Benchmark tests and spin adaptation for the particle-particle Random Phase Approximation

Yang Yang, Stephan N. Steinmann, and Degao Peng

Department of Chemistry, Duke University, NC 27708, U.S.

Helen van Aggelen[*]

Department of Chemistry, Duke University, NC 27708, U.S. and

Department of Inorganic and Physical Chemistry,

Ghent University, 9000 Ghent, Belgium

Weitao Yang[†]

Department of Chemistry and Department of Physics, Duke University, NC 27708, U.S.

(Dated: October 17, 2013)

1

Abstract

The particle-particle random phase approximation (pp-RPA) provides an approximation to the correlation energy in DFT via the adiabatic connection (van Aggelen, Yang, and Yang, http://arxiv.org/abs/1306.4957). It has virtually no delocalization error nor static correlation error for single-bond systems. However, with its formal $O(N^6)$ scaling, the pp-RPA is computationally expensive. In this paper, we implement a spin-separated and spin-adapted pp-RPA algorithm, which reduces the computational cost by a significant factor. We then perform benchmark tests on the G2/97 enthalpies of formation database, DBH24 reaction barrier database and four test sets for non-bonded interactions (HB6/04, CT7/04, DI6/04 and WI9/04). For the G2/97 database, the pp-RPA gives a significantly smaller mean absolute error (8.3 kcal/mol) than the direct particle-hole RPA (ph-RPA) (22.7 kcal/mol). Furthermore, the error in the pp-RPA is nearly constant with the number of atoms in a molecule, while the error in the ph-RPA keeps increasing. For chemical reactions involving typical organic closed-shell molecules, pp- and ph-RPA both give accurate reaction energies. Similarly, both RPAs perform well for reaction barriers and nonbonded interactions. These results suggest that pp-RPA gives reliable energies in chemical applications. The adiabatic connection formalism based on pairing matrix fluctuation is therefore expected to lead to widely applicable and accurate density functionals.

---

*Electronic address: `Helen.VanAggelen@UGent.be`

†Electronic address: `Weitao.Yang@duke.edu`

I. INTRODUCTION

The random phase approximation (RPA) [1] has attracted increasing interests within the quantum chemistry community during the last decade in its formulation as a density functional approximation [2–4]. The RPA has its roots in many-body theories such as Green's function theory [5, 6] or the coupled cluster theory [7, 8]. In density functional theory (DFT), the RPA represents a sophisticated functional, obtained when coupling the adiabatic connection [9, 10] with the fluctuation dissipation theorem [11]. Therefore, the RPA forms a connection between DFT and many-body methods. It is attractive for its lower computational cost ($O(N^4)$ with resolution of identity [12, 13]) compared to most correlated wave function methods, and because it overcomes some failures persistent in commonly used density functional approximations, i.e. the long-range dispersion interaction error [14, 15] and the static correlation error [16, 17].

Previously, the "RPA" mainly referred to the well-known particle-hole channel of the random phase approximation (ph-RPA) [2, 4, 18, 19], especially the direct ph-RPA without exchange, which we will simply note as ph-RPA in the rest of the paper. Recently, the particle-particle counterpart — the particle-particle random phase approximation (pp-RPA) — has been introduced to calculate the correlation energy of atomic and molecular systems [20]. By coupling the adiabatic connection with the pairing matrix fluctuation, the pp-RPA provides an approximate correlation energy [20]. The difference between ph-RPA and pp-RPA can also be viewed from a diagrammatic perspective, which identifies the ph- and the pp-RPA as the sum of all "ring" diagrams and all "ladder" diagrams, respectively [6] (Figure 1). As the summation of all ladder diagrams, the pp-RPA is equivalent to the ladder channel of coupled cluster doubles (ladder-CCD) [7, 21, 22].

The pp-RPA has many interesting features, most notably, in contrast to ph-RPA [17], it has virtually no delocalization error for general systems, in addition to virtually no static correlation error for single-bond systems. It thus satisfies the flat-plane condition [23]. This suggests that the pp-RPA can be a source of inspiration for developing density functionals. The adiabatic connection formulated in previous work [20] demonstrates that the pp-RPA may be a starting point for constructing density functionals based on the pairing matrix fluctuation. Preliminary assessments have shown that pp-RPA is at least as accurate as ph-RPA for calculating enthalpies of formation and van der Waals interactions [20]. A recent
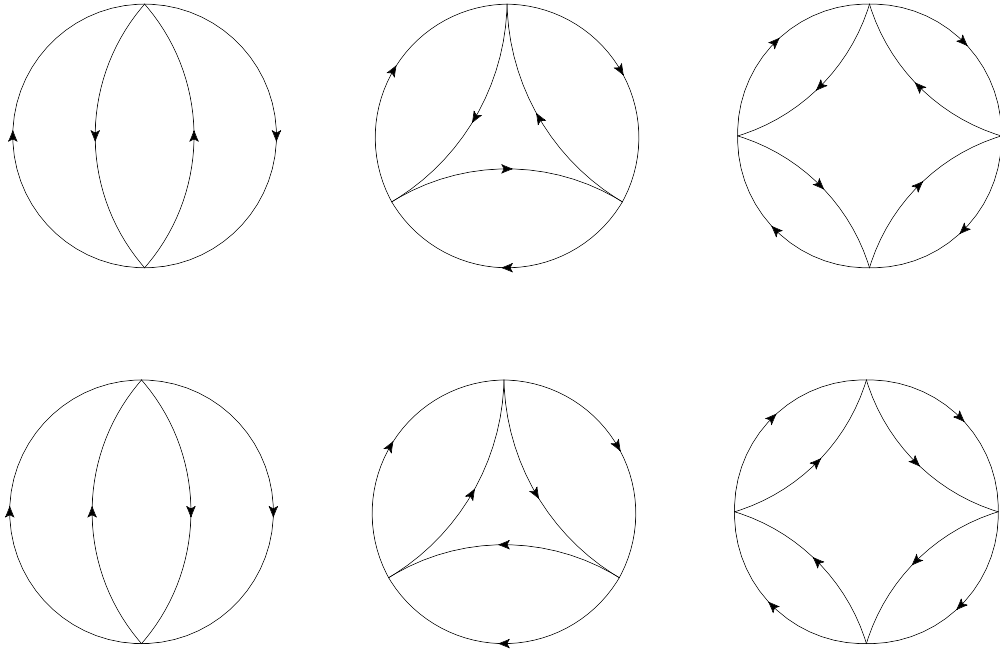
3

FIG. 1: Ring diagrams (upper row) and ladder diagrams (lower row) from the second-order to the fourth-order (left to right). With exchange, the second-order ring diagram is the same as the second-order ladder diagram and they are both exact. However, the direct ph-RPA, which we abbreviate as ph-RPA in this paper, does not contain exchange and therefore is not exact even to the second order.

work by Hesselmann [24] also elucidates the importance of third-order ladder diagrams in molecular systems. However, in order to better gauge the performance of the pp-RPA, more benchmark tests are required.

In this paper, we first derive the spin separation and the spin adaptation for the pp-RPA to alleviate the computational burden. Then we carry out benchmark tests to assess the performance of pp-RPA compared to the well-known ph-RPA. These benchmarks include the Gaussian-2/97 (G2/97) enthalpies of formation database [25, 26], the DBH24 reaction barrier database [27, 28], the HB6/04, CT7/04, DI6/04 and WI9/04 nonbonded interaction databases [29].

II. THEORY

A standard way of deriving the pp-RPA is from the adiabatic connection in terms of the pairing matrix fluctuation [20]. It enables us to adopt DFT references and gives explicitly the correlation energy expression. However, when a Hartree-Fock (HF) reference is used, we can derive using the equation of motion (EOM) method [30, 31], and in this paper, we will follow the EOM approach. Compared with the adiabatic connection formalism, the EOM method is easier for expressing spin separation and spin adaptation.

The EOM for the pp-RPA can be written as

$$[\delta O, [\hat{H}, O^\dagger]] = (E_n^{N+2} - E_0^N - 2\nu)[\delta O, O^\dagger], \tag{1}$$

where $\hat{H}$ is the Hamiltonian operator with the chemical potential taken into consideration

$$\hat{H} = \hat{h} + \hat{V} - \nu\hat{N} = \sum_{pq}\sum_{\sigma\tau}\langle p_\sigma|h - \nu N|q_\tau\rangle p_\sigma^\dagger q_\tau + \frac{1}{2}\sum_{pqrs}\sum_{\sigma\tau\zeta\theta}(p_\sigma q_\tau|V|r_\zeta s_\theta)p_\sigma^\dagger q_\tau^\dagger s_\theta r_\zeta, \tag{2}$$

with $\hat{h}$ the core Hamiltonian, $\hat{V}$ the Coulomb repulsion operator, $\hat{N}$ the electron number operator and ν the chemical potential. In this paper, we use p, q, r, s for generic orbitals, i, j, k, l for occupied orbitals, a, b, c, d for unoccupied orbitals, Greek letters σ, τ, ζ, θ, α and β for spins. $p_\sigma^\dagger$ is the creation operator of the spin orbital p_σ and q_τ is the annihilation operator of the spin orbital q_τ. We define that indices with α spins are always bigger than those with β spins. $O^\dagger$ is a two-electron addition excitation operator which is chosen to be

$$O_n^\dagger = \sum_{a_\sigma > b_\tau} X_{a_\sigma b_\tau}^n a_\sigma^\dagger b_\tau^\dagger + \sum_{i_\sigma > j_\tau} Y_{i_\sigma j_\tau}^n j_\tau^\dagger i_\sigma^\dagger, \tag{3}$$

and δO is a trial two-electron removal de-excitation operator that can be

$$\delta O_n = b_\tau a_\sigma \qquad \text{with} \quad a_\sigma > b_\tau, \tag{4}$$

or

$$\delta O_n = i_\sigma j_\tau \qquad \text{with} \quad i_\sigma > j_\tau. \tag{5}$$

For the first trial operator, when we evaluate with a Hartree-Fock (HF) ground state, the

equation of motion gives

$$\sum_{c_\zeta > d_\theta} \langle \Phi_{HF}^N | [b_\tau a_\sigma, [H, c_\zeta^\dagger d_\theta^\dagger]] | \Phi_{HF}^N \rangle X_{c_\zeta d_\theta}^n + \sum_{k_\zeta > l_\theta} \langle \Phi_{HF}^N | [b_\tau a_\sigma, [H, l_\theta^\dagger k_\zeta^\dagger]] | \Phi_{HF}^N \rangle Y_{k_\zeta l_\theta}^n$$

$$= (E_n^{N+2} - E_0^N - 2\nu) \sum_{c_\zeta > d_\theta} \langle \Phi_{HF}^N | [b_\tau a_\sigma, c_\zeta^\dagger d_\theta^\dagger] | \Phi_{HF}^N \rangle X_{c_\zeta d_\theta}^n + \sum_{k_\zeta > l_\theta} \langle \Phi_{HF}^N | [b_\tau a_\sigma, l_\theta^\dagger k_\zeta^\dagger] | \Phi_{HF}^N \rangle Y_{k_\zeta l_\theta}^n$$

$$= (E_n^{N+2} - E_0^N - 2\nu) X_{a_\sigma b_\tau}^n. \tag{6}$$

For the second trial operator, the equation of motion gives

$$\sum_{c_\zeta > d_\theta} \langle \Phi_{HF}^N | [i_\sigma j_\tau, [H, c_\zeta^\dagger d_\theta^\dagger]] | \Phi_{HF}^N \rangle X_{c_\zeta d_\theta}^n + \sum_{k_\zeta > l_\theta} \langle \Phi_{HF}^N | [i_\sigma j_\tau, [H, l_\theta^\dagger k_\zeta^\dagger]] | \Phi_{HF}^N \rangle Y_{k_\zeta l_\theta}^n$$

$$= (E_n^{N+2} - E_0^N - 2\nu) \sum_{c_\zeta > d_\theta} \langle \Phi_{HF}^N | [i_\sigma j_\tau, c_\zeta^\dagger d_\theta^\dagger] | \Phi_{HF}^N \rangle X_{c_\zeta d_\theta}^n + \sum_{k_\zeta > l_\theta} \langle \Phi_{HF}^N | [i_\sigma j_\tau, l_\theta^\dagger k_\zeta^\dagger] | \Phi_{HF}^N \rangle Y_{k_\zeta l_\theta}^n$$

$$= -(E_n^{N+2} - E_0^N - 2\nu) Y_{i_\sigma j_\tau}^n. \tag{7}$$

The above two equations can be cast into a matrix equation

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\dagger & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{X}^n \\ \mathbf{Y}^n \end{bmatrix} = \omega_n \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{X}^n \\ \mathbf{Y}^n \end{bmatrix}, \tag{8}$$

with

$$A_{a_\sigma b_\tau, c_\zeta d_\theta} = \langle \Phi_{HF}^N | \left[b_\tau a_\sigma, \left[H, c_\zeta^\dagger d_\theta^\dagger \right] \right] | \Phi_{HF}^N \rangle, \tag{9a}$$

$$B_{a_\sigma b_\tau, k_\zeta l_\theta} = \langle \Phi_{HF}^N | \left[b_\tau a_\sigma, \left[H, l_\theta^\dagger k_\zeta^\dagger \right] \right] | \Phi_{HF}^N \rangle, \tag{9b}$$

$$C_{i_\sigma j_\tau, k_\zeta l_\theta} = \langle \Phi_{HF}^N | \left[i_\sigma j_\tau, \left[H, l_\theta^\dagger k_\zeta^\dagger \right] \right] | \Phi_{HF}^N \rangle, \tag{9c}$$

and

$$\omega_n = E_n^{N+2} - E_0^N - 2\nu. \tag{10a}$$

Evaluation of the matrix elements gives

$$A_{a_\sigma b_\tau, c_\zeta d_\theta} = \delta_{a_\sigma c_\zeta} \delta_{b_\tau d_\theta} (\epsilon_{a_\sigma} + \epsilon_{b_\tau} - 2\nu) + \langle a_\sigma b_\tau || c_\zeta d_\theta \rangle, \tag{11a}$$

$$B_{a_\sigma b_\tau, k_\zeta l_\theta} = \langle a_\sigma b_\tau || k_\zeta l_\theta \rangle, \tag{11b}$$

$$C_{i_\sigma j_\tau, k_\zeta l_\theta} = -\delta_{i_\sigma k_\zeta} \delta_{j_\tau l_\theta} (\epsilon_{i_\sigma} + \epsilon_{j_\tau} - 2\nu) + \langle i_\sigma j_\tau || k_\zeta l_\theta \rangle, \tag{11c}$$

where $\langle p_\sigma q_\tau | r_\zeta s_\theta \rangle$ is defined as

$$\langle p_\sigma q_\tau | r_\zeta s_\theta \rangle \equiv \delta_{\sigma \zeta} \delta_{\tau \theta} \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_p^*(\mathbf{r}_1) \phi_q^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_r(\mathbf{r}_1) \phi_s(\mathbf{r}_2), \tag{12}$$

and $\langle p_\sigma q_\tau || r_\zeta s_\theta \rangle \equiv \langle p_\sigma q_\tau | r_\zeta s_\theta \rangle - \langle p_\sigma q_\tau | s_\theta r_\zeta \rangle$.

The correlation energy can be expressed as [20]

$$E_c = \sum_n \omega_n^+ - \text{Tr}\mathbf{A} = -\sum_n \omega_n^- - \text{Tr}\mathbf{C}, \tag{13}$$

where ω_n^+'s and ω_n^-'s are positive and negative eigenvalues, respectively. In other words, the correlation energy can be obtained by diagonalizing the matrix and summing over eigenvalues.

Note that the excitation and de-excitation operators in Eqs. (3), (4) and (5) are all based on particle or hole pairs. In terms of spin combinations, there are three different types of pairs — $\alpha\alpha$, $\beta\beta$ and $\alpha\beta$ pairs ($\beta\alpha$ pairs are absent because of index restrictions). In Eq. (11), matrix elements with different types of excitation and de-excitation pairs are all zero, which naturally leads to the spin separated pp-RPA matrix

$$\begin{bmatrix} \mathbf{A}_{\alpha\alpha,\alpha\alpha} & 0 & 0 & \mathbf{B}_{\alpha\alpha,\alpha\alpha} & 0 & 0 \\ 0 & \mathbf{A}_{\alpha\beta,\alpha\beta} & 0 & 0 & \mathbf{B}_{\alpha\beta,\alpha\beta} & 0 \\ 0 & 0 & \mathbf{A}_{\beta\beta,\beta\beta} & 0 & 0 & \mathbf{B}_{\beta\beta,\beta\beta} \\ \mathbf{B}^\dagger_{\alpha\alpha,\alpha\alpha} & 0 & 0 & \mathbf{C}_{\alpha\alpha,\alpha\alpha} & 0 & 0 \\ 0 & \mathbf{B}^\dagger_{\alpha\beta,\alpha\beta} & 0 & 0 & \mathbf{C}_{\alpha\beta,\alpha\beta} & 0 \\ 0 & 0 & \mathbf{B}^\dagger_{\beta\beta,\beta\beta} & 0 & 0 & \mathbf{C}_{\beta\beta,\beta\beta} \end{bmatrix}.$$

Then the eigenvalue problem can be decomposed into three independent eigenvalue problems

$$\begin{bmatrix} \mathbf{A}_{spin} & \mathbf{B}_{spin} \\ \mathbf{B}^\dagger_{spin} & \mathbf{C}_{spin} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{spin} \\ \mathbf{Y}_{spin} \end{bmatrix} = \omega_{spin} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{spin} \\ \mathbf{Y}_{spin} \end{bmatrix}, \tag{14}$$

with $spin = (\alpha\alpha, \alpha\alpha)$ or $(\alpha\beta, \alpha\beta)$ or $(\beta\beta, \beta\beta)$, and the matrix elements are

$$[A_{\alpha\alpha,\alpha\alpha}]_{ab,cd} \equiv A_{a_\alpha b_\alpha, c_\alpha d_\alpha} = \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b - 2\nu) + \langle ab||cd \rangle$$

$$[A_{\alpha\beta,\alpha\beta}]_{ab,cd} \equiv A_{a_\alpha b_\beta, c_\alpha d_\beta} = \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b - 2\nu) + \langle ab|cd \rangle$$

$$[A_{\beta\beta,\beta\beta}]_{ab,cd} \equiv A_{a_\beta b_\beta, c_\beta d_\beta} = \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b - 2\nu) + \langle ab||cd \rangle,$$

$$[B_{\alpha\alpha,\alpha\alpha}]_{ab,ij} \equiv B_{a_\alpha b_\alpha, i_\alpha j_\alpha} = \langle ab||ij \rangle$$

$$[B_{\alpha\beta,\alpha\beta}]_{ab,ij} \equiv B_{a_\alpha b_\beta, i_\alpha j_\beta} = \langle ab|ij \rangle$$

$$[B_{\beta\beta,\beta\beta}]_{ab,ij} \equiv B_{a_\beta b_\beta, i_\beta j_\beta} = \langle ab||ij \rangle,$$

7

$$[C_{\alpha\alpha,\alpha\alpha}]_{ij,kl} \equiv C_{i_\alpha j_\alpha, k_\alpha l_\alpha} = -\delta_{ik}\delta_{jl}(\epsilon_i + \epsilon_j - 2\nu) + \langle ij||kl\rangle$$

$$[C_{\alpha\beta,\alpha\beta}]_{ij,kl} \equiv C_{i_\alpha j_\beta, k_\alpha l_\beta} = -\delta_{ik}\delta_{jl}(\epsilon_i + \epsilon_j - 2\nu) + \langle ij|kl\rangle$$

$$[C_{\beta\beta,\beta\beta}]_{ij,kl} \equiv C_{i_\beta j_\beta, k_\beta l_\beta} = -\delta_{ik}\delta_{jl}(\epsilon_i + \epsilon_j - 2\nu) + \langle ij||kl\rangle,$$

The final eigenvalue set is the union of the three matrices' eigenvalue sets

$$\omega \equiv \omega_{\alpha\alpha,\alpha\alpha} \cup \omega_{\alpha\beta,\alpha\beta} \cup \omega_{\beta\beta,\beta\beta}, \tag{18}$$

and the trace of the original $\mathbf{A}$ (or $\mathbf{C}$) matrix is the sum of the traces of the three smaller $\mathbf{A}_{spin}$ (or $\mathbf{C}_{spin}$) matrices

$$\mathrm{Tr}\mathbf{A} = \sum_{spin} \mathrm{Tr}\mathbf{A}_{spin}, \tag{19a}$$

$$\mathrm{Tr}\mathbf{C} = \sum_{spin} \mathrm{Tr}\mathbf{C}_{spin}. \tag{19b}$$

Therefore, the correlation energy in Eq.(13) can be written as

$$E_c = \sum_{spin} \left[\sum_{n_{spin}} \omega^+_{n_{spin}} - \mathrm{Tr}\mathbf{A}_{spin} \right] = \sum_{spin} \left[-\sum_{n_{spin}} \omega^-_{n_{spin}} - \mathrm{Tr}\mathbf{C}_{spin} \right], \tag{20}$$

The spin separation can be implemented in both restricted and unrestricted cases. For restricted closed-shell singlet cases, the eigenvalue problem can be further simplified by using spin-adapted particle-particle and hole-hole pairs.

The $\alpha\alpha$ pairs and $\beta\beta$ pairs occur only as triplet pairs, while $\alpha\beta$ pairs can combine into either singlet or triplet pairs,

$$\text{Singlet}: \quad \frac{1}{\sqrt{2}} \frac{1}{\sqrt{1+\delta_{pq}}} (p^\dagger_\alpha q^\dagger_\beta + q^\dagger_\alpha p^\dagger_\beta) \tag{21a}$$

$$\text{Triplet}: \quad \frac{1}{\sqrt{2}} (p^\dagger_\alpha q^\dagger_\beta - q^\dagger_\alpha p^\dagger_\beta), \tag{21b}$$

where p and q are generic orbitals that can be both occupied or both unoccupied. Singlet and triplet pair annihilations can be constructed in a similar manner. The spin adapted pp-RPA $\alpha\beta$ matrix for the restricted case is obtained by dividing the $\alpha\beta$ pairs into blocks according to their spin multiplicity:

$$\begin{bmatrix} \mathbf{A}_s & 0 & \mathbf{B}_s & 0 \\ 0 & \mathbf{A}_t & 0 & \mathbf{B}_t \\ \mathbf{B}^\dagger_s & 0 & \mathbf{C}_s & 0 \\ 0 & \mathbf{B}^\dagger_t & 0 & \mathbf{C}_t \end{bmatrix}$$

and the decoupled eigenvalue problem is

$$
\begin{bmatrix} \mathbf{A}_{mult} & \mathbf{B}_{mult} \\ \mathbf{B}^{\dagger}_{mult} & \mathbf{C}_{mult} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{mult} \\ \mathbf{Y}_{mult} \end{bmatrix} = \omega_{mult} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{mult} \\ \mathbf{Y}_{mult} \end{bmatrix}, \tag{22}
$$

where the multiplicity $mult$ is either singlet (s) or triplet (t). The elements in the triplet matrix are

$$
\begin{aligned}
[A_t]_{ab,cd} &= \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b - 2\nu) + \langle ab||cd \rangle \\
[B_t]_{ab,ij} &= \langle ab||ij \rangle \\
[C_t]_{ij,kl} &= -\delta_{ik}\delta_{jl}(\epsilon_i + \epsilon_j - 2\nu) + \langle ij||kl \rangle,
\end{aligned}
$$

with the restriction that $a > b$, $c > d$, $i > j$ and $k > l$. This triplet eigenvalue problem is the same as the $\alpha\alpha$ and $\beta\beta$ cases and therefore gives the same eigenvalue set

$$
\omega_t = \omega_{\alpha\alpha,\alpha\alpha} = \omega_{\beta\beta,\beta\beta} \tag{24a}
$$

$$
\mathrm{Tr}\mathbf{A}_t = \mathrm{Tr}\mathbf{A}_{\alpha\alpha,\alpha\alpha} = \mathrm{Tr}\mathbf{A}_{\beta\beta,\beta\beta} \tag{24b}
$$

The elements in the singlet matrix are

$$
[A_s]_{ab,cd} = \delta_{ac}\delta_{bd}(\epsilon_a + \epsilon_b - 2\nu) + \frac{1}{\sqrt{(1+\delta_{ab})(1+\delta_{cd})}}(\langle ab|cd \rangle + \langle ab|dc \rangle) \tag{25a}
$$

$$
[B_s]_{ab,ij} = \frac{1}{\sqrt{(1+\delta_{ab})(1+\delta_{ij})}}(\langle ab|ij \rangle + \langle ab|ji \rangle) \tag{25b}
$$

$$
[C_s]_{ij,kl} = -\delta_{ik}\delta_{jl}(\epsilon_i + \epsilon_j - 2\nu) + \frac{1}{\sqrt{(1+\delta_{ij})(1+\delta_{kl})}}(\langle ij|kl \rangle + \langle ij|lk \rangle), \tag{25c}
$$

with the restriction that $a \geq b$, $c \geq d$, $i \geq j$ and $k \geq l$.

The linear combination to generate spin-adapted pairs is a unitary transformation. Consequently, the eigenvalues and the traces do not change — the $\alpha\beta$ eigenvalue set is the union of the singlet and triplet eigenvalue sets

$$
\omega_{\alpha\beta,\alpha\beta} \equiv \omega_s \cup \omega_t, \tag{26}
$$

and the trace of $\mathbf{A}_{\alpha\beta,\alpha\beta}$ (or $\mathbf{C}_{\alpha\beta,\alpha\beta}$) matrix equals the sum of the traces of the singlet $\mathbf{A}_s$ (or $\mathbf{C}_s$) and triplet $\mathbf{A}_t$ (or $\mathbf{C}_t$)

$$
\mathrm{Tr}\mathbf{A}_{\alpha\beta,\alpha\beta} = \mathrm{Tr}\mathbf{A}_s + \mathrm{Tr}\mathbf{A}_t \tag{27a}
$$

$$
\mathrm{Tr}\mathbf{C}_{\alpha\beta,\alpha\beta} = \mathrm{Tr}\mathbf{C}_s + \mathrm{Tr}\mathbf{C}_t. \tag{27b}
$$

Therefore the correlation energy in Eqs. (13) and (20) can further be expressed as

$$E_c = 3(\sum_n \omega_{t,n}^+ - \mathrm{Tr}\mathbf{A}_t) + \sum_n \omega_{s,n}^+ - \mathrm{Tr}\mathbf{A}_s = 3(-\sum_n \omega_{t,n}^- - \mathrm{Tr}\mathbf{C}_t) - \sum_n \omega_{s,n}^- - \mathrm{Tr}\mathbf{C}_s.$$
(28)

In the following sections, Eq. (20) and Eq. (28) are used to compute correlation energies for unrestricted and restricted cases, respectively. With the spin separation, the largest matrices are about half the dimension of the original matrices, thus the memory usage and the computational time are roughly reduced to 1/4 and 1/8 compared to the unoptimized algorithm, respectively. For closed-shell singlet systems, with the spin adaptation, the largest matrices are only about a quarter the dimension of the original matrices, therefore, the computational cost are further reduced to 1/16 in memory and 1/64 in time as of the unoptimized algorithm. However, even with the lowered factor, the computational efficiency through direct diagonalization is still $O(N^6)$, with N either number of virtual orbitals or number of occupied orbitals, whichever dominates.

III. COMPUTATIONAL DETAILS

We have implemented the pp-RPA with spin separation and spin adaptation in QM4D [32]. For pp-RPA calculations, HF or Kohn-Sham DFT reference states were computed using the QM4D package. The PBE reference [33] was used for the G2/97 enthalpies of formation database and the DBH24 reaction barrier database. Both the PBE and HF references were tested and compared for the nonbonded interaction databases. Basis set convergence was tested along the cc-pVXZ series, X=D, T, Q, 5, for selected systems in G2/97, and with the aug-cc-pVXZ series, X=D, T, Q, for selected reaction barriers. The basis set convergence of pp-RPA for non-bonded interactions was assessed with aug-cc-pVXZ series, X=D, T, Q (spherical harmonic atomic orbitals) in a locally modified [?] version of CFOUR [34]. After balancing the accuracy and the computational cost, calculations for G2/97, DBH24 and nonbonded interactions adopted the cc-pVTZ, aug-cc-pVTZ and aug-cc-pVDZ basis sets, respectively. The QM4D program uses Cartesian atomic orbitals and removes basis functions with angular momentum higher than "f". For enthalpies of formation and reaction barriers tests, we used cc-pVXZ-RI auxiliary basis sets (basis functions with angular momentum higher than "f" also truncated) in the post-KS pp-RPA to facilitate the atomic orbital to molecular orbital two-electron integral transformation. The ph-RPA calculations
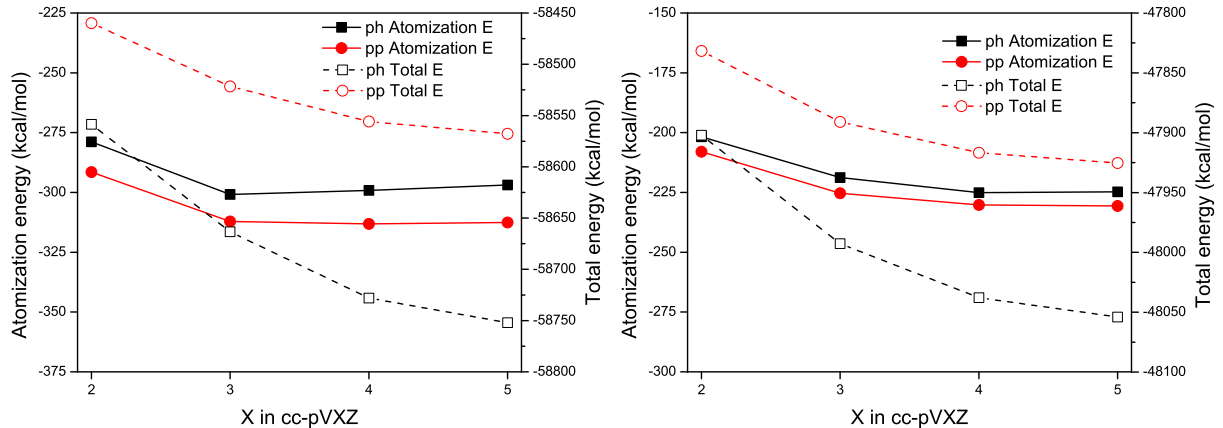
FIG. 2: Basis set convergence for the total energy and atomization energy of HCN (left) and H_2O(right). The total energies for HCN and H_2O converge very slowly with respect to the basis set. The atomization energies, which are plotted with equal intervals between tick marks on the left axis, converge much faster and can be considered converged with the cc-pVTZ basis set.

were carried out in the same way, except that no RI auxiliary basis sets were used. Geometries for the G2/97 benchmark set are taken from Ref. [35], which were optimized using the MP2(full)/6-31G* method. Geometries for the DBH24 set are taken from Ref. [27]. Geometries for the HB6/04, CT7/04, DI6/04 and WI9/04 nonbonded interaction sets are taken from Ref. [29].

IV. RESULTS AND DISCUSSIONS

A. G2/97 enthalpies of formation

1. Basis set convergence test

We choose HCN and H_2O to assess the basis set convergence of total and atomization energies (Figure 2). For both molecules and both types of RPA, the total energy converges slowly and shows as large as ≈ 10 kcal/mol and 20 kcal/mol differences between the cc-pVQZ and the cc-pV5Z basis sets for the pp-RPA and ph-RPA, respectively. By contrast, due to systematic error cancellation, the atomization energy shows about or less than 5 kcal/mol differences between the cc-pVTZ and the cc-pVQZ basis sets. Considering the balance of accuracy and large computational cost, we adopted the cc-pVTZ basis set for benchmarking

11

enthalpies of formation.

2. *Results*

Enthalpies of formation allow for a direct comparison with experimental results and are therefore often used to benchmark electronic structure methods. The closely related zero-point-energy-free atomization energies are somewhat more straightforward to compare to high-level computations and it is therefore customary to report both. In the present work, we investigate the performance of the pp- and ph-RPA for the atomization energies and enthalpies of formation for the G2/97 database [25, 26] (see Table VI in Supplementary Material [36] for the detailed data). The smallest molecule of G2/97 is H_2 and the largest in terms of atoms and number of electrons, are C_4H_{10} and $SiCl_4$, respectively.

Among the whole set, the maximum error (MaxE) for the pp-RPA is -31.1 kcal/mol (C_2F_4), which is half of that for the ph-RPA (63.2 kcal/mol, SiF_4). The mean signed error (MSE) for the pp-RPA is -1.9 kcal/mol which is much smaller than that for the ph-RPA (21.7 kcal/mol). The mean unsigned error (MUE) is 8.3 kcal/mol and 21.7 kcal/mol for the pp-RPA and ph-RPA. These indicate that the error for the pp-RPA is fluctuating around the reference values while the ph-RPA systematically overestimates enthalpies of formation. This behavior also emerges when we plot the signed error for both RPAs with respect to the number of atoms in a molecule (Fig. 3). In contrast to the ph-RPA, the pp-RPA enthalpies of formation show no systematic drift with respect to the number of atoms. The systematic underbinding of the ph-RPA is well known, although previous conclusions were based on data limited to small molecules [4, 37–39]. This error has been ascribed to the ph-RPA's insufficiency to describe the short-range correlation, which may be important when the number of electron pairs changes [3]. Since the number of atoms in a molecule is roughly correlated to the number of electron pairs formed, the increasing error in Fig 3 is in agreement with this argument.

The G2/97 database is often divided into the G2-1 (small molecules) and G2-2 (large molecules) subsets. The G2-2 subset can be further divided into 5 subsets, namely non-hydrogen systems, hydrocarbons, substituted hydrocarbons, inorganic hydrides and radicals. Results are shown in Table I. For the small molecules of the G2-1 subset, the ph-RPA has a relatively low error (MUE = 10.9 kcal/mol) compared to the remaining subsets. The pp-
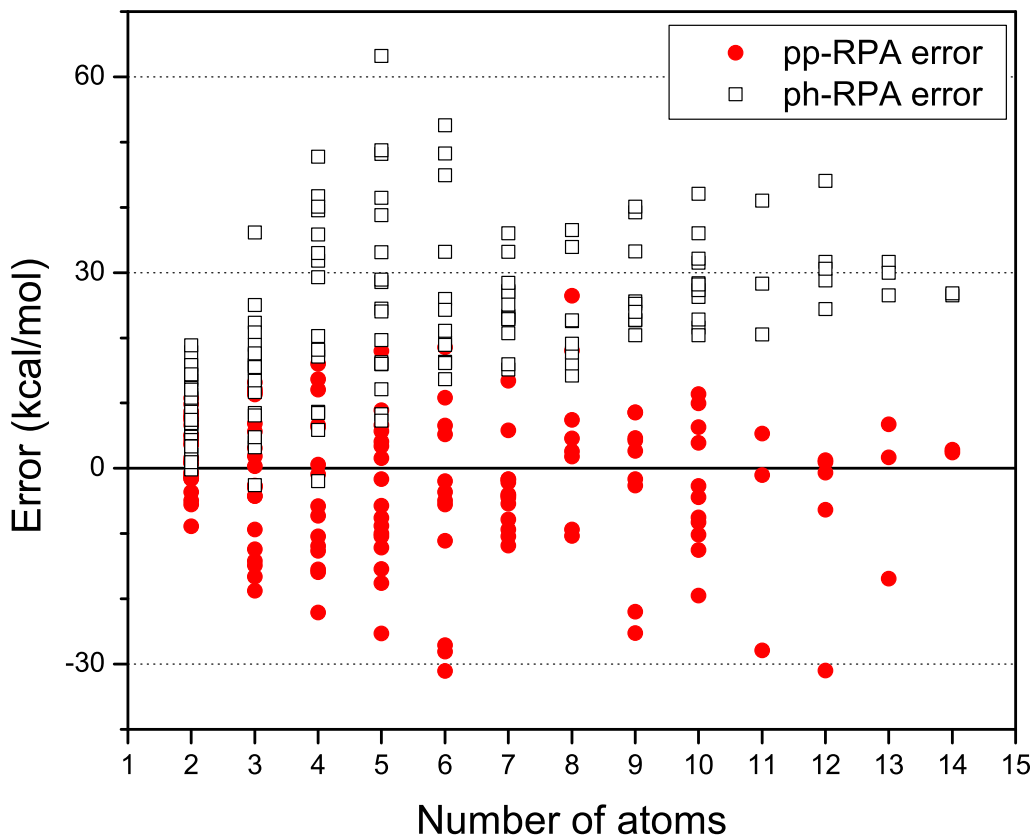
FIG. 3: Signed error with respect to number of atoms for each species in G2/97 enthalpies of formation. Red dots stand for the pp-RPA and black-white squares stand for the ph-RPA. The cc-pVTZ basis set was adopted. The pp-RPA shows smaller errors within about ±30 kcal/mol and the error has a nearly constant trend with respect to number of atoms, while the ph-RPA has errors within about 0-60 kcal/mol and growing with respect to number of atoms.

RPA is already much better than the ph-RPA for this set (MUE = 6.8 kcal/mol). Neither RPA performs well for non-hydrogen systems: the pp-RPA underestimates the enthalpies of formation by 14.6 kcal/mol while the ph-RPA overestimates them by 35.9 kcal/mol. For the two organic subsets (hydrocarbons and substituted hydrocarbons), the pp-RPA performs significantly better than the ph-RPA (MUE = 8 and 27 kcal/mol, respectively). Detailed inspections show that the pp-RPA is good at describing non-cyclic organic compounds. For cyclic compounds such as methylene cyclopropane, bicyclobutane and spiropentane, the error reaches over 10 kcal/mol. For aromatic cyclic systems such as benzene, furan and pyridine, the error even exceeds 20 kcal/mol. When dividing the G2/97 database into three

TABLE I: Mean Signed Errors (MSE) and Mean Unsigned Errors (MUE) (in kcal/mol) of pp-RPA and ph-RPA for subsets in G2/97

subset	MSE		MUE	
	pp-RPA	ph-RPA	pp-RPA	ph-RPA
G2-1	4.7	10.8	6.8	10.9
Non-hydrogen systems	-14.6	35.9	14.6	35.9
Hydrocarbons	-5.6	26.9	7.2	26.9
Substituted hydrocarbons	-3.8	28.2	8.5	28.2
Inorganic hydrids	7.7	2.1	7.7	2.1
Radicals	0.9	19.4	5.1	19.4
Total	-1.9	22.7	8.3	22.7

exclusive subsets for molecules containing only single bonds, molecules with double bonds and molecules featuring a triple bond, the pp-RPA is found to have a low MSE for single bonds (2.2 kcal/mol) but still a substantial MUE (7.9 kcal/mol), while for the ph-RPA both figures of merit are about 19.6 kcal/mol. This indicates that for single bonded systems the pp-RPA over- and underestimates heats of formations to a similar extent. This is no longer true in molecules with multiple bonds: with mostly negative errors, the MSE of the pp-RPA amounts to -8.6 and -4.4 kcal/mol for double and triple bonds, respectively. This indicates that the pp-RPA may be slightly problematic for multiple-bond systems. However, with a mean signed and unsigned error of 0.9 kcal/mol and 5.1 kcal/mol, respectively, the pp-RPA describes radicals very well. Only NO_2, a non-hydrogen system with a double bond, is problematic for the pp-RPA. Just like for closed-shell systems, the ph-RPA overestimates the enthalpies of formation for radicals.

3. Reaction energies from G2/97

Chemically relevant transformations conserving the number of electron pairs might provide a view complementary to the enthalpies of formation. Therefore, we examine 19 reactions involving organic compounds to investigate the performance of the pp- and ph-RPA.

The reactions are divided into four groups, namely, hydrocarbon reactions, substituted hydrocarbon isomerization reactions, substitution reactions and addition reactions (Table II and III). Although the ph-RPA does not predict accurate enthalpies of formation, it describes the enthalpy changes in chemical reactions rather well (MUE = 2.3 kcal/mol), on par with the pp-RPA (MUE = 2.4 kcal/mol). For the four addition reactions, where a double (or triple) bond is converted to two single bonds, the pp-RPA yields significantly larger errors. Already in the enthalpies of formation we have observed the qualitatively different behavior for single and double bonds for the pp-RPA. Our speculation is that the approximate pairing interactions fail to describe intra-electron pair correlation on an equal footing with inter-electron pair correlation. The reasonably "constant" performance for the ph-RPA on the other hand, can be understood considering that the number of electron pairs does not change during these reactions and therefore the major source of error for enthalpies of formation does not play any role.

In the hydrocarbon reactions, the enthalpy difference between allene and propyne, has been used to assess the reliability of density functionals for determining the poly-yne vs. cumulene stability, a very tricky energy difference in general[40]. Even though the sign is correct for both RPAs (in contrast to typical density functional approximations such as B3LYP), the pp-RPA has a large error (2.8 kcal/mol, >200%), overly stabilizing the triple bond of propyne compared to the double bonds in allene. However, the ph-RPA is not affected by such a problem, in agreement with the analysis of the G2/97 set with respect to the bond types. A similar preference for the electron localized geometry is the isomerization of 2-butyne to the more stable, conjugated butadiene. In contrast to these reactions involving a changing degree of electron delocalization, both RPAs perform excellently for the isomerization energy of butane and the isodesmic reaction energies for n-alkanes. For both types of reactions typical density functionals fail dramatically, most likely because of an inaccurate treatment of weak interactions.[41–44]

TABLE II: Reaction energies (in kcal/mol) between molecules taken from the G2/97 database. Basis set: cc-pVTZ. Numbers in the parenthesis indicate the error.

Type	Reaction	Benchmark	pp-RPA	ph-RPA
Hydrocarbon	$CH_3CCH \rightarrow CH_2CCH_2$	1.3	4.1(2.8)	1.8(0.5)
Reactions	$CH_3CCCH_3 \rightarrow CH_2CHCHCH_2$	-8.5	-3.7(4.8)	-5.4(3.1)
	$CH_3CH_2CH_2CH_3 \rightarrow CH_3CH(CH_3)CH_3$	-2.1	-2.5(-0.4)	-1.7(0.4)
	$C_3H_8+CH_4 \rightarrow 2C_2H_6$	2.7	3.3(0.6)	2.3(-0.4)
	$C_4H_{10}+2CH_4 \rightarrow 3C_2H_6$	5.5	7.1(1.6)	5.0(-0.5)
Substituted	$C_2H_4O(oxirane) \rightarrow CH_3CHO$	-27.1	-24.5(2.6)	-24.9(2.2)
Hydrocarbon	$C_2H_5OH \rightarrow CH_3OCH_3$	12.2	12.6(0.4)	11.8(-0.4)
Isomerization	$C_2H_5SH \rightarrow CH_3SCH_3$	2.2	2.2(0.0)	3.4(1.2)
	$(CH_3)_2CHOH \rightarrow C_2H_5OCH_3$	13.5	13.8(0.3)	12.5(-1.0)
	$(CH_3)_2NH \rightarrow CH_3CH_2NH_2$	-6.9	-8.3(-1.4)	-8.0(-1.1)
Substitution	$NCCN+C_2H_6 \rightarrow 2CH_3CN$	-17.2	-15.9(1.3)	-13.8(3.4)
Reactions	$H_2NNH_2+C_2H_6 \rightarrow 2CH_3NH_2$	-13.7	-12.8(0.9)	-11.2(2.5)
	$Cl_2+C_2H_6 \rightarrow 2CH_3Cl$	-19.1	-19.6(-0.5)	-15.2(3.9)
	$Si_2H_6+C_2H_6 \rightarrow 2CH_3SiH_3$	-13	-10.7(2.3)	-7.8(5.2)
	$HOOH+C_2H_6 \rightarrow 2CH_3OH$	-43.4	-44.3(-0.9)	-37.4(6.0)
Addition	$HCl+C_2H_4 \rightarrow C_2H_5Cl$	-17.2	-25.0(-7.8)	-19.1(-1.9)
Reactions	$HCN+C_2H_2 \rightarrow CH_2CHCN$	-42.5	-47.4(-4.9)	-43.8(-1.3)
	$HF+C_2H_2 \rightarrow CH_2CHF$	-22.3	-26.7(-4.4)	-26.6(-4.3)
	$HCl+C_2H_2 \rightarrow CH_2CHCl$	-23.2	-31.6(-8.4)	-28.2(-5.0)

B. DBH24 reaction barriers

1. Basis set convergence test

Figure 4 shows the basis set dependence of the barrier heights of $H + OH \rightarrow O + H_2$ and $HCN \rightarrow HNC$ along the aug-cc-pVXZ basis sets series, with X = D, T, Q. For $H + OH \rightarrow O + H_2$, both RPAs converge well with nearly flat behaviors except for the back-

16

TABLE III: Mean Signed Errors (MSEs) and Mean Unsigned Errors (MUEs) (in kcal/mol) of of reaction energies calculated by pp-RPA and ph-RPA using the cc-pVTZ basis set.

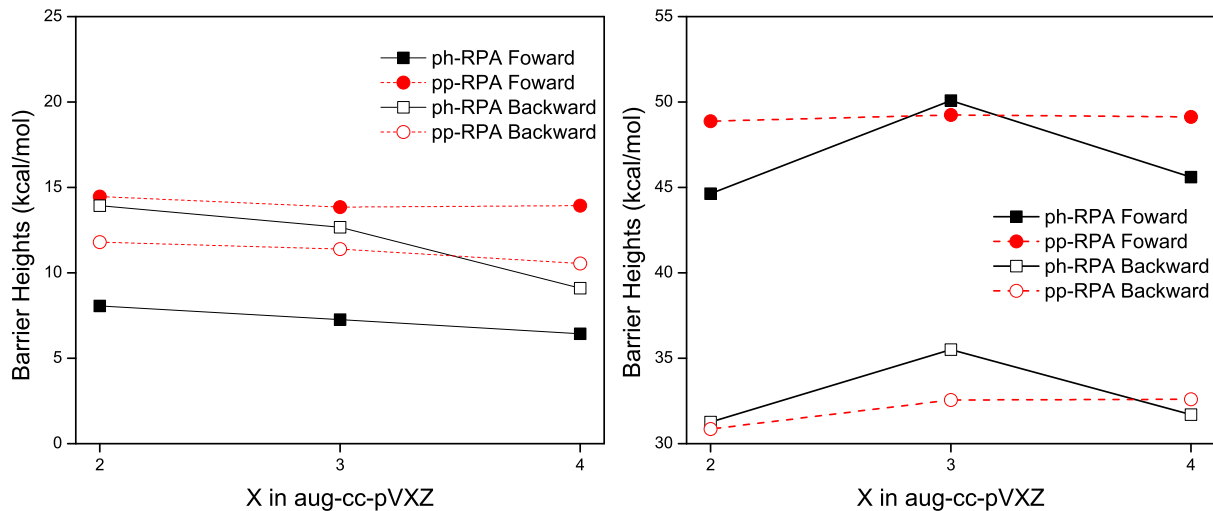| Type	MSE		MUE	
	pp-RPA	ph-RPA	pp-RPA	ph-RPA
Hydrocarbon Reactions	1.9	0.6	2.0	1.0
Substitued Hydrocarbon Isomerizations	0.4	0.2	0.9	1.2
Substitution Reactions	0.6	4.2	1.2	4.2
Addition Reactions	-6.4	-3.1	6.4	3.1
All	-0.6	0.7	2.4	2.3

FIG. 4: Basis set convergence for forward and backward reactions on $H + OH \rightarrow O + H_2$ and $HCN \rightarrow HNC$. The pp-RPA converges well for both cases, while the ph-RPA still has rather large differences between aug-cc-pVTZ and aug-cc-pVQZ.

ward reaction calculated by ph-RPA, which yields a 3 kcal/mol difference between aug-cc-pVTZ and aug-cc-pVQZ. While the pp-RPA has a similar behavior for $HCN \rightarrow HNC$, the ph-RPA has a "bump" at the aug-cc-pVTZ basis of about 4 kcal/mol. These results emphasize that reaction barriers can be very sensitive to basis sets. Nevertheless, considering the computational cost, we choose the aug-cc-pVTZ basis set for the following reaction barrier calculations and expect the results to reflect the correct relative performance.

2. Results

Benchmark values in the DBH24 reaction barrier test set are best estimates either from experiments or highly accurate theoretical methods [28]. The overall performance for the DBH24 database is very similar for the pp-RPA and the ph-RPA (see Table IV and V): the pp-RPA has a slightly smaller mean signed error (-1.11 kcal/mol) than ph-RPA (-1.65 kcal/mol), while the ph-RPA has a slightly smaller mean unsigned error (2.48 kcal/mol vs. 3.19 kcal/mol). Among the four subsets in DBH24, the pp-RPA has the largest mean unsigned error (5.56 kcal/mol) for the HATBH6 subset, which includes three heavy-atom transfer reactions. This is mainly due to the $H + N_2O \rightarrow OH + N_2$ reaction, in which the pp-RPA overestimates the forward barrier by 5.21 kcal/mol and underestimates the backward barrier by 16.82 kcal/mol. The reason is two-fold: firstly, as is shown before, the pp-RPA has difficulties predicting enthalpies of formation for some compounds with double bonds and triple bonds, i.e. the reactants and products are not well described with a 22 kcal/mol error for the reaction energy. Secondly, the pp-RPA does not describe the spin-unpolarized bond-stretching of double and triple bonds well [20], leading to a large error for the transition state. The NSBH6 subset includes three nucleophilic substitution reactions, and both RPAs perform well except for the $OH^- + CH_3F \rightarrow HOCH_3 + F^-$ reaction, which might suffer from delocalization errors in the PBE reference determinant. For UABH6, which includes three unimolecular and association reactions, both RPAs perform well. For HTBH6, which consists of three hydrogen transfer reactions, both methods give accurate reaction barriers, except for $OH + CH_4 \rightarrow CH_3 + H_2O$, where they both underestimate the energy of the transition state. In conclusion, despite some tricky cases, both the pp-RPA and ph-RPA generally provide reliable reaction barriers.

C. HB6/04, CT7/04, DI6/04 and WI9/04 nonbonded interaction

1. Basis set convergence test

To assess the basis set convergence of the pp-RPA, we chose a strong hydrogen bonded system ($HF - HF$), a charge-transfer complex ($H_2O - ClF$), a dipole-dipole and a "pure" van der Waals dimer ($HCl - HCl$ and $CH_4 - Ne$, respectively). Overall, the interaction energies are similar with different basis (aug-cc-pVXZ with X= D, T, Q), although in some cases the

TABLE IV: DBH24 reaction barriers (in kcal/mol) calculated by pp-RPA and ph-RPA using the aug-cc-pVTZ basis set. Numbers in the parenthesis indicate the error.

Database	Reaction	Benchmark	pp-RPA	ph-RPA
HATBH6	$H + N_2O \rightarrow OH + N_2$	17.13	22.34(5.21)	17.50(0.37)
		82.47	65.65(-16.82)	75.49(-6.98)
	$H + ClH \rightarrow HCl + H$	18.00	21.29(3.29)	18.78(0.78)
		18.00	21.29(3.29)	18.78(0.78)
	$CH_3 + FCl \rightarrow CH_3F + Cl$	6.75	3.04(-3.71)	0.16(-6.59)
		60.00	58.93(-1.07)	57.80(-2.2)
NSBH6	$Cl^- \cdots CH_3Cl \rightarrow ClCH_3 \cdots Cl^-$	13.41	12.00(-1.41)	10.87(-2.54)
		13.41	12.00(-1.41)	10.87(-2.54)
	$F^- \cdots CH_3Cl \rightarrow FCH_3 \cdots Cl^-$	3.44	2.11(-1.33)	1.38(-2.06)
		29.42	26.17(-3.25)	26.46(-2.96)
	$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	-2.44	-6.76(-4.32)	-5.89(-3.45)
		17.66	13.79(-3.87)	13.20(-4.46)
UABH6	$H + N_2 \rightarrow HN_2$	14.36	17.33(2.97)	14.23(-0.13)
		10.61	10.34(-0.27)	10.54(-0.07)
	$H + C_2H_4 \rightarrow CH_3CH_2$	1.72	3.95(2.23)	2.31(0.59)
		41.75	43.49(1.74)	43.20(1.45)
	$HCN \rightarrow HNC$	48.07	49.24(1.17)	50.07(2.00)
		32.82	32.55(-0.27)	35.50(2.68)
HTBH6	$OH + CH_4 \rightarrow CH_3 + H_2O$	6.7	0.4(-6.3)	4.2(-2.5)
		19.6	16.6(-3.0)	11.3(-8.3)
	$H + OH \rightarrow O + H_2$	10.7	13.9(3.2)	7.3(-3.4)
		13.1	11.4(-1.7)	12.7(-0.4)
	$H + H_2S \rightarrow H_2 + HS$	3.6	5.5(1.9)	2.8(-0.8)
		17.3	14.4(-2.9)	18.6(1.3)

TABLE V: Mean Signed Errors (MSEs) and Mean Unsigned Errors (MUEs) (in kcal/mol) of DBH24 reaction barriers and its four subsets calculated by pp-RPA and ph-RPA

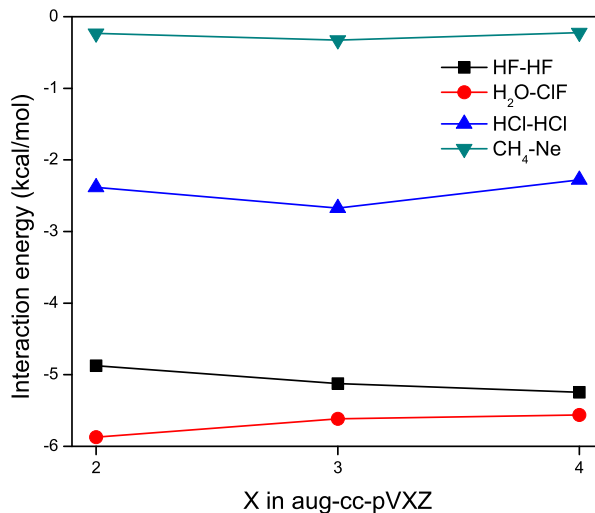| Database	MSE		MUE	
	pp-RPA	ph-RPA	pp-RPA	ph-RPA
HATBH6	-1.64	-2.30	5.56	2.95
NSBH6	-2.60	-3.00	2.60	3.00
UABH6	1.26	1.09	1.44	1.15
HTBH6	-1.47	-2.37	3.14	2.79
DBH24	-1.11	-1.65	3.19	2.48

FIG. 5: Basis set convergence for nonbonded interactions calculated by pp-RPA. Different basis sets give similar interaction energies.

interaction energy does not change monotonically with respect to the basis (see Figure 5). Considering the high computational cost, we adopt the aug-cc-pVDZ basis set and correct for the basis set superposition error (BSSE) according to the Boys-Bernardi counterpoise correction [45] (see Supplementary Material [36] for results without BSSE correction). Table VI confirms that counterpoise corrected aug-cc-pVDZ results overall agree well with the aug-cc-pVQZ results and therefore we expect them to reflect the correct relative performance of the ph- and pp-RPA.

TABLE VI: Mean Signed Errors (MSEs) and Mean Unsigned Errors (MUEs) (in kcal/mol) of HB6/04, CT7/04, DI6/04 and WI9/04 nonbonded interaction by pp-RPA and ph-RPA. If not stated otherwise, the aug-cc-pVDZ basis set has been applied and except for aug-cc-pVQZ the basis set superposition error is corrected for.

| Database | ph-RPA | | | | pp-RPA | | | | | |
| | HF | | PBE | | HF | | PBE | | HF/aug-cc-pVQZ [a] | |
	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
HB8/04	1.14	1.14	1.82	1.82	-0.03	0.48	-0.63	0.76	-0.38	0.42
CT7/04	1.70	1.70	4.15	4.15	0.26	0.26	-0.24	1.22	-0.10	0.37
DI6/04	1.29	1.29	1.57	1.57	0.38	0.40	-0.34	0.37	-0.25	0.25
WI9/04	0.32	0.32	0.29	0.29	0.17	0.17	-0.04	0.15	-0.01	0.04
Total	1.05	1.05	1.86	1.86	0.20	0.31	-0.28	0.60	-0.15	0.25

[a]Calculations were performed with spherical harmonic basis functions. BSSEs were not corrected. Results reported are based on selected systems that are within computational capability. See Table VII-X in Supplementary Material [36] for detailed data.

2. Results

Results for nonbonded interactions are shown in Table VI. Both HF reference and PBE reference are investigated for the two RPAs. Overall, with the aug-cc-pVDZ basis, the HF reference gives slightly better results than the PBE reference for both RPAs. This is expected as the DFT reference has the wrong asymptotic behavior, leading to an overestimation of correlation effects [46]. After BSSE correction, the pp-RPA gives a slightly smaller deviation than the ph-RPA. The two RPAs perform similarly for all four types of interactions, with the pp-RPA using the Hartree-Fock determinant gives the best agreement with the benchmark data. Therefore, it can be concluded that the pp-RPA describes weak interactions equally well as the more popular ph-RPA. This may be due to the accurate second-order energy expansion of the ladder diagram [20], which is commonly believed to dominate the van der Waals interaction. The relationship between the two types of RPA in describing the asymptotic van der Waals awaits further investigation.

V. CONCLUSIONS

We have implemented the spin separation and spin adaptation for the pp-RPA, which factors the pp-RPA eigenvalue problem into several smaller ones. This simplification reduces the computational cost, allowing thorough benchmark tests on the G2/97 enthalpies of formation, DBH24 reaction barriers, and four nonbonded interaction databases. Our results demonstrate that the pp-RPA performs significantly better than the ph-RPA for enthalpies of formation: in contrast to the increasing error of the ph-RPA with the number of atoms in a molecule, the pp-RPA has a nearly constant error. For reaction enthalpies, barriers heights and non bonded interactions, the pp-RPA and ph-RPA perform essentially equally well. These benchmark tests indicate that the pp-RPA is a promising method even for larger systems, although systems with multiple bonds tend to be relatively problematic. This shortcoming might be overcome in further development: the general success of the pp-RPA suggests that the pairing interaction in conjunction with the adiabatic connection formalism forms a promising framework for developing new density functionals.

---

[1] D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952).

[2] A. Hesselmann and A. Gorling, Molecular Physics **109**, 2473 (2011).

[3] H. Eshuis, J. Bates, and F. Furche, Theoretical Chemistry Accounts **131**, 1084 (2012).

[4] X. Ren, P. Rinke, C. Joas, and M. Scheffler, Journal of Materials Science **47**, 7447 (2012).

[5] J. M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960).

[6] J.-P. Blaizot, *Quantum theory of finite systems*, MIT Press, Cambridge, Mass., 1986.

[7] J. Čížek, The Journal of Chemical Physics **45**, 4256 (1966).

[8] G. E. Scuseria, T. M. Henderson, and D. C. Sorensen, The Journal of Chemical Physics **129**, 231101 (2008).

[9] D. Langreth and J. Perdew, Solid State Communications **17**, 1425 (1975).

[10] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[11] F. Furche and T. V. Voorhis, The Journal of Chemical Physics **122**, 164106 (2005).

[12] H. Eshuis, J. Yarkony, and F. Furche, The Journal of Chemical Physics **132**, 234114 (2010).

[13] X. Ren et al., New Journal of Physics **14**, 053020 (2012).

[14] A. Szabo and N. S. Ostlund, The Journal of Chemical Physics **67**, 4351 (1977).

[15] H. Eshuis and F. Furche, The Journal of Physical Chemistry Letters **2**, 983 (2011).

[16] T. M. Henderson and G. E. Scuseria, Molecular Physics **108**, 2511 (2010).

[17] P. Mori-Sánchez, A. J. Cohen, and W. Yang, Phys. Rev. A **85**, 042507 (2012).

[18] F. Furche, The Journal of Chemical Physics **129**, 114105 (2008).

[19] H. Eshuis and F. Furche, The Journal of Chemical Physics **136**, 084105 (2012).

[20] H. van Aggelen, Y. Yang, and W. Yang, Phys. Rev. A **88**, 030501 (2013).

[21] D. Peng, S. N. Steinmann, H. van Aggelen, and W. Yang, The Journal of Chemical Physics **139**, 104112 (2013).

[22] G. E. Scuseria, T. M. Henderson, and I. W. Bulik, The Journal of Chemical Physics **139**, 104113 (2013).

[23] A. J. Cohen, P. Mori-Snchez, and W. Yang, Science **321**, 792 (2008).

[24] A. Hesselmann, The Journal of Chemical Physics **134**, 204107 (2011).

[25] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, The Journal of Chemical Physics **94**, 7221 (1991).

[26] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, The Journal of Chemical Physics **106**, 1063 (1997).

[27] J. Zheng, Y. Zhao, and D. G. Truhlar, Journal of Chemical Theory and Computation **3**, 569 (2007).

[28] J. Zheng, Y. Zhao, and D. G. Truhlar, Journal of Chemical Theory and Computation **5**, 808 (2009).

[29] Y. Zhao and D. G. Truhlar, Journal of Chemical Theory and Computation **1**, 415 (2005).

[30] D. J. Rowe, Rev. Mod. Phys. **40**, 153 (1968).

[31] P. Ring and P. Schuck, *The Nuclear Many-Body Problem*, Physics and astronomy online

library, Springer, 2004.

[32] An in-house program for qm/mm simulations (http://www.qm4d.info).

[33] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[34] Cfour, coupled-cluster techniques for computational chemistry, a quantum-chemical program package by j.f. stanton, j. gauss, m.e. harding, p.g. szalay with contributions from a.a. auer, r.j. bartlett, u. benedikt, c. berger, d.e. bernholdt, y.j. bomble, l. cheng, o. christiansen, m. heckert, o. heun, c. huber, t.-c. jagau, d. jonsson, j. juselius, k. klein, w.j. lauderdale, d.a. matthews, t. metzroth, l.a. muck, d.p. o'neill, d.r. price, e. prochnow, c. puzzarini, k. ruud, f. schiffmann, w. schwalbach, s. stopkowicz, a. tajti, j. vazquez, f. wang, j.d. watts and the integral packages molecule (j. almlof and p.r. taylor), props (p.r. taylor), abacus (t. helgaker, h.j. aa. jensen, p. jorgensen, and j. olsen), and ecp routines by a. v. mitin and c. van wullen. for the current version, see http://www.cfour.de.

[35] G2/97, http://www.cse.anl.gov/OldCHMwebsiteContent/compmat/g2-97.htm.

[36] Supplementary Material, available at...

[37] F. Furche, Phys. Rev. B **64**, 195120 (2001).

[38] A. Ruzsinszky, J. P. Perdew, and G. I. Csonka, Journal of Chemical Theory and Computation **6**, 127 (2010).

[39] J. Paier et al., New Journal of Physics **14**, 043002 (2012).

[40] H. L. Woodcock, H. F. Schaefer, and P. R. Schreiner, J. Phys. Chem. A **106**, 11923 (2002).

[41] S. Grimme, Angew. Chem., Int. Ed. **45**, 4460 (2006).

[42] M. D. Wodrich, C. Corminboeuf, and P. v. R. Schleyer, Org. Lett. **8**, 3631 (2006).

[43] S. N. Steinmann, M. Wodrich, and C. Corminboeuf, Theor. Chem. Acc. **127**, 429 (2010).

[44] J.-W. Song, T. Tsuneda, T. Sato, and K. Hirao, Org. Lett. **12**, 1440 (2010).

[45] S. Boys and F. Bernardi, Molecular Physics **19**, 553 (1970).

[46] W. Cencek and K. Szalewicz, The Journal of Chemical Physics **139**, 024104 (2013).