Reduced-cost linear-response CC2 method based on natural orbitals and natural auxiliary functions

Dávid Mester,[a] Péter R. Nagy, and Mihály Kállay[b]
MTA-BME Lendület Quantum Chemistry Research Group, Department of Physical Chemistry and Materials Science, Budapest University of Technology and Economics, P.O. Box 91, H-1521 Budapest, Hungary

(Received 16 February 2017; accepted 28 April 2017; published online 16 May 2017)

A reduced-cost density fitting (DF) linear-response second-order coupled-cluster (CC2) method has been developed for the evaluation of excitation energies. The method is based on the simultaneous truncation of the molecular orbital (MO) basis and the auxiliary basis set used for the DF approximation. For the reduction of the size of the MO basis, state-specific natural orbitals (NOs) are constructed for each excited state using the average of the second-order Møller–Plesset (MP2) and the corresponding configuration interaction singles with perturbative doubles [CIS(D)] density matrices. After removing the NOs of low occupation number, natural auxiliary functions (NAFs) are constructed [M. Kállay, J. Chem. Phys. **141**, 244113 (2014)], and the NAF basis is also truncated. Our results show that, for a triple-zeta basis set, about 60% of the virtual MOs can be dropped, while the size of the fitting basis can be reduced by a factor of five. This results in a dramatic reduction of the computational costs of the solution of the CC2 equations, which are in our approach about as expensive as the evaluation of the MP2 and CIS(D) density matrices. All in all, an average speedup of more than an order of magnitude can be achieved at the expense of a mean absolute error of 0.02 eV in the calculated excitation energies compared to the canonical CC2 results. Our benchmark calculations demonstrate that the new approach enables the efficient computation of CC2 excitation energies for excited states of all types of medium-sized molecules composed of up to 100 atoms with triple-zeta quality basis sets. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4983277]

I. INTRODUCTION

The excited states of molecular systems are important for many areas of chemistry, biology, and physics. Not only well-established disciplines, such as spectroscopy, photochemistry, or photobiology, but also emerging fields like energy conversion and storage utilize some phenomenon related to excited states. Consequently, the development of efficient theoretical methods for excited states is one of the major objectives of modern quantum chemistry.

Nowadays the most accurate electronic structure theory is the coupled-cluster approach introduced by Čížek a half a century ago.[1] The hierarchical coupled-cluster (CC) methods based on the exponential parameterization of the wave function, that is, the CC singles and doubles (CCSD), the CC singles, doubles, and triples (CCSDT), . . ., enable the consideration of electron correlation with arbitrary accuracy. The CC methods can also be generalized to excited states, but it is not a trivial task. One option is to invoke linear-response (LR) theory, which was first extended to CC theory by Monkhorst and co-workers[2,3] and later by Koch *et al.*[4,5] An alternative approach, which is equivalent to LR-CC theory for excitation energies, is the equation-of-motion (EOM) CC method developed by Bartlett and co-workers.[6,7] The theory, implementation, and performance of the various LR-CC and EOM-CC methods are reviewed in several publications in the literature.[8–11] Concerning bigger molecules, even the relatively cheap CCSD approach is too expensive, and further approximations are required. A simple approximate CCSD approach is the second-order CC (CC2) method proposed and first implemented by Christiansen *et al.*[12–14] and later perfected by Hättig and co-workers.[15–17] The CC2 method supplies excitation energies and transition moments with a moderate error, at least for valence states, with respect to higher-order CC methods.[18,19] It is also worth mentioning the configuration interaction singles with perturbative doubles [CIS(D)] approach introduced by Head-Gordon *et al.*,[20] which can also be regarded as an approximate CC2 method.[21] The CIS(D) approach improves the configuration interaction singles (CIS) excitation energy with a perturbative correction for the double excitations and scales as the fifth power of the systems size as the CC2 method, but, in practice, it is considerably less expensive than CC2. Another closely related approach is the second-order algebraic diagrammatic construction [ADC(2)] method of Schirmer,[22,23] which has been recently extensively studied by Hättig[21] and Dreuw and co-workers.[24] The ADC(2) approach can be regarded as an iterative, Hermitized generalization of CIS(D), which is close in accuracy to CC2.

Despite the excellent numerical results, the applicability of CC methods is limited due to the steep scaling of their costs with the system size. Thus, for excited states of extended molecular systems, currently the time-dependent density functional theory (TD-DFT)[25] is the method of choice,

[a]Electronic mail: mester.david@mail.bme.hu
[b]Electronic mail: kallay@mail.bme.hu

even if the reliability of the TD-DFT results is frequently in question,[26,27] especially for Rydberg and charge transfer (CT) excitations, just as for conjugated π-electron systems. An alternative solution is the reduction of computational expenses of CC methods, such as CC2, which provides results consistently better than TD-DFT.[28–30] In the case of most reduced-cost CC2 approaches, the bottleneck related to the processing of four-center electron repulsion integrals (ERIs) is removed exploiting density fitting (DF) or Cholesky-decomposition (CD) techniques, which can also be combined with Laplace transform (LT) approximations. In the DF approach introduced by Shavitt and co-workers[31] and further developed by Whitten[32] and Dunlap,[33] the four-center ERIs are approximated as products of three- and two-center integrals with the aid of an auxiliary (fitting) basis set. For electron correlation methods, this significantly decreases the expenses of the integral transformation steps as well as the memory demand and the number of input-output (I/O) operations for the correlation calculations.[15,34] The DF approximation was adapted for CC2 by Hättig and co-workers,[15] who also developed analytic gradients for CC2 employing DF.[16,17] In the tensor hypercontraction (THC) scheme of Sherrill, Martínez, and their co-workers, which is a generalization of the DF technique, an even lower-order representation of the ERI tensor is used.[35–37] The method was also successfully applied to ground-[38] and excited-state[39] CC2 calculations. In the CD approach proposed by Koch *et al.*,[40,41] the four-center ERI tensor is decomposed as a product of triangular matrices neglecting the columns/rows that give negligible contributions. The efficiency of the approximation was also demonstrated for CC2.[42–44] The LT approximation developed by Almlöf and Häser[45–47] eliminates the orbital energy denominators appearing in many-body methods like CC2. Together with the DF and further approximations, it can be efficiently used for the scaling reduction of the CC2 method.[48,49]

The computational costs of CC2 calculations can also be brought down by local approximations pioneered by Pulay and co-workers.[50,51] Although the literature of local correlation methods is intensively growing, applications to excited states are relatively scarce. The first excited-state local approach was proposed by Crawford *et al.*,[52,53] who generalized the ground-state local CCSD method of Werner and co-workers[54] to EOM-CCSD. Subsequently several excited-state CC methods utilizing local approximations were published by Korona, Schütz, and their associates.[49,55–60] First, a local EOM-CCSD method was developed, in which the dominant configurations contributing to the given excited state were identified by analyzing the CIS wave function.[55] In later publications, the development of local CC2 methods was reported,[56] which was also extended to the calculation of molecular properties[57] and combined with LT techniques.[49,58–60] It is also worthwhile mentioning the local CC2 methods of Hättig *et al.* utilizing pair natural orbitals[61–63] and the recent local framework for calculating excitation energies (LoFEx) approach of Baudin and Kristensen,[64] as well as the chain of spheres exchange[65] and pair natural orbitals[66] CCSD method of Izsák *et al.*

A widely used technique for decreasing the costs of CC calculations, the frozen natural orbital (NO) approach, is based on the reduction of the size of the molecular orbital (MO) space where the corresponding equations are solved. In this method, a one-particle density matrix is constructed using a more approximate correlation method, the matrix is diagonalized, and the orbitals with small occupation number, i.e., eigenvalue, are dropped from the resulting NO basis.[67–69] While the approach is extensively used for ground-state CCSD and higher-order CC calculations,[70–73] only a couple of applications have been reported for excited states,[53,74] and, to the best of knowledge, no attempt has been made so far to accelerate CC2 calculations with frozen NO approximations. Motivated by the NO technique, a related approach was recently put forward by one of us for the reduction of the auxiliary function basis set employed in DF methods.[75] The so-called natural auxiliary function (NAF) approach constructs a fitting basis that is optimal for the given integrals and thereby allows us to shrink the basis to as small as possible. It was successfully applied to the cost reduction of the second-order Møller–Plesset (MP2) method and the direct random-phase approximation (dRPA)[75] also in combination with local correlation approximations.[76–78] Although it was anticipated in Ref. 75 that significant speedup can also be achieved for CC2 calculations, the NAF technique has not yet been utilized for the CC2 method.

In this paper, we report the development of an efficient and robust reduced-cost LR-CC2 method which is based on the combination of the NO and NAF approaches. The errors introduced are analyzed in detail, and the applicability of the method is demonstrated for representative examples.

II. THEORY

A. The linear-response CC2 approach

The CC2 correlation energy for a restricted Hartree–Fock (HF) reference can be expressed as

$$\Delta E_{\mathrm{CC2}} = \sum_{aibj} [2(ai|bj) - (aj|bi)](t_{ij}^{ab} + t_i^a t_j^b), \tag{1}$$

where $i, j, \ldots$ $(a, b, \ldots)$ denote the occupied (virtual) spatial orbitals, and $(ai|bj)$ is a two-electron integral in the Mulliken notation. The equations for the t_i^a single and t_{ij}^{ab} double excitation amplitudes were derived retaining the CCSD singles equations, while approximating the doubles equations to be correct through first order with assuming the singles to be zeroth-order parameters.[12] The resulting equations read as

$$\Omega_{\mu_1} = \langle \mu_1 | \hat{H} + [\hat{H}, T_2] | \mathrm{HF} \rangle = 0, \tag{2}$$

$$\Omega_{\mu_2} = \langle \mu_2 | \hat{H} + [F, T_2] | \mathrm{HF} \rangle = 0, \tag{3}$$

where Ω is the CC2 residual, $|\mathrm{HF}\rangle$ denotes the HF determinant, and $|\mu_n\rangle$ stands for n-fold excited determinants. $T_2 = \frac{1}{4} \sum_{aibj} t_{ij}^{ab} a^+ b^+ i^- j^-$ is the cluster operator of double excitations with a^+ and i^- as the creation and annihilation operators corresponding to MOs a and i, respectively. For convenience, the T_2 operator together with the single excitation cluster operator, $T_1 = \sum_{ai} t_i^a a^+ i^-$, will be denoted in a general form as $T_n = \frac{1}{(n!)^2} \sum_{\mu_n} t_{\mu_n} \tau_{\mu_n}$, where t_{μ_n} is the cluster amplitude associated with the τ_{μ_n} excitation operator. In the above equations, F is the Fockian, while $\hat{H}$ is the similarity-transformed Hamiltonian, which is obtained from the original Hamiltonian H

as

$$\hat{H} = \mathrm{e}^{-T_1} H \mathrm{e}^{T_1}. \tag{4}$$

The T_1-transformed two-electron MO integrals can be expressed in a closed form as

$$(pq\hat{|}rs) = \sum_{tuxy} (\mathbf{1} - \mathbf{t}_1^{\mathrm{T}})_{tp} (\mathbf{1} + \mathbf{t}_1)_{uq} (\mathbf{1} - \mathbf{t}_1^{\mathrm{T}})_{xr} (\mathbf{1} + \mathbf{t}_1)_{ys} (tu|xy), \tag{5}$$

where $p, q, \ldots$ are the generic orbital indices, and $\boldsymbol{t}_1$ is an $n_\mathrm{b} \times n_\mathrm{b}$ matrix with n_b as the size of the basis. The elements of matrix $\mathbf{t}_1$ are zero except for its virtual-occupied block, where they are equal to the t_{μ_1} amplitudes.

The simple equations for the doubles amplitudes, Eq. (3), can be regarded as the doubles equations of the MP2 method with an effective Hamiltonian, and the doubles amplitudes can simply be expressed as

$$t_{ij}^{ab} = \frac{(ai\hat{|}bj)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} = \frac{(ai\hat{|}bj)}{D_{ij}^{ab}}, \tag{6}$$

where ε_p is the corresponding orbital energy. Substituting these into Eq. (2), it is obvious that the doubles amplitudes can be calculated on-the-fly, and their storage can be avoided.[15] In practice, the equations for the single excitation amplitudes,

$$\Omega_{ai} = \sum_{bjc} (jb\hat{|}ac)\hat{t}_{ij}^{cb} - \sum_{bjk} (jb\hat{|}ik)\hat{t}_{kj}^{ab} + \sum_{bj} \hat{F}_{jb} \hat{t}_{ij}^{ab} + \hat{F}_{ai} + (\varepsilon_a - \varepsilon_i) t_i^a = 0, \tag{7}$$

are iterated, where $\hat{t}_{ij}^{ab} = 2t_{ij}^{ab} - t_{ji}^{ab}$, and the matrix elements of the similarity-transformed Fock-operator can be calculated as

$$\hat{F}_{jb} = \sum_{ai} [2(bj|ai) - (aj|bi)] t_i^a, \tag{8}$$

$$\hat{F}_{ai} = \sum_{bj} [2(ai|jb) - (ab|ji)] t_j^b - \sum_{bjk} [2(bj|ik) - (bk|ij)] t_k^a t_j^b + \sum_{bjc} [2(bj|ca) - (ba|cj)] t_j^b t_i^c - \sum_{bjck} [2(ck|bj) - (bk|cj)] t_k^c t_i^b t_j^a. \tag{9}$$

Introducing the DF approximation, the four-center ERIs can be written in the

$$(pq|rs) = \sum_{PQ} (pq|P)(P|Q)^{-1}(Q|rs) \tag{10}$$

form, where P and Q stand for the elements of the auxiliary basis, whereas $(pq|P)$ and $(P|Q)$ are the three- and two-center Coulomb integrals, respectively, and $(P|Q)^{-1}$ is a simplified notation for the corresponding element of the inverse of the two-center Coulomb integral matrix. In practice, the latter inverse matrix is factorized and rewritten, e.g., as the product of the inverse square root matrices. The four-center ERIs can then be recast in the

$$(pq|rs) = \sum_{Q} J_{pq}^{Q} J_{rs}^{Q} \tag{11}$$

form with

$$J_{pq}^{Q} = \sum_{P} (pq|P)(P|Q)^{-1/2}. \tag{12}$$

Alternatively, we can also use the CD of the inverse matrix for the factorization. The similarity-transformed two-electron integrals, $(pq\hat{|}rs)$, can also be written in a form similar to Eq. (11) since the singles amplitudes can be absorbed into the elements of the **J** list as

$$\hat{J}_{pq}^{Q} = \sum_{rs} (\mathbf{1} - \mathbf{t}_1^{\mathrm{T}})_{rp} (\mathbf{1} + \mathbf{t}_1)_{sq} J_{rs}^{Q}. \tag{13}$$

Note that the occupied-virtual block of $\hat{\mathbf{J}}$ is not affected by the similarity transformation, that is, $\hat{J}_{ia}^{P} = J_{ia}^{P}$, while the other blocks loose their permutational symmetries because $\hat{J}_{pq}^{P} \neq \hat{J}_{qp}^{P}$ any more.

Our implementation of the ground-state CC2 equations is based on the algorithm of Hättig and co-workers.[15] Since our intention is to decrease the costs of the CC2 method partly by reducing the MO space, the important difference is that our implementation uses MO integrals as opposed to the atomic orbital-based algorithm of Ref. 15. This required the slight modification of the working equations and the algorithm, which are presented in Table I together with the scaling of the various steps given in terms of the number of occupied (n_occ) and virtual (n_virt) orbitals and the number of auxiliary functions (n_aux). The rate-determining step of the procedure is still the construction of the $(ai\hat{|}bj)$ integral list (step 6) and the contraction of the doubles amplitudes (step 7), the operation count of which scales as $n_\mathrm{occ}^2 n_\mathrm{virt}^2 n_\mathrm{aux}$.

Concerning excited states, the CC LR theory calculates the excitation energies as the eigenvalues of the Jacobi-matrix which is defined by the derivative of the residual with respect to the cluster amplitudes as

$$A_{\mu_i,\nu_j} = \frac{\partial \Omega_{\mu_i}}{\partial t_{\nu_j}} = \begin{pmatrix} \langle \mu_1 | [\hat{H}, \tau_{\nu_1}] + [[\hat{H}, \tau_{\nu_1}], T_2] | \mathrm{HF} \rangle & \langle \mu_1 | [\hat{H}, \tau_{\nu_2}] | \mathrm{HF} \rangle \\ \langle \mu_2 | [\hat{H}, \tau_{\nu_1}] | \mathrm{HF} \rangle & \delta_{\mu_2\nu_2} \varepsilon_{\mu_2} \end{pmatrix}, \tag{14}$$

where $\varepsilon_{\mu_2} = -D_{ij}^{ab}$. Similar to the ground-state problem, the storage of the doubles amplitudes can also be avoided at the solution of the eigenvalue equation of matrix **A** if an effective Jacobian,

$$A_{\mu_1\nu_1}^{\mathrm{eff}}(\omega_{\mathrm{CC2}}) = A_{\mu_1\nu_1} - \frac{A_{\mu_1\gamma_2} A_{\gamma_2\nu_1}}{\varepsilon_{\gamma_2} - \omega_{\mathrm{CC2}}}, \tag{15}$$

is introduced, where ω_{CC2} is the CC2 excitation energy.[15] The pseudo-eigenvalue equation of matrix $\mathbf{A}^{\mathrm{eff}}$,

$$\sigma(\omega_{\mathrm{CC2}}, \mathbf{r}) = \mathbf{A}^{\mathrm{eff}}(\omega_{\mathrm{CC2}})\mathbf{r} = \omega_{\mathrm{CC2}}\mathbf{r}, \tag{16}$$

is only solved in the space spanned by the single excitations, whose coefficients are included in vector **r**. The working equations for the σ vector can be derived by straightforward differentiation of Eq. (7) with respect to t_{μ_1}. The structure of the resulting expressions is very similar to that of the corresponding ground-state terms, thus, the excited-state equations can be solved by defining intermediates of the same type. Only the number of intermediates will be higher since each term in Eq. (7) including n pieces of t_{μ_1} amplitudes results in n terms of similar structure upon differentiation. Consequently, the algorithm presented in Table I is also applicable *mutatis mutandis* to the evaluation of the σ vector.

TABLE I. Working equations and their scaling for the CC2 algorithm.

Step	Scaling
1 Construct intermediates $\hat{X}^Q_{ai}$ and $\hat{X}^Q_{ij}$	
$\hat{X}^Q_{ai} = \sum_b J^Q_{ab} t^b_i$	$n_{\mathrm{occ}} n^2_{\mathrm{virt}} n_{\mathrm{aux}}$
$\hat{X}^Q_{ij} = \sum_a J^Q_{ai} t^a_j$	$n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
2 Compute the $\hat{J}^Q_{ji}$ and $\hat{J}^Q_{ai}$ lists and the contribution of one term in $\hat{F}_{ai}$ to Ω_{ai}	
$\hat{J}^Q_{ij} = \hat{X}^Q_{ij} + J^Q_{ij}$	$n^2_{\mathrm{occ}} n_{\mathrm{aux}}$
$\Omega_{ai} \leftarrow -\sum_{Qj} \hat{X}^Q_{aj} \hat{J}^Q_{ji}$	$n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
$\hat{J}^Q_{ai} = \hat{X}^Q_{ai} + J^Q_{ai} - \sum_j \hat{J}^Q_{ij} t^a_j$	$2n_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}} + n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
3 Construct intermediate X^Q	
$X^Q = \sum_{ai} J^Q_{ai} t^a_i$	$n_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
4 Compute further contributions of $\hat{F}_{ai}$ to Ω_{ai}	
$\Omega_{ai} \leftarrow \sum_{Qj} (\sum_k \hat{X}^Q_{jk} t^a_k) \hat{J}^Q_{ji}$	$2n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
$\Omega_{ai} \leftarrow \sum_Q \hat{J}^Q_{ai} X^Q + t^a_i (\varepsilon_a - \varepsilon_i)$	$n_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}} + n_{\mathrm{occ}} n_{\mathrm{virt}}$
5 Calculate $\hat{F}_{jb}$	
$\hat{F}_{jb} = \sum_Q \hat{J}^Q_{bj} X^Q - \sum_{Qi} \hat{X}^Q_{ji} J^Q_{bi}$	$n_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}} + n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
6 Construct the $(ai\hat{\vert}bj)$ list	
$(ai\hat{\vert}bj) = \sum_Q \hat{J}^Q_{ai} \hat{J}^Q_{bj}$	$n^2_{\mathrm{occ}} n^2_{\mathrm{virt}} n_{\mathrm{aux}}$
7 Compute and contract $\hat{t}^{ab}_{ij}$, calculate the contributions of $\hat{F}_{jb}$ to Ω_{ai}	
$\hat{t}^{ab}_{ij} = [2(ai\hat{\vert}bj) - (aj\hat{\vert}bi)]/D^{ab}_{ij}$	$n^2_{\mathrm{occ}} n^2_{\mathrm{virt}}$
$\hat{Y}^Q_{ai} = \sum_{bj} \hat{t}^{ab}_{ij} J^Q_{bj}$	$n^2_{\mathrm{occ}} n^2_{\mathrm{virt}} n_{\mathrm{aux}}$
$\Omega_{ai} \leftarrow \sum_{bj} \hat{t}^{ab}_{ij} \hat{F}_{jb}$	$n^2_{\mathrm{occ}} n^2_{\mathrm{virt}}$
8 Calculate further contributions of $\hat{t}^{ab}_{ij}$ to Ω_{ai}	
$\Omega_{ai} \leftarrow \sum_{Qb} (\sum_j J^Q_{bj} t^a_j + J^Q_{ba}) \hat{Y}^Q_{bi} = \sum_{Qb} \hat{J}^Q_{ba} \hat{Y}^Q_{bi}$	$2n_{\mathrm{occ}} n^2_{\mathrm{virt}} n_{\mathrm{aux}} + n^2_{\mathrm{virt}} n_{\mathrm{aux}}$
$\Omega_{ai} \leftarrow -\sum_{Qj} \hat{J}^Q_{ji} \hat{Y}^Q_{aj}$	$n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$
9 Calculate CC2 correlation energy	
$\Delta E_{\mathrm{CC2}} = \sum_{Qai} (\hat{J}^Q_{ai} + X^Q t^a_i - \sum_j \hat{X}^Q_{ij} t^a_j) \hat{Y}^Q_{ai}$	$4n_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}} + n^2_{\mathrm{occ}} n_{\mathrm{virt}} n_{\mathrm{aux}}$

An important feature of matrix $\mathbf{A}^{\mathrm{eff}}$ is that it also depends on the excitation energy of the given excited state, and hence, the eigenvalue equation cannot be solved simultaneously for all the considered states using the conventional Davidson-type diagonalization techniques. We implemented a modified version of the algorithm proposed by Hättig and co-workers.[15] In the first step, an integral direct, multi-state DF-CIS calculation is run, and the corresponding CIS excitation energies and vectors will be used as initial guess in the subsequent calculations. Then, for each excited state, the eigenvector is preoptimized by a modified Davidson algorithm using the root-following technique[79] to avoid convergence to a wrong solution. Finally, the non-linear problem is solved by a direct inversion in the iterative subspace (DIIS) algorithm, which is also applied to solve the ground-state CC2 equations. Our experience showed that in particular cases, the above algorithm finds the wrong root, that is, we arrive at the same CC2 solution starting from different CIS eigenvectors. We could overcome this problem by projecting out the single excitation part of the CC2 eigenvectors converged previously from the initial guess CIS vectors.

B. Natural orbitals

In the NO approximation, a lower-level wave function is chosen, let us denote it by Ψ, and the virtual-virtual block of the one-particle density matrix,

$$D_{ab} = \langle \Psi | a^+ b^- | \Psi \rangle, \tag{17}$$

as well as the occupied-occupied block of the hole density matrix,

$$\overline{D}_{ij} = \langle \Psi | i^- j^+ | \Psi \rangle, \tag{18}$$

are constructed. The eigenvectors of these matrices are referred to as the virtual (VNOs) and occupied NOs (ONOs), respectively, whereas the corresponding eigenvalues are interpreted as the occupation numbers of the orbitals. It is supposed that the orbitals of low population give a small contribution to the correlation and can be dropped. The corresponding truncation thresholds for the virtual and occupied NO sets will be denoted by $\varepsilon_{\mathrm{VNO}}$ and $\varepsilon_{\mathrm{ONO}}$, respectively.

For the construction of the NOs for ground-state CC2 calculations, the MP2 density matrix is a plausible choice. For excited states, the selection of an appropriate wave function Ψ is less trivial. Here the CIS(D) approach seems to be a good choice since it is an excited-state generalization of the MP2 method and accounts for about the same amount of correlation for the excited state as the MP2 method does for the ground state. Moreover, the calculation of CIS(D) density matrices is relatively cheap, and the computation time required is similar to that of a CC2 iteration. Of course, we should keep in mind that the ground-state CC2 amplitudes are also necessary to solve the equations for the excited states, and it is highly desirable to solve both equations in the same MO basis. For this reason, we construct a state-specific MO basis for each excited state in which both the ground- and excited-state equations are solved. Consequently, a single, "state-averaged" (hole) density matrix is required for each excited state for the construction

of the NOs. For this purpose, we simply take the average of the MP2 ($\mathbf{D}^{\mathrm{MP2}}$) and CIS(D) ($\mathbf{D}^{\mathrm{CIS(D)}}$) densities, i.e., we diagonalize the $\mathbf{D} = (\mathbf{D}^{\mathrm{MP2}} + \mathbf{D}^{\mathrm{CIS(D)}})/2$ density matrices, and the similar expression holds for the hole densities. After the diagonalization of the (hole) density matrices and the selection of the NOs of sufficiently high occupation number, the truncated NO basis is canonicalized so that we will be able to employ the equations derived for canonical MOs. The elements of the latter basis will be denoted by $\tilde{a}$, $\tilde{i}$, ..., and $\tilde{\mathbf{J}}$ will stand for the $\mathbf{J}$ list, the MO indices of which are transformed to the canonicalized NO basis.

The required blocks of the one-particle MP2 (hole) density matrices can be calculated as

$$\overline{D}_{ij}^{\mathrm{MP2}} = \sum_{kab} \left(2T_{ik}^{ab} T_{jk}^{ab} - T_{ik}^{ab} T_{jk}^{ba} \right), \tag{19}$$

$$D_{ab}^{\mathrm{MP2}} = \sum_{ijc} \left(2T_{ij}^{ca} T_{ij}^{cb} - T_{ij}^{ca} T_{ij}^{bc} \right), \tag{20}$$

where T_{ij}^{ab} is a first-order doubles amplitude defined as

$$T_{ij}^{ab} = \frac{(ai|bj)}{D_{ij}^{ab}}. \tag{21}$$

The algorithm for the construction of MP2 densities is presented in Table II. The bottleneck is the assembly of the $(ai|bj)$ integral list, which scales as $n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2 n_{\mathrm{aux}}$. There are two other fifth-power scaling operations in the algorithm. The evaluation of the virtual-virtual block of the density matrix scaling as $n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^3$ is also relatively expensive, while the computation time needed for the calculation of the occupied-occupied block of the hole density matrix is in general considerably shorter. Please note that the MP2 densities must be constructed only once, independently of the number of excited states.

The CIS(D) density matrix can be split up as the sum of the CIS density matrix and the contribution of the perturbative (D) correction, $\mathbf{D}^{\mathrm{CIS(D)}} = \mathbf{D}^{\mathrm{CIS}} + \mathbf{D}^{(\mathrm{D})}$. The CIS density matrices read as

$$\overline{D}_{ij}^{\mathrm{CIS}} = \sum_{a} c_i^a c_j^a, \tag{22}$$

$$D_{ab}^{\mathrm{CIS}} = \sum_{i} c_i^a c_i^b, \tag{23}$$

TABLE II. Working equations and their scaling for the evaluation of MP2 density matrices.

Step	Scaling
1 Assemble $(ai\|bj)$ list	
$(ai\|bj) = \sum_Q J_{ai}^Q J_{bj}^Q$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2 n_{\mathrm{aux}}$
2 Compute T_{ij}^{ab}	
$T_{ij}^{ab} = (ai\|bj)/D_{ij}^{ab}$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2$
3 Compute intermediate X_{ij}^{ab}	
$X_{ij}^{ab} = 2T_{ij}^{ab} - T_{ij}^{ba}$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2$
4 Compute density matrices	
$\overline{D}_{ij}^{\mathrm{MP2}} = \sum_{kab} T_{ik}^{ab} X_{jk}^{ab}$	$n_{\mathrm{occ}}^3 n_{\mathrm{virt}}^2$
$D_{ab}^{\mathrm{MP2}} = \sum_{ijc} T_{ij}^{ca} X_{ij}^{cb}$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^3$

where c_i^a is a CIS coefficient, while we define $\mathbf{D}^{(\mathrm{D})}$ analogously to its MP2 counterpart as

$$\overline{D}_{ij}^{(\mathrm{D})} = \sum_{kab} (2c_{ik}^{ab} c_{jk}^{ab} - c_{ik}^{ab} c_{jk}^{ba}), \tag{24}$$

$$D_{ab}^{(\mathrm{D})} = \sum_{ijc} (2c_{ij}^{ca} c_{ij}^{cb} - c_{ij}^{ca} c_{ij}^{bc}) \tag{25}$$

with c_{ij}^{ab} as a CIS(D) doubles coefficient. The latter were given by Head-Gordon and co-workers[20] in a spin-orbital basis. The corresponding expression for spatial orbitals can easily be derived by spin-integration and exploiting the relationships of doubles amplitudes[80] and reads as

$$c_{ij}^{ab} = \frac{\sum_c [(ac|bj)c_i^c + (ac|bi)c_j^c] - \sum_k [(kj|ai)c_k^b + (kj|bi)c_k^a]}{D_{ij}^{ab} + \omega_{\mathrm{CIS}}}, \tag{26}$$

where ω_{CIS} stands for the CIS excitation energy. The calculation of CIS(D) densities follows the route shown in Table III. As it can be seen, similar to the evaluation of the MP2 densities, the rate-determining operation is the construction of a four-index intermediate as a product of three-index quantities, which scales as $n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2 n_{\mathrm{aux}}$. However, in addition to the inexpensive CIS contribution, there are two further fourth-power scaling terms (step 2) in contrast to the MP2 density matrices, where the occupied-occupied and virtual-virtual blocks of the three-center integral list are contracted with CIS coefficients. Since the CIS(D) coefficients, and thus the density matrices, are state-dependent, these quantities have to be evaluated for each excited state.

Our numerical experience showed that further approximations can be introduced at the construction of the $\mathbf{D}^{\mathrm{MP2}}$ and $\mathbf{D}^{\mathrm{CIS(D)}}$ matrices. As it will be demonstrated in Sec. III B, the calculated excitation energies hardly change if we neglect the exchange-like contributions to the density matrices, that is, instead of Eqs. (19), (20), (24), and (25), we define the density

TABLE III. Working equations and their scaling for the evaluation of CIS(D) density matrices.

Step	Scaling
1 Construct CIS contribution to density matrices	
$\overline{D}_{ij}^{\mathrm{CIS(D)}} \leftarrow \sum_a c_i^a c_j^a$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}$
$D_{ab}^{\mathrm{CIS(D)}} \leftarrow \sum_i c_i^a c_i^b$	$n_{\mathrm{occ}} n_{\mathrm{virt}}^2$
2 Compute intermediate Y_{ai}^Q	
$Y_{ai}^Q \leftarrow \sum_c J_{ac}^Q c_i^c$	$n_{\mathrm{occ}} n_{\mathrm{virt}}^2 n_{\mathrm{aux}}$
$Y_{ai}^Q \leftarrow -\sum_j J_{ij}^Q c_j^a$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}} n_{\mathrm{aux}}$
3 Compute intermediate V_{ij}^{ab}	
$V_{ij}^{ab} = \sum_Q Y_{ai}^Q J_{bj}^Q$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2 n_{\mathrm{aux}}$
4 Calculate c_{ij}^{ab} coefficients	
$c_{ij}^{ab} = (V_{ij}^{ab} + V_{ji}^{ba})/(D_{ij}^{ab} + \omega_{\mathrm{CIS}})$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2$
5 Compute intermediate X_{ij}^{ab}	
$X_{ij}^{ab} = 2c_{ij}^{ab} - c_{ij}^{ba}$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2$
6 Compute density matrices	
$\overline{D}_{ij}^{\mathrm{CIS(D)}} \leftarrow \sum_{kab} c_{ik}^{ab} X_{jk}^{ab}$	$n_{\mathrm{occ}}^3 n_{\mathrm{virt}}^2$
$D_{ab}^{\mathrm{CIS(D)}} \leftarrow \sum_{ijc} c_{ij}^{ca} X_{ij}^{cb}$	$n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^3$

matrices as

$$\overline{D}_{ij}^{\mathrm{MP2}} = \sum_{kab} 2T_{ik}^{ab} T_{jk}^{ab}, \tag{27}$$

$$D_{ab}^{\mathrm{MP2}} = \sum_{ijc} 2T_{ij}^{ca} T_{ij}^{cb}, \tag{28}$$

$$\overline{D}_{ij}^{(\mathrm{D})} = \sum_{kab} 2c_{ik}^{ab} c_{jk}^{ab}, \tag{29}$$

$$D_{ab}^{(\mathrm{D})} = \sum_{ijc} 2c_{ij}^{ca} c_{ij}^{cb}. \tag{30}$$

The advantage of these equations is that the evaluation of intermediates X_{ij}^{ab} (see Tables II and III) can be avoided, and the calculation of the density matrices can be reduced to the multiplication of a matrix with its transpose, whose operation can be efficiently performed. Furthermore, a simple and fast algorithm can be designed for the evaluation of the above equations if the T_{ij}^{ab} or c_{ij}^{ab} list does not fit into the memory. In this case, at the construction of the MP2 densities (see Table II), one of the virtual indices of the doubles amplitudes is split up into blocks, hereafter denoted by $\mathcal{A}$, so that the amplitudes with the remaining three indices can be kept in the memory. The outmost loop in the algorithm runs over $\mathcal{A}$, and in the inner loops, the doubles amplitudes with one virtual index in the given index range are calculated, and their contribution to the density matrices is computed.

The evaluation of the CIS(D) density matrices deserves somewhat more attention (see Table III) since the size of the J_{ab}^{Q} list can also easily exceed the size of the available memory. In this case, the list is blocked according to its auxiliary index so that the J_{ab}^{Q} integrals with Q in a block can be held in the memory, and the first contribution to intermediate Y_{ai}^{Q} (step 2) is evaluated within a loop performed over the blocks. Analogously to the MP2 algorithm, one of the virtual indices of the four-index intermediates is also restricted, and the subsequent steps (steps 3 to 6) are carried out in a loop running over the corresponding $\mathcal{A}$ blocks. A further difficulty is that the symmetrization of intermediate V_{ij}^{ab} is necessary to obtain the c_{ij}^{ab} coefficients (step 4), and it requires the recalculation of particular V_{ij}^{ab} matrix elements if the out-of-core algorithm is executed.

C. Natural auxiliary functions

With the aid of the NAF approach,[75] the size of the auxiliary basis set can be reduced in a way reminiscent of the NO technique via a rank reduction of the corresponding three-center integral matrix $\mathbf{J}$. If the latter is regarded as a two-index matrix with composite row index pq and column index Q, we can, in principle, perform its singular value decomposition (SVD) as

$$\mathbf{J} = \mathbf{M}\Sigma\mathbf{N}^{\mathrm{T}}, \tag{31}$$

where $\mathbf{M}$ and $\mathbf{N}$ are the matrices composed of the left and right singular vectors (SVs), and Σ is a diagonal matrix with the singular values on its diagonal. In practice, we only need the singular values and right SVs of $\mathbf{J}$, which are identical, respectively, to the square root of the eigenvalues and eigenvectors of matrix

$$\mathbf{W} = \mathbf{J}\mathbf{J}^{\mathrm{T}} \tag{32}$$

and can be more efficiently obtained by diagonalizing $\mathbf{W}$ than by performing the SVD of matrix $\mathbf{J}$. The eigenvectors of matrix $\mathbf{W}$, that is, the right SVs of $\mathbf{J}$, can be considered as the linear combination coefficients of new auxiliary functions, which we call the NAFs. Similar to the NO approach, the NAFs with eigenvalues lower than a threshold ($\varepsilon_{\mathrm{NAF}}$) are dropped, and the remaining $\overline{n}_{\mathrm{aux}}$ pieces of eigenvectors are collected in matrix $\overline{\mathbf{N}}$. If the auxiliary function index of matrix $\mathbf{J}$ is now transformed by matrix $\overline{\mathbf{N}}$ as

$$\overline{\mathbf{J}} = \mathbf{J}\overline{\mathbf{N}}, \tag{33}$$

the resulting $\overline{\mathbf{J}}$ matrix is the best rank $\overline{n}_{\mathrm{aux}}$ approximation of $\mathbf{J}$ in the least-squares sense.

The most computation-intensive operation of a CC2 calculation scales as $n_{\mathrm{occ}}^2 n_{\mathrm{virt}}^2 \overline{n}_{\mathrm{aux}}$ (see Table I), but the scaling of the other expensive terms is also linear in $\overline{n}_{\mathrm{aux}}$. Furthermore, the storage requirements also grow linearly with the number of fitting functions. Consequently, using a NAF basis, the decrease in the computation time is expected to be proportional to the $\overline{n}_{\mathrm{aux}}/n_{\mathrm{aux}}$ ratio.

An important question is at what stage the NAF basis should be constructed and what $\mathbf{J}$ should be used for that purpose. In principle, we can change for a NAF basis right at the beginning of the calculation, prior to the evaluation of the MP2 and CIS(D) density matrices. The advantage of this strategy would be that the expenses of the latter steps would also be reduced. However, as it is discussed in Ref. 75, the NAF basis is an auxiliary basis that is optimal for the given MO basis, and hence, it is expedient to construct the NAFs from the $\mathbf{J}$ matrix whose MO indices are already transformed to the final MO basis, in our case from matrix $\tilde{\mathbf{J}}$. The second question is what block of $\tilde{\mathbf{J}}$ is to be used for the determination of the NAFs. Of course, the NAFs constructed from the entire $\tilde{\mathbf{J}}$ will be the exact right SVs of the matrix. However, we can also build approximate NAF bases from particular blocks of $\tilde{\mathbf{J}}$, for instance, we can compute matrix $\mathbf{W}$ only from the $\tilde{J}_{ai}^{Q}$ or $\tilde{J}_{pi}^{Q}$ integrals. In this way, we can reduce the time spent on the construction of the NAF basis, but probably more NAFs have to be retained to obtain results of similar accuracy as the NAFs computed, e.g., from the $\tilde{J}_{ai}^{Q}$ list will be lower-quality fitting functions for the $(ab|ij)$-type integrals. These aspects will be analyzed in Sec. III B.

D. General algorithm

To conclude this section, we overview our general algorithm for the present reduced-cost LR-CC2 approach, which is as follows:

1. Solve HF equations
2. Solve CIS equations for all the excited states using a multi-state, integral-direct algorithm
3. Loop over excited states
   - 3.a. Compute state-averaged one-particle density matrices (Tables II and III)
   - 3.b. Transform the MO indices of $\mathbf{J}$ to the pseudo-canonical NO basis
   - 3.c. Evaluate matrix $\mathbf{W}$ from $\tilde{\mathbf{J}}$ [Eq. (32)], select NAFs to be retained

3.d. Transform the auxiliary function index of $\tilde{\mathbf{J}}$ to the NAF basis [Eq. (33)]
3.e. Save integral list $\bar{\mathbf{J}}$ to disk
End loop

4. Loop over excited states
4.a. Retrieve integral list $\bar{\mathbf{J}}$ from disk
4.b. Solve the ground- and excited-state CC2 equations (Table I)
End loop

Note that the CC2 equations are solved in a different loop than the one in which the densities, the NAF basis, and the transformed integrals are constructed. The rationale behind this is that, in the first loop, it enables the design of an algorithm where the original $\mathbf{J}$ list is read from a disk or recalculated only once for the integral transformation. The transformed $\bar{\mathbf{J}}$ integral list is saved to the disk and retrieved in the second loop. However, the size of the $\bar{\mathbf{J}}$ lists is significantly, by more than an order of magnitude smaller than that of matrix $\mathbf{J}$, thus the reading and processing of the $\bar{\mathbf{J}}$ list do not cause any complication in the CC2 calculations.

III. RESULTS

A. Computational details

The proposed reduced-cost CC2 approach has been implemented in the Mrcc suite of quantum chemical programs and is available in the current release of the package.[81]

For the test calculations, Dunning's correlation consistent basis sets augmented with diffuse functions (aug-cc-pVXZ, where X = D, T, Q) were used,[82–84] and the corresponding auxiliary bases developed by Weigend *et al.* were employed in both the HF and the subsequent calculations.[85–87] In the CIS and CC2 calculations, the core orbitals were kept frozen.

For monitoring the convergence of the excitation energies with the truncation thresholds, a test set of small molecules was assembled including acetone, acetamide, benzene, butadiene, formaldehyde, and formamide from the benchmark set of Thiel and co-workers,[88,89] as well as the ethylene-tetrafluoroethylene system (at separations of 5 and 10 Å) from Dreuw *et al.*[90] We endeavored to consider all important types of excitations. The valence excited states ($n \rightarrow \pi^*$, $\sigma \rightarrow \pi^*$, and $\pi \rightarrow \pi^*$, including also the $\pi \rightarrow \pi^*$ states of the conjugated butadiene and the aromatic benzene) were those identified by Thiel and co-workers;[88,89] to consider CT excitations, the ones assigned by Dreuw *et al.*[90] for the ethylene-tetrafluoroethylene system were chosen, whereas the Rydberg states of formamide and acetamide were selected by us.

To also gain some insight into the performance of the new method for medium-sized molecules, another test set was also used, which is representative of all the important excitations types as well. The geometries of 6,6′-difluoro-indigo,[91] a bithiophene derivative,[28] and N-methyl-2,3-benzcarbaxole[28] were taken from the literature, while those for the remaining azobenzene, diphenylamine, hydrazone dye, and 4-(*N*,*N*-dimethylamino)-3-hydroxyflavone molecules were optimized with the Mrcc program at the DFT level applying the B3LYP[92] functional and the cc-pVDZ basis set. Further benchmark calculations were carried out for even larger systems to demonstrate the applicability of the approach. In these calculations, the bisimide derivative studied by Grimme and co-workers[91] as well as the phenothiazine-isoalloxazine dyad,[56] two borondipyrromethene-flavin derivatives [Flv(a) and Flv(b)],[57] and the cyanoacrylic acid derivative (D21L6)[58] test system of Schütz *et al.* was considered. The structural formulas of the molecules and the coordinates for the molecules optimized by us are available in the supplementary material.

The reported computation times are wall-clock times measured on a computer with 128 GB of main memory and a 6-core 3.5 GHz Intel Xeon E5-1650 processor.

B. Convergence with the truncation thresholds

First, we analyze the convergence of the CC2 excitation energies with the cutoff parameters, $\varepsilon_{\mathrm{VNO}}$ and $\varepsilon_{\mathrm{NAF}}$, of the VNOs and NAFs for the small-molecule test set using the aug-cc-pVTZ basis. The effect of the truncation of the ONO space will be studied later for bigger systems because even for the largest molecules (benzene and the ethylene-tetrafluoroethylene model) considered here the number of occupied orbitals is rather small and cannot be decreased. We focus on the results obtained with the aug-cc-pVTZ basis since at least triple-ζ basis sets including diffuse functions are required for semi-quantitative excitation energies, while the quadruple-ζ and bigger basis sets are usually not worth using together with the simplistic CC2 approach and are not practical for large molecules. Nevertheless, results calculated with double- and quadruple-ζ basis sets will also be presented later.

To study the convergence with the number of virtual orbitals, the VNO space constructed by diagonalizing the density matrices including (Eqs. (20) and (25)) or excluding (Eqs. (28) and (30)) the exchange-like terms were gradually decreased, and the error with respect to the canonical CC2 excitation energy was calculated. The results are visualized in Fig. 1. Inspecting the plots we can observe that the error introduced by dropping the VNOs is moderate and only slightly increases in most cases up to a truncation of about 60%, where the mean absolute error (MAE) obtained with density matrices including (excluding) the exchange-like terms is only 0.016 (0.019) eV, and the maximum absolute error (MAX) still does not exceed 0.044 (0.060) eV. For the $\pi \rightarrow \pi^*$, Rydberg, and CT excitations, the error starts increasing steeply at around 60%, while for the excited states of other types, the results deteriorate more slowly. The increase of the errors, apart from small fluctuations on the meV scale, is monotonic up to 60%; however, the sign of the errors can differ even for different states of the same molecule or different excited states of the same type. The difference of the results obtained with and without the exchange contributions is very small. The average difference is 2 to 3 meV for any truncation of the VNO space with the S_2 state of acetamide exhibiting the largest deviation, 0.030 eV, at the truncation of 73%, which causes too large error anyway to be used in practical applications. As a consequence, the exchange-excluding density matrices can safely be employed.

The behavior of the excitation energies as a function of the size of the NAF basis was analyzed similarly, first without any truncation of the VNO space. The NAFs constructed from the

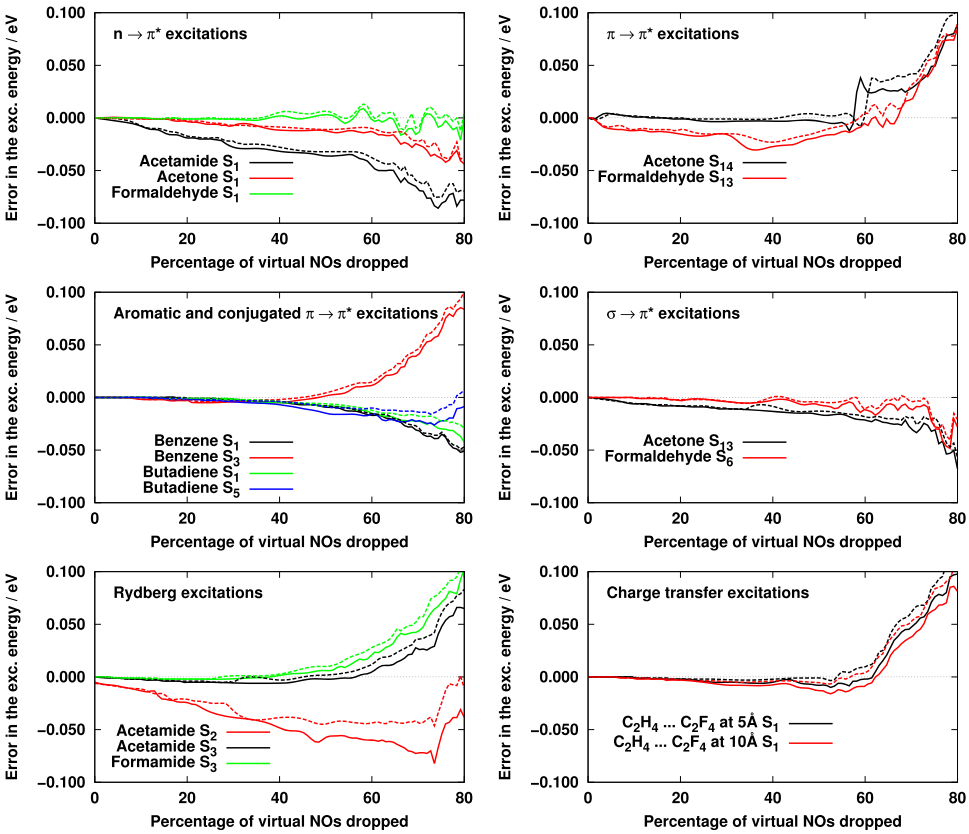

FIG. 1. Error of CC2 excitation energies as a function of the number of VNOs dropped with the aug-cc-pVTZ basis set. Dashed (solid) line: results obtained with density matrices including (excluding) the exchange-like contributions.

J^Q_{ai}, J^Q_{pi}, and J^Q_{pq} lists were tested, however, the results calculated using the first integral list are relatively poor and not discussed here. The errors of the approximate CC2 excitation energies with respect to the canonical ones are displayed in Fig. 2. As it can be seen, in contrast to the VNOs, the NAF approximation introduces a very small error up to 40% truncation, where the MAE (MAX) is 0.002 (0.004) eV with NAFs derived from the J^Q_{pi} list, while the corresponding value for the case of the

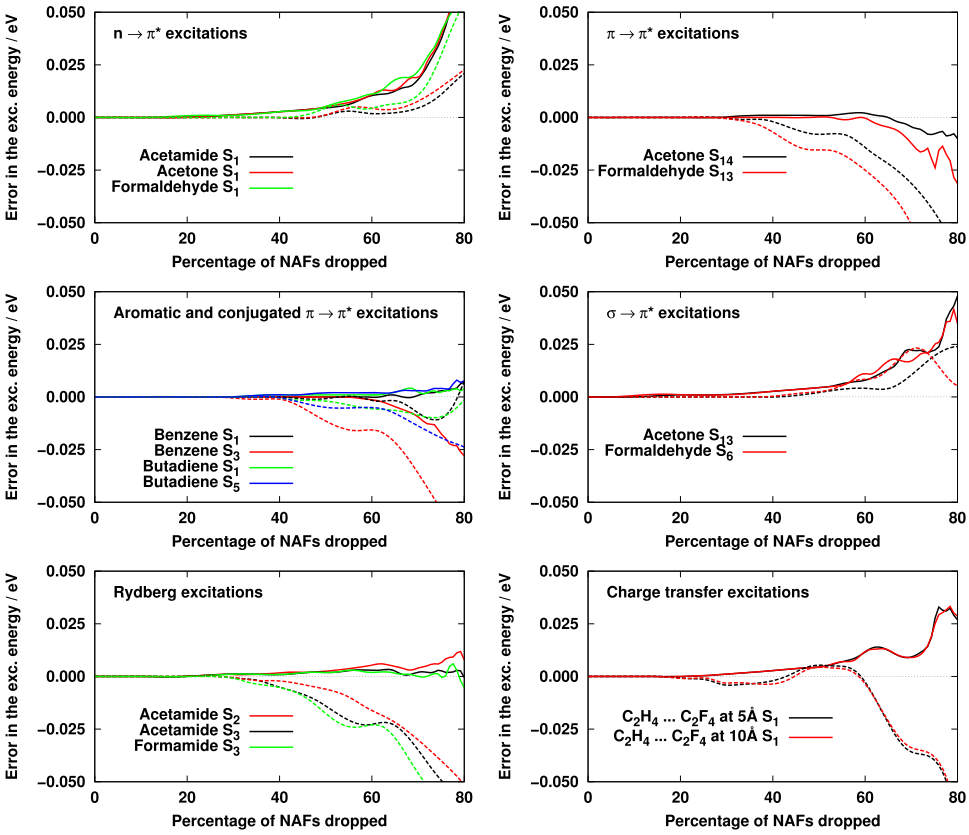

FIG. 2. Error of CC2 excitation energies as a function of the number of NAFs dropped with the aug-cc-pVTZ basis set. Dashed (solid) line: results obtained with **W** matrices calculated from the J^Q_{pi} (J^Q_{pq}) lists.

J^Q_{pq} list is just 0.001 (0.002) eV. Dropping further NAFs, the approach based on the J^Q_{pq} integrals is obviously more robust, with the exception of the $n \rightarrow \pi^*$ excitations, and the error for the latter scheme is significantly lower than with the approach using the J^Q_{pi} list. At the truncation of 60%, the MAE (MAX) is 0.005 (0.011) eV with the J^Q_{pq}-based and 0.010 (0.024) eV with the J^Q_{pi}-based approach. Beyond 60%, the quality of the results drops rapidly for most types of excited states, especially with the latter scheme. These results suggest that it is worthwhile constructing the NAF basis from the J^Q_{pq} list even if $n^2_{\mathrm{virt}} n^2_{\mathrm{aux}}$ additional operations are required to calculate matrix $\mathbf{W}$ [Eq. (32)] with respect to the case when $\mathbf{W}$ is computed just from the J^Q_{pi} integrals.

Since our intention is to combine the NO and NAF approaches, it is also important to study their joint effect, which also facilitates the selection of reliable default values for our truncation thresholds. As it is discussed in Sec. II C, if the NO and NAF approximations are applied simultaneously, it is more beneficial to construct first the NO basis then the NAFs. Accordingly, taking into account the above results, we first calculate the VNOs from the density matrices excluding the exchange terms and truncate the VNO space using threshold $\varepsilon_{\mathrm{VNO}}$. Then, we construct the NAF basis with the $\tilde{J}^Q_{\tilde{p}\tilde{q}}$ integral list and truncate it according to $\varepsilon_{\mathrm{NAF}}$. At the determination of the default truncation threshold, our main purpose was to develop a robust method and to keep the MAE of the approach about an order of magnitude smaller than the intrinsic error of the CC2 method, that is, 0.2 to 0.3 eV.[93,94] We tested numerous combinations of the $\varepsilon_{\mathrm{VNO}}$ and $\varepsilon_{\mathrm{NAF}}$ thresholds with the considered values corresponding close to 60% truncations in Figs. 1 and 2. Relying on the results of these numerical experiments, which are presented in the supplementary material, we propose $\varepsilon_{\mathrm{VNO}} = 7.5 \times 10^{-5}$ and $\varepsilon_{\mathrm{NAF}} = 0.1$ a.u. as the default values of the cutoff parameters.

The errors of the computed excitation energies and the number of the neglected VNOs and NAFs with the default thresholds are presented in Table IV for the small-molecule test. Looking at the results we can conclude that the MAE (MAX) of the joint NO and NAF approach is 0.023 (0.052) eV, which is considerably lower than the error of the CC2 method itself. The percentage of the dropped VNOs is rather system independent and falls in the narrow range of 56% to 65%. The ratio of the neglected NAFs is even more stable, and it fluctuates in an interval of about 5%. As expected, in the truncated NO basis, considerably fewer auxiliary functions are necessary, and in average, 82.3% of the NAFs can be removed. Being aware of these results, we can envisage a speedup of about 35 for the rate-determining step of the CC2 calculations scaling as $n^2_{\mathrm{occ}} n^2_{\mathrm{virt}} n_{\mathrm{aux}}$, and a reduction of also about 35-times is foreseen in the size of the largest quantity to be stored, the virtual-virtual block of matrix $\mathbf{J}$.

TABLE IV. Canonical CC2 excitation energies (in eV), the error of CC2 excitation energies computed with the present approach (in eV), and the percentage of VNOs and NAFs dropped using the default thresholds with the aug-cc-pVTZ basis set for small molecules.

Molecule	Character	State	ω_{CC2}	Error	Dropped VNOs	Dropped NAFs
Acetamide	$n \rightarrow \pi^*$	S_1	5.605	−0.033	63.6	82.9
	Rydberg	S_2	5.917	−0.031	61.5	82.7
	Rydberg	S_3	6.456	0.034	61.8	82.6
Acetone	$n \rightarrow \pi^*$	S_1	4.454	−0.001	65.0	83.4
	$\sigma \rightarrow \pi^*$	S_{13}	9.110	−0.007	64.7	83.6
	$\pi \rightarrow \pi^*$	S_{14}	9.212	0.052	62.4	83.4
Benzene	$\pi \rightarrow \pi^*$	S_1	5.220	−0.017	63.4	84.0
	$\pi \rightarrow \pi^*$	S_3	6.452	0.035	63.9	84.5
Butadiene	$\pi \rightarrow \pi^*$	S_1	6.134	−0.009	63.2	83.7
	$\pi \rightarrow \pi^*$	S_5	7.064	−0.004	63.8	84.6
Formaldehyde	$n \rightarrow \pi^*$	S_1	3.996	0.041	58.8	80.3
	$\sigma \rightarrow \pi^*$	S_6	9.191	0.010	59.2	80.6
	$\pi \rightarrow \pi^*$	S_{13}	10.698	0.022	57.7	81.2
Formamide	Rydberg	S_3	6.697	0.034	58.5	80.9
CT[a] at 5Å	CT	S_1	6.728	0.013	56.1	79.5
CT[a] at 10Å	CT	S_1	6.740	0.020	56.8	79.7
Average				0.023[b]	61.3	82.3
Maximum				0.052[c]	65.0	84.6
Minimum				0.001[d]	56.1	79.5

[a]Ethylene-tetrafluoroethylene model.
[b]Mean absolute error.
[c]Maximum absolute error.
[d]Minimum absolute error.

To gain some impression of the basis set dependence of the results, we also performed test calculations for the small-molecule test set with the aug-cc-pVDZ and aug-cc-pVQZ basis sets using the default thresholds ($\varepsilon_{\mathrm{VNO}} = 7.5 \times 10^{-5}$ and $\varepsilon_{\mathrm{NAF}} = 0.1$ a.u.). The error measures, such as the MAE, MAX, and the minimum absolute error (MIN), for the excitation energies are summarized in Table V, while the detailed results can be found in the supplementary material. As we can presume, with the double-ζ basis in average, a smaller number of VNOs are dropped (32.9%) than with the triple-ζ one (61.3%), while with the quadruple-ζ basis, more VNOs are neglected (78.5%). The tendency is similar but not so pronounced for the NAFs, where the average number of dropped functions only changes with a couple of percents when increasing the cardinal number of the basis set. The MAE and MAX are about

TABLE V. Error measures for the CC2 excitation energies (in eV) and the average, maximum, minimum percentage of VNOs and NAFs dropped using the default thresholds with various basis sets for small molecules.

	Error			Dropped VNOs			Dropped NAFs		
Basis	MAE	MAX	MIN	Avg.	Max.	Min.	Avg.	Max.	Min.
aug-cc-pVDZ	0.051	0.100	0.022	32.9	39.2	25.0	81.4	83.1	80.0
aug-cc-pVTZ	0.023	0.052	0.001	61.3	65.0	56.1	82.3	84.6	79.5
aug-cc-pVQZ	0.047	0.106	0.004	78.5	81.0	75.8	88.6	90.5	86.7

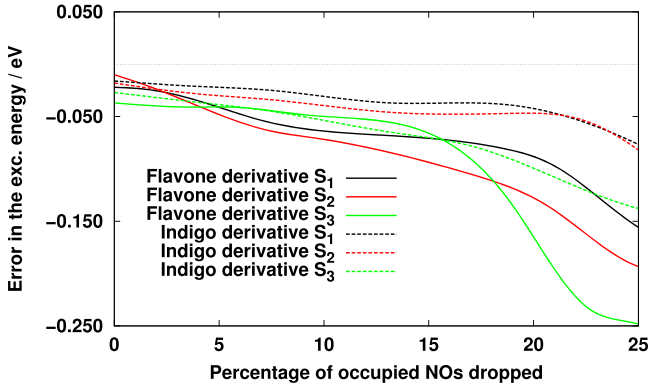

FIG. 3. Error of CC2 excitation energies as a function of the number of ONOs dropped with the aug-cc-pVTZ basis set using $\varepsilon_{\mathrm{VNO}} = 7.5 \times 10^{-5}$ and $\varepsilon_{\mathrm{NAF}}$ = 0.1 a.u. thresholds.

twice as large with aug-cc-pVDZ and aug-cc-pVQZ basis sets than with aug-cc-pVTZ but are still moderate and remarkably lower than the intrinsic error of CC2. We note that with the aug-cc-pVQZ basis set, the truncation of the VNO and NAF spaces is expected to result in a speedup of about 190 for the most expensive operations and a reduction of about 190-times in the size of the integral list.

We also investigated the truncation of the ONO basis. To this intent, we selected two bigger molecules, 6,6′-difluoro-indigo and 4-(*N*,*N*-dimethylamino)-3-hydroxyflavone, for which the number of correlated occupied orbitals is sufficiently high, 53 for both molecules. Fixing the VNO and NAF truncation thresholds to their default values, the number of neglected ONOs was systematically varied. The errors of the CC2 excitation energies with respect to the canonical ones are shown in Fig. 3. The error that arose at the 10% reduction of the ONO space is in average as large as the combined error of the VNO and NAF approximations (0.02 eV, see Table VI) even if the latter spaces are much more aggressively truncated. Thus, at around 10%, the overall error of the three approximations is about 0.05 eV, but it starts growing rapidly after the truncation of 15%. It means that only a couple of occupied orbitals can be dropped, and unfortunately the gain in the computation time is rather limited. The approximation is probably more efficient for large systems, for which, however, CC2 calculations are currently not feasible even with the present reduced-cost scheme. Therefore, in the following test calculations, the occupied MO space will not be truncated.

C. Representative examples

Our approach was further tested for bigger molecules to check if the conclusions drawn for the small systems are also valid in the general case. The set includes dye molecules which are well-established test molecules from the literature.[30,95,96] The 6,6′-difluoro-indigo, the bithiophene derivative, and the *N*-methyl-2,3-benzcarbaxole molecules stem from the test suite of Grimme and co-workers;[28,91,97] the azobenzene, the diphenylamine, and the hydrazone dye were collected from the work of Matsuura *et al.*;[98] whereas the 4-(*N*,*N*-dimethylamino)-3-hydroxyflavone molecule featuring a dominant CT excited state was taken from the studies of Mély and co-workers.[99,100] Applying our default settings, for each molecule the six lowest singlet states were considered, among which all important types of excited state can be found including several Rydberg and two CT states. The numerical results are collected in Table VI.

The MAE (MAX) of the computed excitation energies with respect to the canonical results is 0.022 (0.048) eV, which is practically identical to the corresponding value, 0.023 (0.051) eV, obtained for the small-molecule test set. The same holds for the average number of the neglected VNOs and NAFs, which are 61.1% and 82.6%, respectively, for these molecules, while for the smaller ones, the 61.3% and 82.3% values were found. The spread of the results is also similar for the smaller and the bigger systems: the deviations of the minimum and maximum number of dropped VNOs and NAFs are 6% and 4%, respectively, which should be compared to the differences of 9% and 5% obtained for the small-molecule test set. These data support that the errors introduced by our approximations are rather stable and independent of the system size, and our approach with the default settings can be applied as a black-box method for arbitrary excited states of extended systems.

As it has been discussed above, a theoretical speedup of 35 can be expected in the most expensive step of a CC2 iteration with the aug-cc-pVTZ basis, however, the overhead due to evaluation of the density matrices is notable. To estimate the overall speedup factors for our approach, we measured the wall-clock times for the canonical CC2 calculations including the CIS iterations, the solution of the ground-state CC2, and the six response equations. These values were divided by the wall times measured for the entire reduced-cost CC2 calculations excluding the solution of the HF equations, that is, the wall time spent on the CIS calculation, the construction of the MP2 and the six CIS(D) density matrices, the integral transformations, reading and writing of the $\bar{\mathbf{J}}$ integral list, and the solution of the ground- and excited-state CC2 equations for each excited state. The calculated speedups are also reported in Table VI. The average speedup with respect to the canonical run is about 18, but in particular cases much larger increase in the speed of the calculation can be observed. Of course, we would await that the speedup factors vary in a tighter interval since the number of the neglected VNOs and NAFs does so. However, the number of iterations needed to converge the CC2 equations can significantly change when working in the reduced spaces. Anyway, the computation time is reduced by an order of magnitude even in the worst cases, and it is easy to see that the speedup is even better if more excited states are targeted.

To further demonstrate the applicability of our approach, benchmark calculations were carried out for even larger systems. The considered molecules, that is, the bisimide derivative;[28,91,96] the flavin derivatives, Flv(a) and Flv(b);[57] the phenothiazine-isoalloxazine dyad;[49,56,58,64] and the cyanoacrylic acid derivative D21L6[58,59] were previously studied in the literature at the CC2 level using small bases, however, CC2 calculations with the aug-cc-pVTZ or similar basis sets have not been possible so far. The results are compiled in Table VII.

The computed excitation energies and the assignation of the excited states are in line with the results available in the literature. The percentage of the VNOs and NAFs dropped

TABLE VI. Canonical CC2 excitation energies (in eV), the error of CC2 excitation energies computed with the present approach (in eV), and the percentage of VNOs and NAFs dropped using the default thresholds with the aug-cc-pVTZ basis set for bigger molecules.

Molecule	Number of atoms	Number of basis functions	State	Character	ω_{CC2}	Error	Dropped VNOs	Dropped NAFs	Speedup
Hydrazone dye	21	828	S_1	$\pi \rightarrow \pi^*$	3.562	–0.012	58.1	80.7	14.1
			S_2	CT	3.924	–0.016	59.0	80.9	
			S_3	$n, \sigma \rightarrow \pi^*$	3.947	–0.015	58.9	80.8	
			S_4	$\pi \rightarrow \pi^*$	4.180	–0.019	57.8	80.4	
			S_5	$n, \sigma \rightarrow \pi^*$	4.527	–0.011	58.9	80.8	
			S_6	$n, \sigma \rightarrow \pi^*$	4.543	–0.006	58.8	80.8	
Diphenylamine	24	851	S_1	$\pi \rightarrow \pi^*$	4.334	–0.015	63.3	83.9	17.9
			S_2	$\pi \rightarrow \pi^*$	4.460	–0.024	63.3	83.7	
			S_3	Rydberg	4.627	0.014	63.4	83.8	
			S_4	Rydberg	4.994	0.048	63.3	83.9	
			S_5	Rydberg	5.084	0.037	63.2	84.0	
			S_6	Rydberg	5.229	0.026	63.2	83.9	
Azobenzene	24	874	S_1	$n, \sigma \rightarrow \pi^*$	2.785	0.014	63.7	83.6	12.3
			S_2	$\pi \rightarrow \pi^*$	4.071	–0.018	62.6	83.3	
			S_3	$\pi \rightarrow \pi^*$	4.427	–0.029	62.5	83.2	
			S_4	$\pi \rightarrow \pi^*$	4.435	–0.029	62.6	83.3	
			S_5	$\pi \rightarrow \pi^*$	5.155	–0.042	62.2	83.1	
			S_6	Rydberg	6.057	0.021	62.0	83.5	
6,6′-difluoro-indigo	28	1150	S_1	$\pi \rightarrow \pi^*$	2.952	–0.016	58.3	80.9	12.8
			S_2	$\pi \rightarrow \pi^*$	3.436	–0.018	58.2	80.8	
			S_3	$\pi \rightarrow \pi^*$	3.684	–0.027	58.2	80.9	
			S_4	$\pi \rightarrow \pi^*$	3.725	–0.027	58.2	80.9	
			S_5	$\pi \rightarrow \pi^*$	4.274	–0.023	58.1	80.9	
			S_6	$\pi \rightarrow \pi^*$	4.775	–0.037	58.1	80.9	
Bithiophene derivative	29	1135	S_1	$\pi \rightarrow \pi^*$	3.843	–0.026	59.4	82.3	17.8
			S_2	$\pi \rightarrow \pi^*$	4.631	–0.025	59.4	82.3	
			S_3	$\pi \rightarrow \pi^*$	4.940	–0.034	59.2	82.2	
			S_4	$\pi \rightarrow \pi^*$	5.446	0.002	59.2	82.4	
			S_5	Rydberg	5.538	0.009	59.1	82.3	
			S_6	$\pi \rightarrow \pi^*$	5.602	–0.027	59.2	82.3	
N-methyl-2,3-benzcarbaxole	31	1127	S_1	$\pi \rightarrow \pi^*$	3.332	–0.013	63.3	83.6	20.5
			S_2	$\pi \rightarrow \pi^*$	4.062	–0.021	63.4	83.6	
			S_3	Rydberg	4.468	0.020	63.1	83.8	
			S_4	$\pi \rightarrow \pi^*$	4.510	–0.037	63.1	83.6	
			S_5	Rydberg	4.835	0.011	63.2	83.9	
			S_6	Rydberg	4.898	–0.017	63.2	83.9	
Hydroxyflavone derivative	36	1311	S_1	CT	3.382	–0.022	63.0	83.1	28.1
			S_2	$\pi \rightarrow \pi^*$	4.124	–0.030	63.0	83.1	
			S_3	$n, \sigma \rightarrow \pi^*$	4.154	–0.020	63.0	83.2	
			S_4	$\pi \rightarrow \pi^*$	4.227	–0.010	63.1	83.2	
			S_5	Rydberg	4.330	0.018	62.7	83.3	
			S_6	Rydberg	4.866	–0.037	62.6	83.3	
Average						0.022[a]	61.1	82.6	17.6
Maximum						0.048[b]	63.7	84.0	28.1
Minimum						0.002[c]	57.8	80.4	12.3

[a]Mean absolute error.
[b]Maximum absolute error.
[c]Minimum absolute error.

is still close to what we obtain for the smaller molecules but seems to slightly increase with the size of the system. The basis set includes at least 2000 atomic orbitals for all the molecules, and the number of basis functions far exceeds 3000 for the most extended system. Looking at the wall times required for the CIS and CC2 iterations, we can see that the expenses of the solution of the CC2 equations in the reduced spaces are similar to those of the initial canonical CIS calculation. In addition to the calculation of the state-averaged density matrices, two integral transformations and, for the construction of the NAF basis, the evaluation of matrix **W** are required. The time of these operations amounts to 10% to 15% of the computation time spent on the calculation of the densities. Our timings show that the excitation energy of a single excited state can be

TABLE VII. CC2 excitation energies computed with the present approach (in eV), the percentage of VNOs and NAFs dropped, and computation times (in min) using the default thresholds with the aug-cc-pVTZ basis set for big molecules.

Molecule	Number of atoms	Number of basis functions	State	Character	ω_{CC2}	Dropped VNOs	Dropped NAFs	t_{CIS}[a]	$t_{\bar{\mathbf{J}}}$[b]	t_{CC2}[c]
Flv(a)	51	2001	S1	$\pi \rightarrow \pi^*$	2.750	60.9	82.1	15.5	81.6	3.7
			S2	$\pi \rightarrow \pi^*$	2.995	60.8	82.0		81.6	3.6
			S3	$n, \sigma \rightarrow \pi^*$	3.339	61.1	82.1		81.6	3.4
			S4	$\pi \rightarrow \pi^*$	3.613	60.9	82.0		81.6	3.5
Dyad	53	2051	S1	$\pi \rightarrow \pi^*$	3.063	61.4	82.7	15.4	84.1	3.5
			S2	CT	3.373	61.7	82.8		84.1	3.5
			S3	$\pi \rightarrow \pi^*$	3.506	61.2	82.6		84.1	3.4
			S4	CT	3.826	61.2	82.7		84.1	3.5
Bisimide derivative	60	2346	S1	$\pi \rightarrow \pi^*$	2.543	61.5	82.5	24.6	179.3	7.0
			S2	$\pi \rightarrow \pi^*$	3.438	61.5	82.6		178.7	6.7
			S3	$\pi \rightarrow \pi^*$	3.707	61.5	82.6		179.2	6.4
			S4	$\pi \rightarrow \pi^*$	3.793	61.5	82.6		178.9	6.4
Flv(b)	78	2829	S1	$\pi \rightarrow \pi^*$	2.603	63.8	83.3	53.9	416.3	12.9
			S2	$\pi \rightarrow \pi^*$	2.978	63.9	83.3		416.2	12.9
			S3	$n, \sigma \rightarrow \pi^*$	3.279	64.0	83.4		416.2	13.0
			S4	$\pi \rightarrow \pi^*$	3.329	63.8	83.3		415.9	13.0
D21L6	98	3412	S1	CT	2.622	65.2	84.4	42.8	1093.2	29.1
			S2	$\pi \rightarrow \pi^*$	3.442	65.1	84.3		1092.9	30.0
Average						62.3	82.9			
Maximum						65.2	84.4			
Minimum						60.8	82.0			

[a]Average wall time for a CIS iteration using a multi-state, integral-direct algorithm.
[b]Average wall time for the construction of the $\bar{\mathbf{J}}$ integral list. The calculation of the MP2 density is distributed equally among the excited states, that is, $t_{\bar{\mathbf{J}}} = t_{\mathbf{CIS(D)}} + t_{\mathbf{MP2}}/n + t_{\mathbf{W}} + t_{\mathrm{Int}}$, where $t_{\mathbf{CIS(D)}}$, $t_{\mathbf{MP2}}$, and $t_{\mathbf{W}}$ are the wall times required for the evaluation of the CIS(D) and MP2 density matrices and the $\mathbf{W}$ matrix, respectively, n stands for the number of excited states, and t_{Int} is the wall time required for the integral transformations.
[c]Average wall time for a CC2 iteration.

computed within less than a day for a molecule of more than 80 atoms with around 3000 basis functions.

IV. CONCLUSIONS

An efficient approach has been presented for the calculation of CC2 excitation energies. Our scheme successfully combines the natural orbital and natural auxiliary function approximations. For the calculation of the NOs, state-averaged density matrices are constructed using approximate MP2 and CIS(D) densities for the ground and excited state, respectively. For the calculation of NAFs, the use of the entire integral list turned out to be the most advantageous choice.

The errors introduced by the various approximations have been analyzed for a representative test set including all important types of excitations. The thresholds for the selection of NOs and NAFs have been determined. Our results suggest that the average error obtained with the proposed cutoff parameters is low, 0.02 eV, with a triple-ζ basis and practically independent of the system size. In average, 60% and 80% of the virtual NOs and NAFs can be dropped, respectively, and these numbers also seem system-independent. For these calculations, the CC2 iterations are by about a factor of 35 faster, and even if the evaluation of the density matrices requires some extra time, the overall speedup is of more than an order of magnitude.

Our results demonstrate that the proposed approach enables the routine use of the CC2 method for molecules of more than 50 atoms with basis sets including more than 2000 basis functions. We have also carried out CC2 excitation energy calculations with reliable basis sets for systems of about 100 atoms on a single workstation computer. Such calculations were previously not possible within an acceptable time frame.

SUPPLEMENTARY MATERIAL

See supplementary material for molecular structures and geometries and excitation energies computed with various thresholds and basis sets.

ACKNOWLEDGMENTS

The authors are indebted to Professor Péter G. Szalay for useful discussions. The computing time granted on the Hungarian HPC Infrastructure at NIIF Institute, Hungary, is gratefully acknowledged.

[1]J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[2]H. J. Monkhorst, Int. J. Quantum Chem. **S11**, 421 (1977).
[3]E. Dalgaard and H. J. Monkhorst, Phys. Rev. A **28**, 1217 (1983).
[4]H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).
[5]H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **93**, 3345 (1990).
[6]H. Sekino and R. J. Bartlett, Int. J. Quantum Chem. **S18**, 255 (1984).
[7]J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).
[8]M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).
[9]R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[10]R. J. Bartlett, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **3**, 126 (2012).

[11] K. Sneskov and O. Christiansen, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 566 (2011).
[12] O. Christiansen, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **243**, 409 (1995).
[13] K. Hald, C. Hättig, D. L. Yeager, and P. Jørgensen, Chem. Phys. Lett. **328**, 291 (2000).
[14] O. Christiansen, H. Koch, P. Jørgensen, and T. Helgaker, Chem. Phys. Lett. **263**, 530 (1996).
[15] C. Hättig and F. Weigend, J. Chem. Phys. **113**, 5154 (2000).
[16] C. Hättig, J. Chem. Phys. **118**, 7751 (2003).
[17] A. Köhn and C. Hättig, J. Chem. Phys. **119**, 5021 (2003).
[18] D. Kánnár and P. G. Szalay, J. Chem. Theory Comput. **10**, 3757 (2014).
[19] D. Kánnár, A. Tajti, and P. G. Szalay, J. Chem. Theory Comput. **13**, 202 (2017).
[20] M. Head-Gordon, R. J. Rico, M. Oumi, and T. J. Lee, Chem. Phys. Lett. **219**, 21 (1994).
[21] C. Hättig, Adv. Quantum Chem. **50**, 37 (2015).
[22] J. Schirmer, Phys. Rev. A **26**, 2395 (1982).
[23] J. Schirmer and A. B. Trofimov, J. Chem. Phys. **120**, 11449 (2004).
[24] M. Wormit, D. R. Rehn, P. H. P. Harbach, J. Wenzel, C. M. Krauter, E. Epifanovsky, and A. Dreuw, Mol. Phys. **112**, 774 (2014).
[25] M. E. Casida, "Recent advances in density functional methods," in *Computational Chemistry: Reviews of Current Trends*, edited by D. P. Chong (World Scientific, Singapore, 1999), Vol. 1.
[26] A. Dreuw and M. Head-Gordon, Chem. Rev. **105**, 4009 (2005).
[27] S. Grimme and M. Parac, ChemPhysChem **4**, 292 (2003).
[28] L. Goerigk and S. Grimme, J. Chem. Phys. **132**, 184103 (2010).
[29] N. O. C. Winter, N. K. Graf, S. Leutwyler, and C. Hättig, Phys. Chem. Chem. Phys. **15**, 6623 (2013).
[30] D. Jacquemin, I. Duchemin, and X. Blase, J. Chem. Theory Comput. **11**, 5340 (2015).
[31] S. F. Boys, G. B. Cook, C. M. Reeves, and I. Shavitt, Nature **178**, 1207 (1956).
[32] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
[33] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
[34] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
[35] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **137**, 044103 (2012).
[36] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, J. Chem. Phys. **137**, 224106 (2012).
[37] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martínez, J. Chem. Phys. **137**, 221101 (2012).
[38] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **138**, 124111 (2013).
[39] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez, J. Phys. Chem. B **117**, 12972 (2013).
[40] H. Koch and A. M. Sánchez de Merás, J. Chem. Phys. **113**, 508 (2000).
[41] H. Koch, A. Sánchez de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).
[42] T. B. Pedersen, A. M. Sánchez de Merás, and H. Koch, J. Chem. Phys. **120**, 8887 (2004).
[43] P. Baudin, J. Sánchez de Marín, I. García Cuesta, and A. M. Sánchez de Merás, J. Chem. Phys. **140**, 104111 (2014).
[44] D. R. Nascimento and A. E. DePrince III, J. Chem. Theory Comput. **12**, 5834 (2016).
[45] J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).
[46] M. Häser and J. Almlöf, J. Chem. Phys. **96**, 489 (1992).
[47] M. Häser and J. Almlöf, J. Chem. Phys. **110**, 3660 (1999).
[48] N. O. C. Winter and C. Hättig, J. Chem. Phys. **134**, 184101 (2011).
[49] D. Kats and M. Schütz, J. Chem. Phys. **131**, 124117 (2009).
[50] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[51] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).
[52] T. D. Crawford and R. A. King, Chem. Phys. Lett. **366**, 611 (2002).
[53] A. Kumar and T. D. Crawford, J. Phys. Chem. A **121**, 708 (2017).
[54] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[55] T. Korona and H.-J. Werner, J. Chem. Phys. **118**, 3006 (2003).
[56] D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **125**, 104106 (2006).
[57] D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **127**, 064107 (2007).
[58] K. Freundorfer, D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **133**, 244110 (2010).
[59] K. Ledermüller, D. Kats, and M. Schütz, J. Chem. Phys. **139**, 084111 (2013).
[60] K. Ledermüller and M. Schütz, J. Chem. Phys. **140**, 164113 (2014).
[61] B. Helmich and C. Hättig, J. Chem. Phys. **135**, 214106 (2011).
[62] B. Helmich and C. Hättig, J. Chem. Phys. **139**, 084114 (2012).
[63] B. Helmich and C. Hättig, Comput. Theor. Chem. **1040-1041**, 35 (2014).
[64] P. Baudin and K. Kristensen, J. Chem. Phys. **144**, 224106 (2016).
[65] A. K. Dutta, F. Neese, and R. Izsák, J. Chem. Phys. **144**, 034102 (2016).
[66] A. K. Dutta, F. Neese, and R. Izsák, J. Chem. Phys. **145**, 034102 (2016).
[67] W. Meyer, J. Chem. Phys. **58**, 1017 (1973).
[68] P. J. Hay, J. Chem. Phys. **59**, 2468 (1973).
[69] R. Ahlrichs, H. Lischka, V. Staemmler, and W. Kutzelnigg, J. Chem. Phys. **62**, 1225 (1975).
[70] A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128**, 164101 (2008).
[71] A. E. DePrince and C. D. Sherrill, J. Chem. Theory Comput. **9**, 293 (2013).
[72] Z. Rolik and M. Kállay, J. Chem. Phys. **134**, 124111 (2011).
[73] Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, and M. Kállay, J. Chem. Phys. **139**, 094105 (2013).
[74] A. Landau, K. Khistyaev, S. Dolgikh, and A. I. Krylov, J. Chem. Phys. **132**, 014109 (2010).
[75] M. Kállay, J. Chem. Phys. **141**, 244113 (2014).
[76] M. Kállay, J. Chem. Phys. **142**, 204105 (2015).
[77] P. R. Nagy, G. Samu, and M. Kállay, J. Chem. Theory Comput. **12**, 4897 (2016).
[78] T. Kjærgaard, J. Chem. Phys. **146**, 044103 (2017).
[79] W. Butscher and W. E. Kammer, J. Comput. Phys. **20**, 313 (1975).
[80] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, J. Chem. Phys. **86**, 2881 (1987).
[81] MRCC, a quantum chemical program suite written by M. Kállay, Z. Rolik, J. Csontos, I. Ladjánszki, L. Szegedy, B. Ladóczki, G. Samu, K. Petrov, M. Farkas, P. Nagy, D. Mester, and B. Hégely, see also Ref. 73 as well as http://www.mrcc.hu/.
[82] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[83] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[84] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).
[85] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[86] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).
[87] F. Weigend, J. Comput. Chem. **29**, 167 (2008).
[88] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **128**, 134110 (2008).
[89] M. R. Silva-Junior, S. P. Sauer, M. Schreiber, and W. Thiel, Mol. Phys. **108**, 453 (2010).
[90] A. Dreuw, J. L. Weisman, and M. Head-Gordon, J. Chem. Phys. **119**, 2943 (2003).
[91] L. Goerigk, J. Moellmanna, and S. Grimme, Phys. Chem. Chem. Phys. **11**, 4611 (2009).
[92] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[93] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **256**, 185 (1996).
[94] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).
[95] D. Jacquemin, V. Wathelet, E. A. Perpéte, and C. Adamo, J. Chem. Theory Comput. **5**, 2420 (2009).
[96] D. Jacquemin, E. A. Perpéte, I. Ciofini, C. Adamo, R. Valero, Y. Zhao, and D. G. Truhlar, J. Chem. Theory Comput. **6**, 2071 (2010).
[97] M. Parac and S. Grimme, J. Phys. Chem. A **106**, 6844 (2002).
[98] A. Matsuura, H. Sato, W. Sotoyama, A. Takahashi, and M. Sakurai, J. Mol. Struct.: THEOCHEM **860**, 119 (2008).
[99] D. A. Yushchenko, O. B. Vadzyuk, S. O. Kosterin, G. Duportail, Y. Mély, and V. G. Pivovarenko, Anal. Biochem. **369**, 218 (2007).
[100] C. A. Kenfack, A. S. Klymchenko, G. Duportail, A. Burgerc, and Y. Mély, Phys. Chem. Chem. Phys. **14**, 8910 (2012).