T\mathbf{A}\mathbf{r}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{p}_0 & \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{r}_n \\ \mathbf{r}_n^T\mathbf{A}\mathbf{p}_0 & \mathbf{r}_n^T\mathbf{A}\mathbf{p}_1 & \cdots & \cdots & \mathbf{r}_n^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{r}_n^T\mathbf{A}\mathbf{r}_n \end{pmatrix} \begin{pmatrix} \alpha_0^{(n)} \\ \alpha_1^{(n)} \\ \vdots \\ \vdots \\ \alpha_{n-1}^{(n)} \\ \alpha_n^{(n)} \end{pmatrix} = \begin{pmatrix} \mathbf{p}_0^T\mathbf{r}_n \\ \mathbf{p}_1^T\mathbf{r}_n \\ \vdots \\ \vdots \\ \mathbf{p}_{n-1}^T\mathbf{r}_n \\ \mathbf{r}_n^T\mathbf{r}_n \end{pmatrix} \tag{136}$$

that is equivalent to

$$\begin{pmatrix} \mathbf{p}_0^T\mathbf{A}\mathbf{p}_0 & 0 & ... & 0 & 0 & 0 \\ 0 & \mathbf{p}_1^T\mathbf{A}\mathbf{p}_1 & ... & 0 & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{p}_{n-2}^T\mathbf{A}\mathbf{p}_{n-2} & 0 & 0 \\ 0 & 0 & ... & 0 & \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{r}_n \\ 0 & 0 & ... & 0 & \mathbf{r}_n^T\mathbf{A}\mathbf{p}_{n-1} & \mathbf{r}_n^T\mathbf{A}\mathbf{r}_n \end{pmatrix} \begin{pmatrix} \alpha_0^{(n)} \\ \alpha_1^{(n)} \\ \vdots \\ \vdots \\ \alpha_{n-2}^{(n)} \\ \alpha_{n-1}^{(n)} \\ \alpha_n^{(n)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 0 \\ \mathbf{r}_n^T\mathbf{r}_n \end{pmatrix} \tag{137}$$

where eqs 132−134 were used. It can be seen from eq 137 that solving the reduced space equation can be avoided, due to the fact that all the information necessary to obtain the optimal solution vector is contained in the last three trial vectors.

Due to the form of the subspace equation the optimal solution vector $\mathbf{x}_{n+1}$ may be expressed in terms of a single search direction $\mathbf{p}_n$ and the optimal step length $\alpha_n^{(n)}$ as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n^{(n)} \mathbf{p}_n \tag{138}$$

where

$$\mathbf{p}_n = \mathbf{r}_n - \frac{\mathbf{p}_{n-1}^T \mathbf{A} \mathbf{r}_n}{\mathbf{p}_{n-1}^T \mathbf{A} \mathbf{p}_{n-1}} \mathbf{p}_{n-1} \tag{139}$$

$$\alpha_n^{(n)} = \frac{\mathbf{p}_n^T \mathbf{r}_n}{\mathbf{p}_n^T \mathbf{A} \mathbf{p}_n} \tag{140}$$

Therefore only three vectors: $\mathbf{x}_n$, $\mathbf{r}_n$ and $\mathbf{p}_{n-1}$ need to be stored in each iteration. Equations 132−134 are valid for n increased by one, and the iteration procedure of the CG algorithm is established.

The CG algorithm in iteration $n+1$ may be viewed as an iterative subspace algorithm that uses the basis of the trial vectors $\{\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_{n-1}, \mathbf{r}_n\}$. The iterative subspace algorithm in eq 28 is set up in terms of trial vectors consisting of $\{\mathbf{r}_0, \mathbf{r}_1, ..., \mathbf{r}_{n-1}, \mathbf{r}_n\}$. From eqs 114 and 115 it is seen that these two sets of trial vectors span the same space and that the CG and the iterative subspace algorithms of eq 28 therefore give the same iteration sequence.

APPENDIX B. CONJUGATE RESIDUAL ALGORITHM AND ITS PRECONDITIONING

B.1. Conjugate Residual (CR) Algorithm. In this section, the CR algorithm[19] is discussed (described in detail in ref 38). The CR algorithm is an iterative method for solving a set of linear equations of a form in eq 110, where $\mathbf{A}$ is a symmetric but not necessarily positive definite matrix. The residual for a general vector $\mathbf{x}$ is given as

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x} \tag{141}$$

The solution to eq 110 is obtained by minimization of the squared residual norm:

$$g(\mathbf{x}) = \mathbf{r}^T \mathbf{r} \tag{142}$$

as

$$\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0 \tag{143}$$

We shortly summarize the CR algorithm with emphasis on the development that allows us to connect to iterative subspace algorithms that are used in quantum chemistry. We refer to the paper by Ziółkowski et al.,[38] where a more thorough development of CR is given, and the connection to the DIIS algorithm[48,49] (that is commonly used in quantum chemistry) is presented.

After n iterations of the CR algorithm, the optimal solution vector $\mathbf{x}_n$, n residuals $\mathbf{r}_i$, and n optimal directions $\mathbf{p}_i$ are known. It may be shown that the previous optimal directions and residuals fulfill the relations:[38]

$$\mathbf{r}_i^T \mathbf{A} \mathbf{p}_j = 0,\ \mathbf{r}_i^T \mathbf{A} \mathbf{r}_j = 0,\ i,j = 0,1\ ,\ ...,\ n,\ i > j \tag{144}$$

$$\mathbf{r}_i^T \mathbf{A}^2 \mathbf{p}_j = 0,\ i,j = 0,1\ ,\ ...,\ n,\ i > j+1 \tag{145}$$

$$\mathbf{p}_i^T \mathbf{A}^2 \mathbf{p}_j = \mathbf{p}_i^T \mathbf{A}^2 \mathbf{p}_j \delta_{ij},\ i,j = 0,1\ ,\ ...,\ n-1 \tag{146}$$

In iteration $n+1$, the new trial vector may be written as a general vector in the space spanned by the previous optimal search directions $\{\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_{n-1}\}$, and the current residual $\mathbf{r}_n$:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=0}^{n-1} \alpha_i^{(n)} \mathbf{p}_i + \alpha_n^{(n)} \mathbf{r}_n \tag{147}$$

Minimizing $g(\mathbf{x}_{n+1})$ with respect to the $n+1$ free parameters leads to a subspace equation:

$$\begin{pmatrix} \mathbf{p}_0^T\mathbf{A}^2\mathbf{p}_0 & \mathbf{p}_0^T\mathbf{A}^2\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_0^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{p}_0^T\mathbf{A}^2\mathbf{r}_n \\ \mathbf{p}_1^T\mathbf{A}^2\mathbf{p}_0 & \mathbf{p}_1^T\mathbf{A}^2\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_1^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{p}_1^T\mathbf{A}^2\mathbf{r}_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_0 & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_1 & \cdots & \cdots & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{r}_n \\ \mathbf{r}_n^T\mathbf{A}^2\mathbf{p}_0 & \mathbf{r}_n^T\mathbf{A}^2\mathbf{p}_1 & \cdots & \cdots & \mathbf{r}_n^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{r}_n^T\mathbf{A}^2\mathbf{r}_n \end{pmatrix} \begin{pmatrix} \alpha_0^{(n)} \\ \alpha_1^{(n)} \\ \vdots \\ \vdots \\ \alpha_{n-1}^{(n)} \\ \alpha_n^{(n)} \end{pmatrix} = \begin{pmatrix} \mathbf{p}_0^T\mathbf{A}\mathbf{r}_n \\ \mathbf{p}_1^T\mathbf{A}\mathbf{r}_n \\ \vdots \\ \vdots \\ \mathbf{p}_{n-1}^T\mathbf{A}\mathbf{r}_n \\ \mathbf{r}_n^T\mathbf{A}\mathbf{r}_n \end{pmatrix} \tag{148}$$

that is equivalent to

$$\begin{pmatrix} \mathbf{p}_0^T\mathbf{A}^2\mathbf{p}_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \mathbf{p}_1^T\mathbf{A}^2\mathbf{p}_1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_{n-2} & 0 & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{r}_n \\ 0 & 0 & \cdots & 0 & \mathbf{r}_n^T\mathbf{A}^2\mathbf{p}_{n-1} & \mathbf{r}_n^T\mathbf{A}^2\mathbf{r}_n \end{pmatrix} \begin{pmatrix} \alpha_0^{(n)} \\ \alpha_1^{(n)} \\ \vdots \\ \vdots \\ \alpha_{n-2}^{(n)} \\ \alpha_{n-1}^{(n)} \\ \alpha_n^{(n)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 0 \\ \mathbf{r}_n^T\mathbf{A}\mathbf{r}_n \end{pmatrix} \tag{149}$$

where relations eqs 144−146 were used. From eq 149 it may be seen that also in the CR algorithm only the three last trial vectors are necessary to obtain the optimal solution vector in iteration $n+1$.

The optimal solution vector $\mathbf{x}_{n+1}$ may be expressed in terms of a single search direction $\mathbf{p}_n$ and of the optimal step length $\alpha_n^{(n)}$ as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n^{(n)}\mathbf{p}_n \tag{150}$$

where

$$\mathbf{p}_n = \mathbf{r}_n - \frac{\mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{r}_n}{\mathbf{p}_{n-1}^T\mathbf{A}^2\mathbf{p}_{n-1}}\mathbf{p}_{n-1} \tag{151}$$

$$\alpha_n^{(n)} = \frac{\mathbf{p}_n^T\mathbf{A}\mathbf{r}_n}{\mathbf{p}_n^T\mathbf{A}^2\mathbf{p}_n} \tag{152}$$

and

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n^{(n)}\mathbf{A}\mathbf{p}_n \tag{153}$$

Equations 144−146 are valid for n increased by one, and the iteration procedure of the CR algorithm is established.

In the CR algorithm, we may thus discard all but the last direction without information lost, similarly to what is found in the CG algorithm. The main difference between the CG and the CR algorithms is that step lengths in the CG algorithm are determined from a minimization of $f(\mathbf{x})$ in eq 111, whereas in the CR algorithm the step length is determined from a minimization of $g(\mathbf{x})$ in eq 143. The minimization of $f(\mathbf{x})$ is only unique for a symmetric positive matrix $\mathbf{A}$, while the minimization of $g(\mathbf{x})$ is unique also for a nonsingular symmetric but not necessarily positive definite matrix. The CG algorithm therefore can only be safely applied for a positive definite matrix $\mathbf{A}$, while the CR algorithm can also be applied when the matrix $\mathbf{A}$ is not positive definite.

B.2. Preconditioned CR Algorithm. To improve the convergence of the CR algorithm, the set of linear equations in eq 110 may be solved in a preconditioned form, where a coordinate transformation is introduced to produce a new set of linear equations which has a lower condition number. To carry out this transformation, eq 110 is multiplied with the transpose of a nonsingular matrix: $\mathscr{P}$

$$\mathscr{P}^T\mathbf{A}\mathscr{P}\mathbf{Y} - \mathscr{P}^T\mathbf{b} = 0 \tag{154}$$

where $\mathbf{Y} = \mathscr{P}^{-1}\mathbf{x}$. We may now solve eq 154 using the CR algorithm and transform the solution to the original coordinates. Alternatively, eq 154 may be solved in the original basis using modified CR equations. To do this, we write the residual in the $\mathbf{Y}$ basis as

$$\mathbf{r}^{\mathscr{P}} = \mathscr{P}^T\mathbf{b} - \mathscr{P}^T\mathbf{A}\mathscr{P}\mathbf{Y} = \mathscr{P}^T\mathbf{r} \tag{155}$$

Optimal directions are now determined by minimizing

$$\mathbf{g}^{\mathscr{P}}(\mathbf{x}) = (\mathbf{r}^{\mathscr{P}})^T\mathbf{r}^{\mathscr{P}} = \mathbf{r}^T\mathscr{P}\mathscr{P}^T\mathbf{r} = \mathbf{r}^T\mathbf{C}^{-1}\mathbf{r} \tag{156}$$

where

$$\mathbf{C}^{-1} = \mathscr{P}\mathscr{P}^T \tag{157}$$

The preconditioned analogue to eqs 150−153 becomes[38]

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n^{(n)}\mathbf{C}^{-1}\mathbf{p}_n \tag{158}$$

where

$$\mathbf{p}_n = \mathbf{r}_n - \frac{\mathbf{p}_{n-1}^T\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{r}_n}{\mathbf{p}_{n-1}^T\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{p}_{n-1}}\mathbf{p}_{n-1} \tag{159}$$

$$\alpha_n^{(n)} = \frac{\mathbf{p}_n^T\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{r}_n}{\mathbf{p}_n^T\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{p}_n} \tag{160}$$

and

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n^{(n)}\mathbf{A}\mathbf{C}^{-1}\mathbf{p}_n \tag{161}$$

Choosing $\mathscr{P}^T$ such that $\mathbf{C}$ is a good approximation to $\mathbf{A}$ ensures that the linear equations are solved in a basis where the matrix $\mathbf{A}$ has a lower condition number.

From eqs 158 and 159, it seems like the preconditioned CR algorithm may be applied whenever an easily invertible matrix $\mathbf{C}$ that is a good approximation to the $\mathbf{A}$ matrix can be identified. However, the step lengths in the preconditioned CR algorithm are determined from a minimization of $g^{\mathscr{P}}(\mathbf{x})$ in eq 156, and this minimization is uniquely defined only when $\mathbf{C}^{-1}$ corresponds to a coordinate transformation and can be decomposed as in eq 157. The decomposition in eq 157 requires that $\mathbf{C}^{-1}$ is positive

definite, and we thus conclude that the CR preconditioned equations in eqs 158−161 can be safely applied only for a positive definite preconditioner **C**.

APPENDIX C. MACDONALD'S THEOREM FOR THE GROUND-STATE RESPONSE EIGENVALUE EQUATION

For a ground state, the response eigenvalue equation may be expressed as

$$\mathbf{E}^{[2]}\mathbf{X} = \mathbf{S}^{[2]}\mathbf{X}\boldsymbol{\omega} \tag{162}$$

where the positive $\mathbf{S}^{[2]}$-norm eigensolutions $\mathbf{X}_+ = \{\mathbf{X}_{1+}, \mathbf{X}_{2+}, ..., \mathbf{X}_{p+}, ...\}$ are associated with positive eigenvalues $\{\omega_1, \omega_2, ..., \omega_p, ...\}$. We assume that the eigenvalues are sorted in ascending order. Similarly, the negative $\mathbf{S}^{[2]}$-normed eigensolutions $\mathbf{X}_- = \{\mathbf{X}_{1-}, \mathbf{X}_{2-}, ..., \mathbf{X}_{p-}, ...\}$ are associated with the negative eigenvalues $\{-\omega_1, -\omega_2, ..., -\omega_p, ...\}$. The matrix **X** is the collection of positive and negative eigenvectors $[\mathbf{X}_+, \mathbf{X}_-]$, and $\boldsymbol{\omega}$ is a diagonal matrix collecting the eigenvalues in corresponding order.

Let us now compare the solution to eq 162 for two paired subspaces:

$$\mathbf{b}' = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_N, \mathbf{b}_1^{\mathrm{P}}, \mathbf{b}_2^{\mathrm{P}}, ..., \mathbf{b}_N^{\mathrm{P}}\} \tag{163}$$

$$\mathbf{b}'' = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_N, \mathbf{b}_{N+1}, \mathbf{b}_1^{\mathrm{P}}, \mathbf{b}_2^{\mathrm{P}}, ..., \mathbf{b}_N^{\mathrm{P}}, \mathbf{b}_{N+1}^{\mathrm{P}}\} \tag{164}$$

where $\mathbf{b}''$ is obtained from $\mathbf{b}'$ by adding the paired vectors $\mathbf{b}_{N+1}$ and $\mathbf{b}_{N+1}^{\mathrm{P}}$. The solution to eq 162 in $\mathbf{b}'$ and $\mathbf{b}''$ spaces may be expressed as (see eq 21):

$$(\mathbf{X}')^{\dagger}\mathbf{E}^{[2]}\mathbf{X}' = \begin{pmatrix} \boldsymbol{\omega}'_{\mathrm{exc}} & 0 \\ 0 & \boldsymbol{\omega}'_{\mathrm{exc}} \end{pmatrix}; \quad (\mathbf{X}')^{\dagger}\mathbf{S}^{[2]}\mathbf{X}' = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \tag{165}$$

$$(\mathbf{X}'')^{\dagger}\mathbf{E}^{[2]}\mathbf{X}'' = \begin{pmatrix} \boldsymbol{\omega}''_{\mathrm{exc}} & 0 \\ 0 & \boldsymbol{\omega}''_{\mathrm{exc}} \end{pmatrix}; \quad (\mathbf{X}'')^{\dagger}\mathbf{S}^{[2]}\mathbf{X}'' = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \tag{166}$$

where $\mathbf{X}'$ contains N paired eigenvectors in the $\mathbf{b}'$ subspace and $\mathbf{X}''$ contains $(N+1)$ paired eigenvectors in the $\mathbf{b}''$ subspace. The positive eigenvalues are collected in the diagonal matrices $\boldsymbol{\omega}'_{\mathrm{exc}}$ and $\boldsymbol{\omega}''_{\mathrm{exc}}$ and represent subspace approximations to the eigenvalues $\boldsymbol{\omega}$ in eq 162. Since $\mathbf{b}'$ is a subset of $\mathbf{b}''$, we may express the eigenvectors in the $\mathbf{b}'$ subspace in terms of the eigenvectors of the $\mathbf{b}''$ subspace as

$$\mathbf{X}'_{p+} = \sum_{q=1}^{N+1} [a_{p+q+}\mathbf{X}''_{q+} + a_{p+q-}\mathbf{X}''_{q-}] \tag{167}$$

where the normalization of $\mathbf{X}'_{p+}$ in the $\mathbf{S}^{[2]}$-norm implies

$$\sum_{q=1}^{N+1} [|a_{p+q+}|^2 - |a_{p+q-}|^2] = 1 \tag{168}$$

The lowest positive eigenvalue in $\mathbf{b}'$ is equal to

$$\omega'_{1+} = (\mathbf{X}'_{1+})^{\dagger}\mathbf{E}^{[2]}\mathbf{X}'_{1+} \tag{169}$$

However, by means of eq 167, the same eigenvalue can also be expressed as

$$\begin{aligned} \omega'_{1+} &= \sum_{n,m=1}^{N+1} [a_{1+n+}\mathbf{X}''_{n+} + a_{1+n-}\mathbf{X}''_{n-}]^{\dagger}\mathbf{E}^{[2]}[a_{1+m+}\mathbf{X}''_{m+} + a_{1+m-}\mathbf{X}''_{m-}] \\ &= \sum_{m}^{N+1} \omega''_{m+}[|a_{1+m+}|^2 + |a_{1+m-}|^2] \\ &\geq \sum_{m}^{N+1} \omega''_{1+}[|a_{1+m+}|^2 + |a_{1+m-}|^2] \\ &\geq \omega''_{1+} \sum_{m}^{N+1} [|a_{1+m+}|^2 - |a_{1+m-}|^2] \\ &= \omega''_{1+} \end{aligned} \tag{170}$$

We have thus shown that $\omega'_{1+} \geq \omega''_{1+}$ for iterative solutions of the response eigenvalue equation. This result represents a generalization of MacDonald's theorem for a symmetric positive definite eigenvalue equation with a positive definite metric.

APPENDIX D. ALGORITHMS BASED ON BLOCK DIAGONALIZATION OF THE HESSIAN

Another approach for solving the standard response equations has been presented by Casida.[2] This approach is analogous to the approach introduced by Jørgensen et al. in ref 11. In the approach in ref 2, the standard response equation:

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega\begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}\right]\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{pmatrix} \tag{171}$$

is transformed by means of the unitary transformation **U** in eq 22, to become equal to

$$\left[\begin{pmatrix} \mathbf{A}+\mathbf{B} & 0 \\ 0 & \mathbf{A}-\mathbf{B} \end{pmatrix} - \omega\begin{pmatrix} 0 & -\Sigma \\ -\Sigma & 0 \end{pmatrix}\right]\begin{pmatrix} \mathbf{X}' \\ -\mathbf{Y}' \end{pmatrix} = \begin{pmatrix} \mathbf{G}'_1 \\ -\mathbf{G}'_2 \end{pmatrix} \tag{172}$$

where

$$\begin{aligned} &\mathbf{X}' = \mathbf{X}+\mathbf{Y};\ \mathbf{Y}' = \mathbf{X}-\mathbf{Y};\ \mathbf{G}'_1 = \mathbf{G}_1+\mathbf{G}_2; \\ &\mathbf{G}'_2 = \mathbf{G}_1-\mathbf{G}_2 \end{aligned} \tag{173}$$

From eq 172 two separate equations are obtained namely:

$$\begin{aligned} &[(\mathbf{A}+\mathbf{B}) - \omega^2\mathbf{S}(\mathbf{A}-\mathbf{B})^{-1}\Sigma]\mathbf{X}' \\ &\quad = \mathbf{G}'_1 + \omega\Sigma(\mathbf{A}-\mathbf{B})^{-1}\mathbf{G}'_2 \end{aligned} \tag{174}$$

and

$$\begin{aligned} &[(\mathbf{A}-\mathbf{B}) - \omega^2\Sigma(\mathbf{A}+\mathbf{B})^{-1}\Sigma]\mathbf{Y}' \\ &\quad = \mathbf{G}'_2 - \omega\Sigma(\mathbf{A}+\mathbf{B})^{-1}\mathbf{G}'_1 \end{aligned} \tag{175}$$

The solution to the standard response equation is thus replaced by solving two sets of linear equations of half the dimension. When solving eqs 174 and 175, the inverse matrices $(\mathbf{A}-\mathbf{B})^{-1}$ and $(\mathbf{A}+\mathbf{B})^{-1}$ are required. When the matrices $(\mathbf{A}-\mathbf{B})$ and $(\mathbf{A}+\mathbf{B})$ are constructed explicitly, the inverse matrices may be obtained straightforwardly. But, when the matrix dimension is large and the iterative subspace algorithms need to

be used, the separation of eq 172 into eqs 174 and 175 becomes prohibitively inefficient. The response equations are better solved in a form referencing the double dimension in eq 172, as in eq 97, thereby separating the determination of the symmetric and antisymmetric components of the solution vector and maintaining the full coupling between the symmetric and antisymmetric components.

AUTHOR INFORMATION

Corresponding Author
*E-mail: joaka@ifm.liu.se.

ACKNOWLEDGMENT

J.K. and P.J. acknowledge Dr. Jeppe Olsen for very helpful and constructive discussions. J.K. and P.J. acknowledge support from the Lundbeck Foundation and the Danish Center for Scientific Computing. P.N. acknowledges financial support from the Swedish Research Council (grant no. 621-2010-5014).

REFERENCES

(1) Olsen, J.; Jørgensen, P. *J. Chem. Phys.* **1985**, *82*, 3235–3264.
(2) Casida, M. E. In *Recent Advances in Density Functional Methods*, Part 1; Chong, D. P., Ed.; World Scientific: Singapore, 1995; Vol. 1; Chapter 5, pp 155−192.
(3) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. *J. Chem. Phys.* **2001**, *115*, 10323–10334.
(4) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. *J. Chem. Phys.* **2005**, *123*, 194103–194120.
(5) Kristensen, K.; Kauczor, J.; Kjærgaard, T.; Jørgensen, P. *J. Chem. Phys.* **2009**, *131*, 044112(33).
(6) Orr, B. J.; Ward, J. F. *Mol. Phys.* **1971**, *20*, 513–526.
(7) Boyd, R. W. *Nonlinear Optics*, 3rd ed.; Academic Press: Burlington, MA, 2008; pp 155−157.
(8) Bunch, J. R.; Hopcroft, J. E. *Math. Comp.* **1974**, *28*, 231–236.
(9) Atkinson, K. A. *An Introduction to Numerical Analysis*, 2nd ed.; John Wiley & Sons: New York, 1989; pp 511−525.
(10) Trefethen, L. N.; Bau, D., III *Numerical Linear Algrebra*; SIAM: Philadelphia, PA, 1997; pp 172− 178.
(11) Jørgensen, P.; Linderberg, J. *Int. J. Quantum Chem.* **1970**, *4*, 587–602.
(12) Pople, J. A.; Krishnan, R.; Schlegel, H.; Binkley, J. S. *Int. J. Quantum Chem.* **1979**, *13*, 225–241.
(13) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1981**, *75*, 1284–1292.
(14) Wormer, P. E. S.; Visser, F.; Paldus, J. *J. Comput. Phys.* **1982**, *48*, 23–44.
(15) Hestenes, M. R.; Stiefel, E. *J. Res. Natl. Bur. Stand., Sect. A* **1952**, *49*, 409–436.
(16) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C++: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, U.K., 2007; pp 87−92.
(17) Shewchuk, J. R. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*; Carnegie Mellon University: Pittsburgh, PA, 1994.
(18) Saad, Y. *Iterative Methods for Sparse Linear Systems*; SIAM: Philadelphia, PA, 2003; pp 187−194.
(19) Stiefel, E. *Comment. Math. Helv.* **1955**, *29*, 157–179.
(20) Niklasson, A. M. N.; Challacombe, M. *Phys. Rev. Lett.* **2004**, *92*, 193001(4).
(21) Coriani, S.; Høst, S.; Jansík, B.; Thøgersen, L.; Olsen, J.; Jørgensen, P.; Reine, S.; Pawłowski, F.; Helgaker, T.; Sałek, P. *J. Chem. Phys.* **2007**, *126*, 154108(11).
(22) Davidson, E. R. *J. Comput. Phys.* **1975**, *17*, 87–94.
(23) Flament, J. P.; Gervais, H. P. *Int. J. Quantum Chem.* **1979**, *16*, 1347–1356.
(24) Hansen, A. E.; Voigt, B.; Rettrup, S.; Bouman, T. D. *Int. J. Quantum Chem.* **1983**, *23*, 595–611.
(25) Olsen, J.; Jensen, H. J. A.; Jørgensen, P. *J. Comput. Phys.* **1988**, *74*, 265–282.
(26) Olsen, J.; Jørgensen, P.; Simons, J. *Chem. Phys. Lett.* **1990**, *169*, 463–472.
(27) Saue, T. In *Relativistic Electronic Structure Theory - Part 1: Fundamentals*; Schwerdtfeger, P., Ed.; Elsevier, Amsterdam, The Netherlands, 2002; Chapter 7, pp 332−400.
(28) Saue, T.; Jensen, H. J. A. *J. Chem. Phys.* **2003**, *118*, 522–536.
(29) Bast, R.; Jensen, H. J. A.; Saue, T. *Int. J. Quantum Chem.* **2009**, *109*, 2091–2112.
(30) Villaume, S.; Saue, T.; Norman, P. *J. Chem. Phys.* **2010**, *133*, 064105(10).
(31) Kjærgaard, T.; Jørgensen, P.; Olsen, J.; Coriani, S.; Helgaker, T. *J. Chem. Phys.* **2008**, *129*, 054106(23).
(32) Čížek, J.; Paldus, J. *J. Chem. Phys.* **1967**, *47*, 3976–3985.
(33) Larsen, H.; Jørgensen, P.; Olsen, J.; Helgaker, T. *J. Chem. Phys.* **2000**, *113*, 8908–8917.
(34) Axelsson, O. *Iterative Solution Methods*; Cambridge University Press: Cambridge, U.K., 1996; pp 252−254.
(35) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: Chichester, U.K., 2000; pp 543−548.
(36) Parlett, B. N. *The Symmetric Eigenvalue Problem*; Prentice Hall: Englewood Cliff, New Jersey, 1980; pp 75−80.
(37) MacDonald, J. K. L. *Phys. Rev.* **1933**, *43*, 830–833.
(38) Ziółkowski, M.; Weijo, V.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **2008**, *128*, 204105(12).
(39) Helgaker, T.; Jensen, H. J. A.; Jørgensen, P.; *DALTON an ab initio electronic structure program*, release 2.0, 2005; http://www.kjemi.uio.no/software/dalton/dalton.html.
(40) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
(41) *NIST Standard Reference Database Number 69*; NIST: Gaithersburg, MD; http://webbook.nist.gov/chemistry.
(42) *Maestro*, v. 8.5; Schrodinger, LLC: Cambridge, MA, 2008; http://www.schrodinger.com.
(43) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(44) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
(45) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
(46) Genick, U. K.; Soltis, S. M.; Kuhn, P.; Canestrelli, I. L.; Getzoff, E. D. *Nature* **1998**, *392*, 206–209.
(47) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
(48) Pulay, P. *Chem. Phys. Lett.* **1980**, *73*, 393–398.
(49) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556–560.