



Molecular interactions in electron-groups embedding generalized valence bond picture

Ewa Pastorczak¹ · Katarzyna Pernal¹

Received: 9 July 2018 / Accepted: 25 September 2018 / Published online: 11 November 2018
© The Author(s) 2018

Abstract

Theoretical description of molecular interactions remains a challenge for computational chemistry. In particular, systems dominated by static correlation, in which bonds are stretched or twisted, are often beyond capabilities of methods based on a single-electron approximation, being usually a method of choice. For interacting multireference systems, it is necessary to apply either high-level coupled-cluster methods (which however include multireference effects only partially) or base the theoretical description on a multireference wavefunction. Popular multireference methods like CASPT2 (complete active space perturbation theory) do not provide satisfactory results since they may suffer from problems with size consistency and poor accuracy. Recently we have shown that combining a simple multireference wavefunction, perfect-pairing generalized valence bond (GVB) with extended random phase approximation in embedding framework (EERPA) leads to a method EERPA-GVB providing accurate results for challenging multireference systems. In this paper, good performance of EERPA-GVB is confirmed by its application to van der Waals and hydrogen-bonded complexes. In addition, we show that the decomposition of the EERPA-GVB correlation energy into contributions from pairs of geminals can provide useful insight into the investigated interactions.

Keywords Molecular interactions · Random phase approximation · Multireference systems

1 Introduction

The ubiquity of non-covalent interactions and their increasingly appreciated role in such fields as material design [1], catalysis [2], medicine [3] and even photochemistry [4–6] necessitates the development of computational tools able to describe them. Historically, it has been a challenge, in particular due to the long-range nature and the subtlety of the London dispersion, but recently sophisticated coupled-cluster approaches are becoming more computationally affordable [7, 8] and efficient approaches such as the density functional theory (DFT) [9] have developed ways to treat

van der Waals interactions [10]. Molecular interactions can also be computed and analyzed using Symmetry-Adapted Perturbation Theory (SAPT) [11, 12], and other energy decomposition schemes [13, 14] and some progress toward providing the same description for intramolecular interactions has been made [15, 16]. The success of computational methods has not however been extended to systems where the static correlation plays a role. This group includes, apart from somewhat exotic systems like chromium or beryllium dimers, all systems containing significantly stretched or compressed bonds. This is a severe limitation as computational and experimental studies, albeit scarce [17–21], have shown how essential the van der Waals interactions can be in the reaction process.

Capturing molecular interactions in systems where bonds are being twisted or broken is so challenging because one needs to ensure that both the static and the dynamic correlation are accurately described. This usually means one needs to employ a multireference (MR) wavefunction and a good-quality dynamic-correlation correction. What is more, the method should be size-extensive and be able to produce smooth interaction energy surfaces. Favorable scaling with the basis set

Published as part of the special collection of articles In Memoriam of János Ángyán.

✉ Ewa Pastorczak
ewa.pastorczak@p.lodz.pl

✉ Katarzyna Pernal
pernal@gmail.com

¹ Institute of Physics, Lodz University of Technology, ul. Wolczanska 219, 90-924 Lodz, Poland

size is also of value since the description of non-covalent interactions demands high-quality basis sets containing diffuse functions. Popular coupled-cluster methods are of single reference kind and they are prone to fail if applied to interacting strongly correlated systems, unless a high-level CC, with full triples or higher, is employed [22]. The CASPT2 method—often a method of choice for multireference systems—may suffer from lack of size consistency, difficulty with obtaining smooth potential energy curves, and poor accuracy when applied to the description of molecular interactions.

We have recently introduced an Embedding Extended Random Phase Approximation (EERPA) [22] correlation correction and paired it with a simple multireference wavefunction—a strongly orthogonal perfect-pairing generalized valence bond (GVB). GVB wavefunction accounts for electron pair correlation providing correct description of bond breaking process but it lacks long-range correlation. Consequently it is not able to account for weak interactions. Adding dynamic-correlation correction to GVB energy via perturbation theory or linearized multireference coupled-cluster theory allows one to include dispersion interaction component but the accuracy is poor (it must be admitted that the applications are scarce) [23]. EERPA, on the other hand, based on introducing extended random phase approximation correlation correction in an embedding fashion, was shown to be a tailor-made approach for describing intermolecular interactions of multireference systems [22]. It is accurate and numerically stable. What is more, EERPA-GVB is not as computationally demanding as even CCSD(T) method, let alone the higher-level approaches like CCSDT. The cost of computing the EERPA correction is similar to that of the familiar RPA correlation, while the reference GVB wavefunction scales as $N_g M^4$, where M is the basis set size and N_g —the number of geminals [24],

The goal of this paper is to explore further capabilities of EERPA and show its usefulness in predicting not only values of interacting energy but also in getting insight into interaction between monomers. We firstly recap the theoretical framework of EERPA-GVB, then we show the method's robustness using examples of dispersion-dominated and hydrogen-bonded dimers, and finally we analyze the behavior of two van der Waals complexes with twisted or broken bonds.

The perfect-pairing GVB ansatz, being of interest in this paper conforms to the generalized product function form proposed by McWeeny [25] and it reads [26–28]

$$\Psi^{\text{GVB}} = \hat{A} \prod_{I=1}^{N/2} \Psi^I, \quad (1)$$

where N is a number of electrons assumed to be even. Each Ψ^I is an antisymmetric two-electron wavefunction, called a geminal. The antisymmetry of the total wavefunction is assured by the antisymmetrizing operator \hat{A} , which includes

a proper normalization factor. Geminals are constrained to be strongly orthogonal [25, 29]

$$\forall_{I \neq J} \int \Psi^I(\mathbf{x}, \mathbf{x}') \Psi^J(\mathbf{y}, \mathbf{x}') d\mathbf{x}' = 0 \quad (2)$$

(here and elsewhere in the paper \mathbf{x}, \mathbf{y} combines spatial and spin coordinates of a single electron). Additional two constraints imposed on geminals are: singlet symmetry and the constraint that each geminal is expanded in a subspace of only two orbitals. Consequently, each GVB geminal can be written in the following form

$$\forall_I \Psi^I = \Psi^I(\mathbf{x}_1, \mathbf{x}_2) = 2^{-1/2} (c_{1I} \varphi_{1I}(\mathbf{r}_1) \varphi_{1I}(\mathbf{r}_2) + c_{2I} \varphi_{2I}(\mathbf{r}_1) \varphi_{2I}(\mathbf{r}_2)) (\alpha\beta - \beta\alpha). \quad (3)$$

The squares of expansion coefficients c_p present in Eq. (3) are the natural occupation numbers $\{n_p\}$

$$\forall_p c_p^2 = n_p \in [0, 1], \quad (4)$$

$\{\varphi_p\}$ being a set of the natural orbitals pertaining to the ansatz Ψ^{GVB} .

Strong orthogonality of geminals is a restriction imposed on the wavefunction but in return one is rewarded with a simple expression for the electronic energy reading [28, 30]

$$E^{\text{GVB}} = 2 \sum_p^N n_p h_{pp} + \sum_{pq}^N \delta_{I_p I_q} c_p c_q \langle pp|qq \rangle + \sum_{pq}^N (1 - \delta_{I_p I_q}) n_p n_q (2 \langle pq|pq \rangle - \langle pq|qp \rangle), \quad (5)$$

where $\{h_{pp} = \langle \varphi_p | \hat{t} + \hat{v}_{\text{ext}} | \varphi_p \rangle\}$ are one-electron integrals in the representation of the natural orbitals and two-electron integrals are defined using a standard $r_1 r_2 r_1 r_2$ convention. Symbols I_p in Eq. (5) indicate a geminal to which a spinorbital φ_p belongs. Thus, a symbol $\delta_{I_p I_q}$ equals 1 only if both orbitals p and q are used in expansion of the same geminal (cf. Eq. (3)), whereas $(1 - \delta_{I_p I_q})$ is different from zero for two orbitals p and q belonging to two different geminals. A quick look at the GVB energy expression reveals immediately its main appealing features: electron pair correlation is accounted for due to presence of the middle term (GVB goes beyond a single particle picture), the number of CI coefficients undergoing optimization is only equal to N . In addition, active orbitals in GVB are unique and the number of them is N . Taking into account that in widely used multireference methods like CAS-SCF or MCSCF the number of CI coefficients is exponentially growing and the choice of active orbitals is often problematic and arbitrary, it becomes clear that the GVB ansatz is more computationally attractive than the latter methods.

Deficiencies of GVB have been known, see, e.g., Refs. [31–33], and they include a deteriorated performance for molecules described by more than one Lewis structure, failure in dissociating multiple bonds, practical limitation to closed-shell systems and the lack of dispersion energy in weakly interacting systems. Recently we have shown that GVB, when corrected with properly designed correlation energy corrections, yields excellent results for molecules undergoing conformational changes [30, 34], and for molecular interactions [22]. For the convenience of the reader, we will present two correlation energy corrections: ERPA and EERPA. The former is generally applicable and can be seen as extension of the random phase approximation (RPA) [35] correlation energy for multireference wavefunction. EERPA (Embedding ERPA), on the other hand, in its current formulation is applicable to weakly interacting dimers.

Let us begin with the Extended Random Approximation (ERPA). The ERPA equation reads [36–39]

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B} & \mathcal{A} \end{pmatrix} \begin{pmatrix} \mathbf{X}^\nu \\ \mathbf{Y}^\nu \end{pmatrix} = \omega_\nu \begin{pmatrix} -\mathcal{N} & \mathbf{0} \\ \mathbf{0} & \mathcal{N} \end{pmatrix} \begin{pmatrix} \mathbf{X}^\nu \\ \mathbf{Y}^\nu \end{pmatrix}, \quad (6)$$

For a system S the \mathcal{A} , \mathcal{B} , and \mathcal{N} matrices (assumed to be real-valued) are determined from the one- and two-electron reduced density matrices, γ^S and Γ^S , obtained from an assumed reference wavefunction (GVB in this case)

$$\forall_{p>q,r>s} \mathcal{A}_{pqrs}^S = \mathcal{A}_{pqrs}(\gamma^S, \Gamma^S), \quad (7)$$

$$\forall_{p>q,r>s} \mathcal{B}_{pqrs}^S = \mathcal{A}_{pqsr}(\gamma^S, \Gamma^S), \quad (8)$$

$$\forall_{p>q,r>s} \mathcal{N}_{pqrs}^S = (n_p - n_q) \delta_{pr} \delta_{qs} \quad (9)$$

(see Appendix in Ref. [30] for the explicit expressions for the matrix elements of \mathcal{A} for the GVB wavefunction). The eigenvectors $[\mathbf{X}^\nu, \mathbf{Y}^\nu]$ provide approximation to reduced transition density matrices that has led to obtaining the following spin-summed ERPA correlation energy expression [30, 34, 40]

$$E_{\text{corr}}^{\text{ERPA}}(S) = \sum_{p>r,q>s} (1 - \delta_{p,q} \delta_{r,s} \delta_{q,r}) W_{pqrs}^S, \quad (10)$$

$$W_{pqrs}^S = \left\{ \sum_{\nu} (n_r - n_p)(n_s - n_q) (X_{pr}^{\nu} + Y_{pr}^{\nu}) (X_{qs}^{\nu} + Y_{qs}^{\nu}) - \frac{1}{2} [(1 - n_r)n_p + (1 - n_p)n_r] \delta_{rs} \delta_{pq} \right\} \langle pq|rs \rangle. \quad (11)$$

Notice that the expression $(1 - \delta_{p,q} \delta_{r,s} \delta_{q,r})$ in Eq. (10) vanishes if all orbitals p, q, r, s belong to the same geminal. In

this way, intra-geminal correlation energy already included in the GVB ansatz is excluded from $E_{\text{corr}}^{\text{ERPA}}$ and double counting of correlation is avoided. The ERPA-GVB interaction energy for a dimer consisting of monomers A and B reads

$$E_{\text{ERPA-GVB}}^{\text{Int}} = E_{AB}^{\text{GVB}} + E_{\text{corr}}^{\text{ERPA}}(AB) - (E_A^{\text{GVB}} + E_{\text{corr}}^{\text{ERPA}}(A) + E_B^{\text{GVB}} + E_{\text{corr}}^{\text{ERPA}}(B)). \quad (12)$$

Thus, the ERPA equations are solved three times: for a dimer, $S = AB$ and for two monomers, $S = A$ and $S = B$.

In the EERPA-GVB approach, inspired by one-electron reduced density matrix embedding theory [41], energies of monomers are obtained in the same fashion as in ERPA-GVB but the correlation energy description of a dimer is modified in order to account separately for a correlation of electrons in a monomer embedded in the field of another monomer, and for inter-monomer correlation effects. The EERPA correlation has been defined as follows:

$$E_{\text{corr}}^{\text{EERPA}}(AB) = E_{\text{corr}}^A + E_{\text{corr}}^B + E_{\text{corr}}^{AB}. \quad (13)$$

The correlation energy of the monomer A embedded in B , E_{corr}^A , results from solving truncated ERPA equations with the following matrices

$$\mathcal{A}^A = [\mathcal{A}_{pqrs}(\gamma^{AB}, \Gamma^{AB})]_{p>q, r>s, pqrs \in \Omega_A} \quad (14)$$

$$\mathcal{B}^A = [\mathcal{B}_{pqrs}(\gamma^{AB}, \Gamma^{AB})]_{p>q, r>s, pqrs \in \Omega_A} \quad (15)$$

and

$$\mathcal{N}^A = [(n_p - n_q) \delta_{pr} \delta_{qs}]_{p>q, r>s, pqrs \in \Omega_A} \quad (16)$$

where the set Ω_A defines excitations originating from an orbital localized on a monomer A (a set G_A includes all occupied orbitals assigned to A) and ending in either another orbital from G_A or an unoccupied orbital (V is a set of all unoccupied orbitals) or one of the weakly occupied orbital localized on a monomer B , i.e.,

$$\Omega_A = \{p > r, q > s : r, s \in G_A \wedge p, q \in G_A \cup G_B^{\text{weak}} \cup V\}, \quad (17)$$

where

$$G_B^{\text{weak}} = \left\{ \varphi_p : p \in G_B \wedge n_p < \frac{1}{2} \right\}. \quad (18)$$

The correlation energy E_{corr}^A entering Eq. (13) results from

$$E_{\text{corr}}^A = \sum_{\substack{p > r, q > s \\ pqrs \in \Omega_A}} (1 - \delta_{I_p I_q} \delta_{I_r I_s} \delta_{I_q I_r}) W_{pqrs}^A \quad (19)$$

where W_{pqrs}^A is defined in Eq. (11). The energy for a monomer B embedded in A , E_{corr}^B , is obtained analogously. E_{corr}^{AB} , the remaining term in the EERPA correlation energy expression for a dimer (13), accounts for inter-monomer correlation effects and it is obtained by solving ERPA equations for a dimer with the matrices

$$\mathcal{A}^{AB} = [\mathcal{A}_{pqrs}(\gamma^{AB}, \Gamma^{AB})]_{p>q, r>s}, \quad (20)$$

$$\mathcal{B}^{AB} = [\mathcal{B}_{pqrs}(\gamma^{AB}, \Gamma^{AB})]_{p>q, r>s}, \quad (21)$$

and

$$\mathcal{N}^{AB} = [(n_p - n_q) \delta_{pr} \delta_{qs}]_{p>q, r>s}, \quad (22)$$

and computing the correlation energy from the expression

$$E_{\text{corr}}^{AB} = \sum_{\substack{p > r, q > s \\ pqrs \notin \bar{\Omega}_A \\ pqrs \notin \bar{\Omega}_B}} (1 - \delta_{I_p I_q} \delta_{I_r I_s} \delta_{I_q I_r}) W_{pqrs}^{AB}, \quad (23)$$

where the set $\bar{\Omega}_A$ defines intra-monomer contributions

$$\bar{\Omega}_A = \{p > r, q > s : r, s \in G_A \wedge p, q \in G_A \cup V\}, \quad (24)$$

and analogously for $\bar{\Omega}_B$.

The EERPA-GVB interaction energy results from the supermolecular calculation and it is computed as (cf. Eq. (12))

$$E_{\text{EERPA-GVB}}^{\text{Int}} = E_{AB}^{\text{GVB}} + E_{\text{corr}}^{\text{EERPA}}(AB) - (E_A^{\text{GVB}} + E_{\text{corr}}^{\text{ERPA}}(A) + E_B^{\text{GVB}} + E_{\text{corr}}^{\text{ERPA}}(B)), \quad (25)$$

where the correlation energy expression for monomers is given by Eqs. (6)–(11). Thus, the difference between EERPA-GVB and the ERPA-GVB interaction energy given in Eq. (12) amounts to a different approach to the correlation energy for a dimer, which for EERPA is given by Eqs. (13)–(24). It should be stressed that in the limit when a distance between monomers tends to infinity the EERPA and the ERPA correlation energy for a dimer become equivalent and the interaction energy tends to inter-monomer correlation energy E_{corr}^{AB} , which in turn reduces to the second-order dispersion energy [39], i.e.,

$$E_{\text{EERPA-GVB}}^{\text{Int}}(R_{AB} \rightarrow \infty) = E_{\text{ERPA-GVB}}^{\text{Int}}(R_{AB} \rightarrow \infty) = E_{\text{corr}}^{AB} = E_{\text{disp}}^{(2)}(AB), \quad (26)$$

where E_{corr}^{AB} is the inter-monomer correlation term as defined in the EERPA approach, cf. Eq. (13).

2 Weakly bounded complexes in- and out-of-equilibrium geometry in the EERPA-GVB picture

2.1 Computational details

To highlight the properties of the ERPA-GVB and EERPA-GVB approaches, we carried out calculations of interaction energies for a number of weakly interacting dimers bound by hydrogen bonds, such as $\text{NH}_3 \cdots \text{H}_2\text{O}$, hydrogen sulfide and water dimers, and for van der Waals (vdW) complexes where dispersion energy is the driving force, i.e., in $\text{He} \cdots \text{Ne}$, acetylene and ethene dimers. All systems are described by the aug-cc-pVDZ basis set [42]. To judge the accuracy of those results, we computed CCSD(T) energies as implemented in the DALTON software package [43] in the same basis set. We also performed SAPT2+3(CCD) (hereafter referred to as SAPT) computations [44] using Psi4 software [45].

In addition, we performed an analysis of the basis set dependence (in basis sets aug-cc-pV ζ Z, where $\zeta = 2, \dots, 6$) of ERPA-GVB and EERPA-GVB methods using the example of helium dimer.

We have also focused on dimers involving molecules in out-of-equilibrium geometries, i.e., ethene dimer with one of the monomers twisted, ethene–fluorine complex with F–F bond stretched and compressed, and the same complex with ethene molecule twisted. Those computations were also performed in aug-cc-pVDZ basis set.

Interaction energies computed with supermolecular methods were corrected for the basis set superposition error (BSSE) using the Boys' counterpoise correction [46]. GVB computations were performed in a developer version of DALTON software package [43]. ERPA and EERPA corrections were computed in our in-house code interfaced with DALTON. Core orbitals were correlated. The only orbitals included in the “active” set as described in Ref. [30] were those involved in twisting or stretching of the bonds. The equilibrium geometries of studied complexes were taken from NIB database developed by Truhlar et al. [47].

2.2 Results and discussion

The intuitive understanding of what embedding in the EERPA method does is that it counteracts the counterpoise correction, since the correlation energy for a dimer, Eq. (13), includes embedded-monomer correlation terms,

cf. Eq. (19), which are obtained by allowing excitations from one monomer to weakly occupied orbitals on another monomer. One could then wonder whether—if the basis set used is sufficiently large—the counterpoised corrected ERPA and EERPA methods would produce the same results. This is, however, not the case, since the correlation effects included in EERPA are missed in ERPA and this is not related to the basis set size. This conclusion can be illustrated by the example of helium dimer (see Fig. 1) where ERPA-GVB and EERPA-GVB interaction energies nearly parallel each other up to the aug-cc-pV6Z basis set. Notice also that EERPA-GVB curve remains close but above the benchmark MC-ACPF one for all basis sets. Although it is not guaranteed to be the case for all systems, such behavior is a sign of dependable performance of the method.

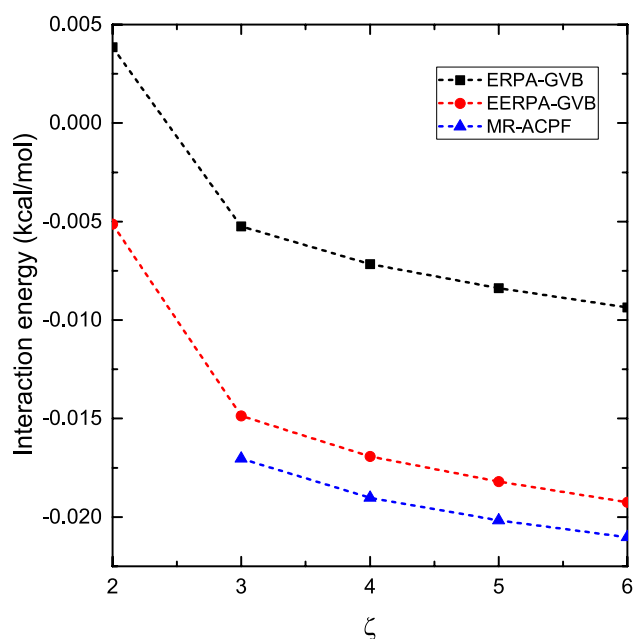


Fig. 1 He₂ interaction energy at $R_{\text{He-He}} = 5.6$ a.u. computed in aug-cc-pV ζ Z basis set. The reference MR-ACPF energies were taken from [48]

Already in [22], we have shown that EERPA-GVB describes accurately the dispersion-dominated systems. Here, we reaffirm this statement, adding that for this type of systems EERPA-GVB tends to be on par or even more accurate than SAPT [11, 44] computations (see Table 1). Interestingly, in terms of SAPT energy decomposition scheme, GVB interaction energies are almost exactly the sum of what SAPT identifies as electrostatic, exchange and induction components. While this is a rather intuitive result, it has not been previously demonstrated, since the perfect-pairing strongly orthogonal GVB method is not, as a rule, used for describing non-covalent interactions. The task of the EERPA correction should be therefore to add the dispersion component as well as mixed terms such as exchange–dispersion, exchange–induction and the exchange–induction–dispersion. The latter three play no significant role in the interaction of systems presented in Table 1, but as we shall see later, this is not always the case.

It is impressive that even in the challenging case of helium–neon interaction, which is purely dispersion-driven, the EERPA-GVB curve stays on top of the CCSD(T) one (see Fig. 2) when even the SAPT energies follows the MP2 curve. The ERPA-GVB minimum is even more shallow than the MP2 one and the GVB method, as expected, does not produce a minimum at all.

All three systems presented in Table 1 are single reference and weakly interacting so this is not surprising that both MP2 and SAPT describe them reasonably well. However, EERPA-GVB is clearly the best performer here, not just in the vdW minimum, but also along the entire curves (see Figs. 2 and 3). The ERPA-GVB method underbinds the complexes, while MP2 and SAPT slightly overestimate the interaction energies.

For hydrogen-bonded systems, EERPA-GVB is less accurate than for van der Waals complexes although it stays superior to the ERPA-GVB, and (obviously) GVB methods. Below we present a set of three hydrogen-bonded systems: water and hydrogen sulfide dimer and water–ammonia complex (see Table 2). For each of them, the electrostatic component is a significant part of the interaction, so already

Table 1 Interaction energies of dispersion-dominated systems in kcal/mol computed in aug-cc-pVDZ basis set

	CCSD(T)	MP2	GVB	ERPA-GVB	EERPA-GVB	SAPT	Elst+Exch+Ind
He...Ne	-0.029	-0.020 (0.009)	0.024 (0.053)	-0.014 (0.015)	-0.029 (0.000)	-0.017 (0.012)	0.020 (-0.049)
(C ₂ H ₂) ₂	-1.12	-1.23 (-0.11)	-0.33 (0.78)	-0.96 (0.16)	-1.16 (-0.04)	-1.21 (-0.10)	-0.33 (0.79)
(C ₂ H ₄) ₂	-1.14	-1.21 (-0.07)	0.54 (1.68)	-0.78 (0.36)	-1.12 (0.02)	-1.23 (-0.09)	0.61 (1.75)

Errors w.r.t. CCSD(T) are given in parentheses. In the last column, a sum of electrostatic, exchange, and induction components (SAPT calculation) is shown. Geometries for complexes are taken from Ref. [47] and the assumed distances between centers of masses of monomers read: 6.02 a.u., 8.41 a.u. and 7.32 a.u. for He...Ne, (C₂H₂)₂ and (C₂H₄)₂, respectively

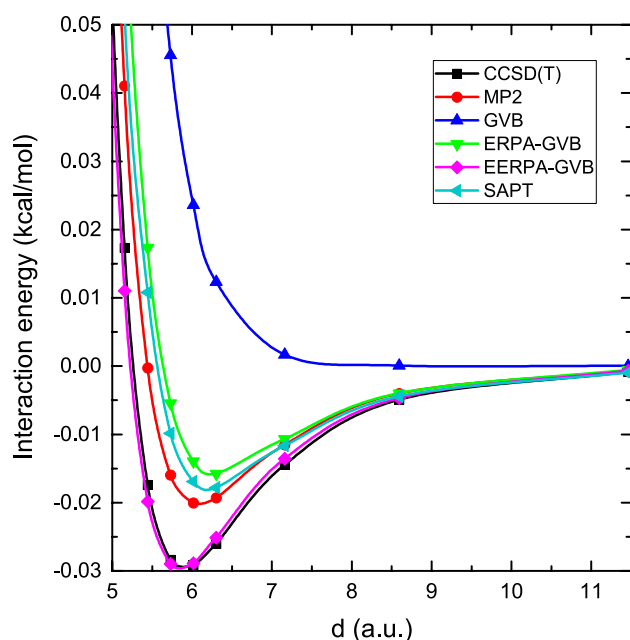


Fig. 2 He...Ne interaction energy curves computed in aug-cc-pVDZ basis set

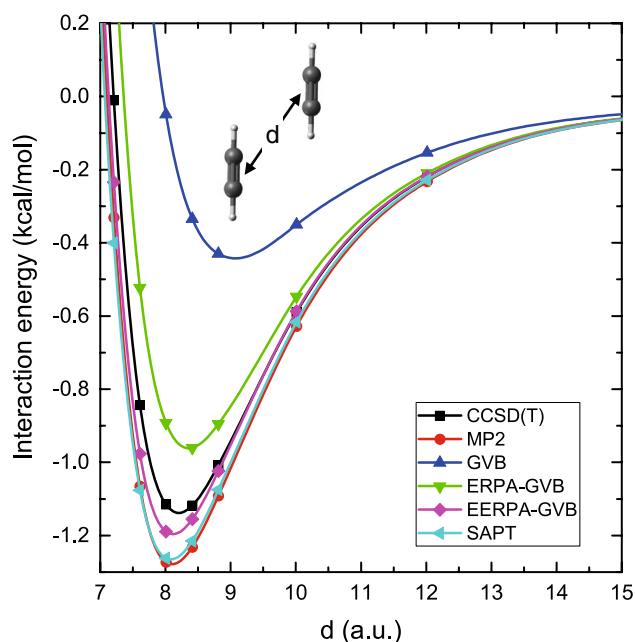


Fig. 3 C₂H₂ dimer interaction energy curves computed in aug-cc-pVDZ basis set

the GVB method reproduces part of the binding. As it has been observed for dispersion-bound complexes, GVB interaction energies stay close to the sums of the electrostatic, exchange and induction components in a SAPT computation. The absolute error of the interaction energies produced by

EERPA-GVB is at most 0.6 kcal/mol, but this may amount to as much as 15% of the binding. What is the reason for this (relative with respect to dispersion-bound systems) inaccuracy? In the first three examples, the exchange–dispersion component of the energy was very small (less than 8%) compared to the dispersion component. For the considered hydrogen-bonded systems, the exchange–dispersion is between 10 and 20% of the dispersion energy. The exchange–dispersion component is always positive and can be interpreted as the change in the exchange interaction introduced by the monomer correlation [49]. While the GVB wavefunction itself is antisymmetric, there is no additional antisymmetry-related constraint put on the correlation energy expression. The two-body ERPA correlation component E_{corr}^{AB} converges to dispersion energy for well-separated monomers, cf. Eq. (26). From those observations, one can conjecture that the inaccuracies are related to exchange–dispersion component, which is not properly accounted for by EERPA. Indeed, for the NH₃...H₂O complex (see inset in Fig. 5), one can see that in the region of no density overlap the E_{AB}^{corr} correlation curve follows the sum of SAPT pure dispersion terms rather than the sum of all terms containing dispersion (dispersion, exchange–dispersion, induction–dispersion and exchange–induction–dispersion).

Regardless of this observation, as evident from Figs. 4 and 5, EERPA-GVB is a reliable method for description of hydrogen-bonded systems. In particular in case of H₂S dimer, it yields a very similar curve as SAPT (see Fig. 4). The shapes of all the curves are correct, and the minima are only slightly too deep.

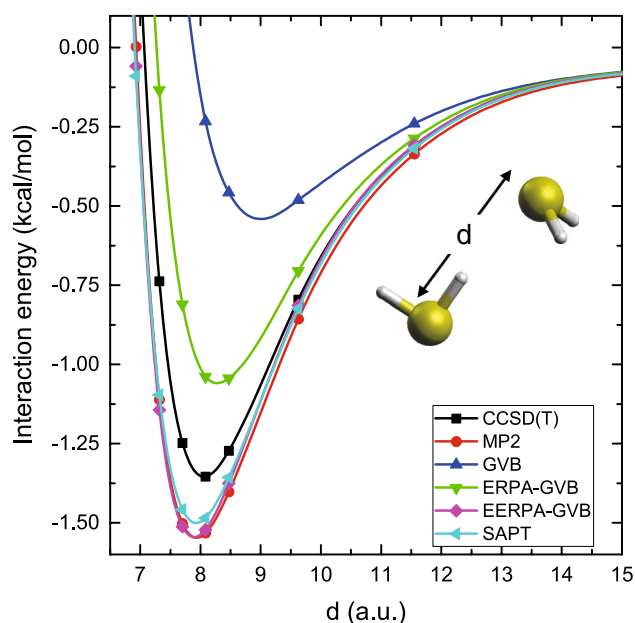
The true advantage of EERPA-GVB lies however elsewhere, i.e., in its ability to accurately describe the interactions of systems out-of-equilibrium geometry when bonds are stretched or broken and one (or both) of the monomers requires multireference description. This ability gives one a nearly unique opportunity to elucidate the effects of the non-covalent interactions on systems attempting chemical reactions.

Let us look again at the ethylene dimer in D_{2d} symmetry. We have established that EERPA-GVB describes this system accurately and that the GVB method is responsible for the description of the electrostatic, exchange and induction component of the interaction. Let us now twist the C–C bond in one of the monomers to $\theta = 90^\circ$ (see Fig. 6). Clearly, in the attraction region the GVB method produces essentially the same results for both geometries—no binding, which means that the total attraction here is related to the dispersion interaction (including the mixed terms). The binding energy of the twisted complex is more than 40% smaller than in the flat dimer. While one often intuitively understands dispersion as a “bulk” interaction between electron clouds, not discriminating between different types of bonds, this is not a full picture: according to a frozen-orbital SAPT analysis

Table 2 Interaction energies of hydrogen-bonded systems in kcal/mol computed in aug-cc-pVDZ basis set

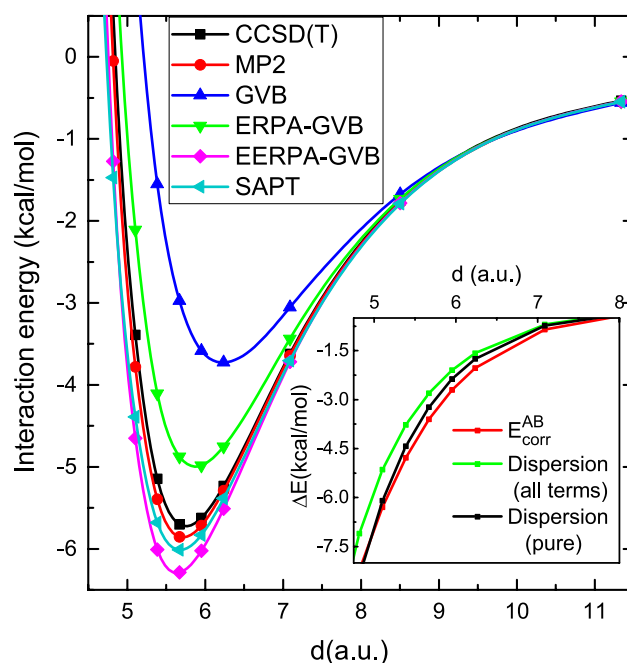
	CCSD(T)	MP2	GVB	ERPA-GVB	EERPA-GVB	SAPT	Elst+Exch+Ind
(H ₂ O) ₂	-4.40	-4.44 (-0.04)	-2.48 (1.92)	-3.76 (0.64)	-4.86 (-0.46)	-4.53 (-0.13)	-2.59 (1.81)
NH ₃ ⋯H ₂ O	-5.70	-5.85 (-0.15)	-2.97 (2.73)	-4.88 (0.83)	-6.29 (-0.59)	-6.01 (-0.31)	-3.21 (2.49)
(H ₂ S) ₂	-1.35	-1.53 (-0.18)	-0.23 (1.12)	-1.04 (0.31)	-1.52 (-0.17)	-1.49 (-0.13)	-0.19 (1.16)

Errors w.r.t. CCSD(T) are given in parentheses. Geometries for complexes are taken from Ref. [47] and the assumed distances between centers of masses of monomers read: 5.56 a.u., 5.67 a.u. and 8.09 a.u. for (H₂O)₂, NH₃⋯H₂O and (H₂S)₂, respectively

**Fig. 4** H₂S dimer interaction energy curves computed in aug-cc-pVDZ basis set

by Cao and Wong [50], the main contributors to dispersion in the D_{2d} ethylene dimer are σ - π bond pairs and dihydrogen contacts (where contacts are understood as significant interactions between electron pairs). The reduced attraction can be partly attributed to the geometry change: one of the σ - σ “contacts” disappears, but most of the reduction has to be attributed to the destruction of C=C π bond. This observation is in agreement with the view of π electrons having larger polarizabilities than σ ones and therefore contributing more to the dispersion interaction [51].

A particularly useful feature of EERPA-GVB is its ability to decompose the correlation energy into contributions from pairs of interacting geminals, which are localized on different monomers. This is possible since the inter-monomer correlation energy given in Eq. (23) and used in EERPA is a sum of inter-geminal terms involving orbitals p, q, r, s of which at least one is assigned to one of a geminal I_A

**Fig. 5** NH₃⋯H₂O complex interaction energy curve computed in aug-cc-pVDZ basis set. On the inset, the EERPA correlation is plotted against the sum of SAPT pure dispersion terms and the sum of SAPT pure dispersion and mixed (dispersion-exchange, exchange-dispersion and exchange-dispersion-induction)

localized on the monomer A and at least one is assigned to a geminal localized on B , I_B

$$pqrs \in I_A \cup I_B \cup V \quad (27)$$

where V indicates a set of unoccupied orbitals. Therefore, contribution to inter-monomer correlation energy from a pair of geminals, one localized on A and the other on B can be extracted from E_{corr}^{AB} by selecting only terms with indices pertaining to Eq. (27). By excluding terms corresponding to a π -bond-geminal interacting with lone-pair geminals localized on the fluorine atom F_1 , closer to the ethylene molecule, one can check how much the π -bond-lone pairs (LP) interaction contributes to the total binding energy of the charge-transfer

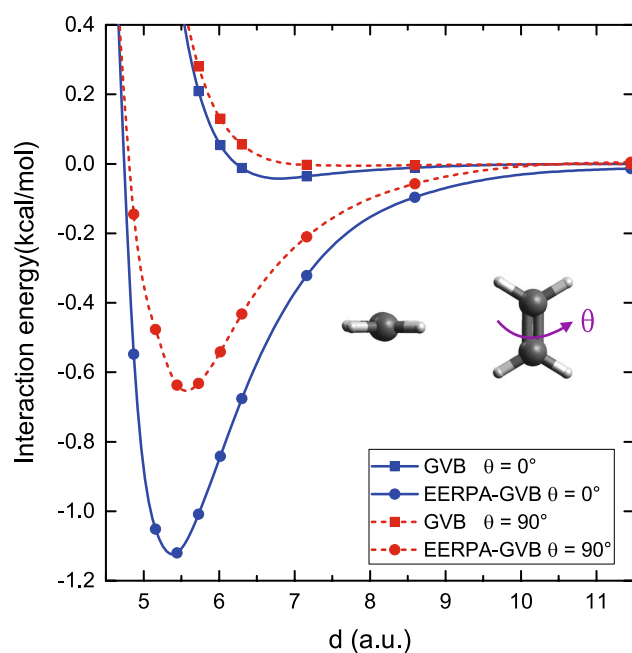


Fig. 6 Ethylene dimer interaction energy curves computed in aug-cc-pVDZ basis set

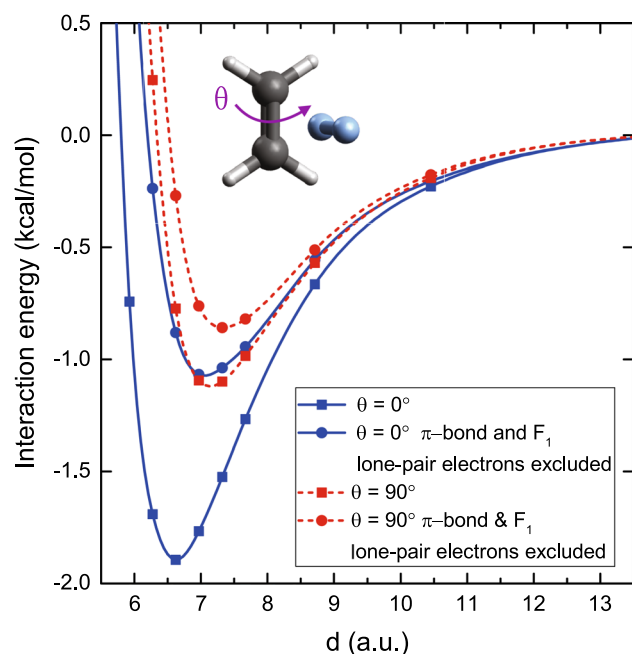


Fig. 7 Ethylene-fluorine complex EERPA-GVB interaction energy curves computed in aug-cc-pVDZ basis set

complex of ethylene and fluorine and how this changes upon the twisting of ethylene molecule. In Fig. 7, we can see that for flat ethylene the π -LP interaction constitutes about 40% of the total binding. This observation does not hold for

twisted ethylene, where the binding energy is small and the role of the corresponding interaction (LP- p orbitals on ethylene) is minor. Hypothesis about the special character of π -bonds in non-covalent interactions is here again reaffirmed.

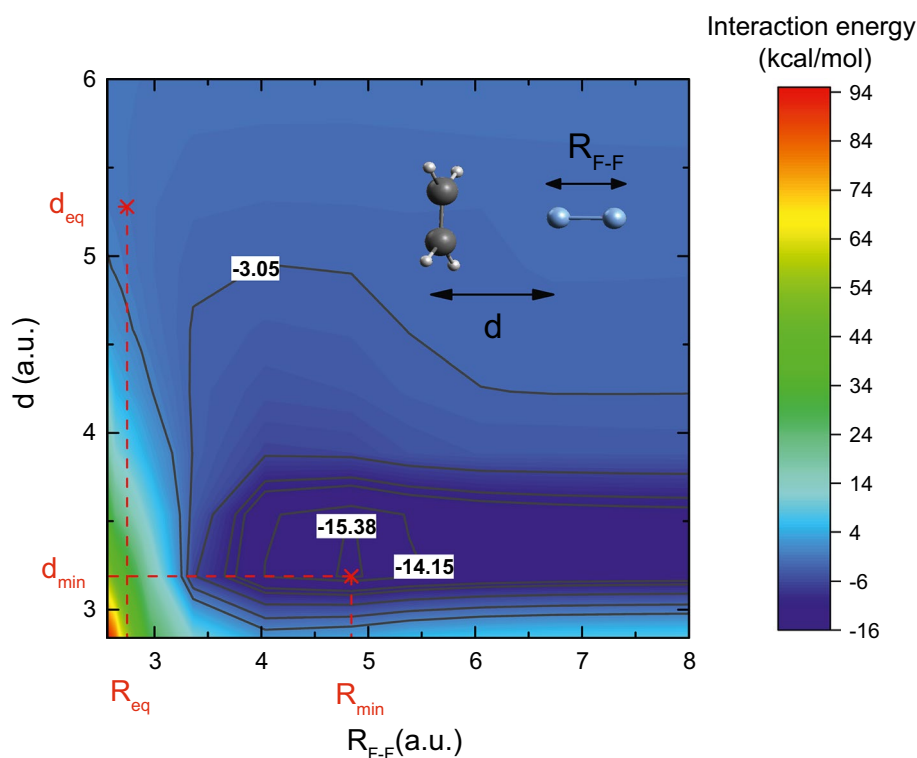
Studying such systems as the complex of fluorine molecule and twisted ethylene brings insight into the interactions of electrons forming different types of bonds, but what is more important, the interactions of molecules out-of-equilibrium geometry have very practical consequences. Namely, they facilitate (or obstruct) chemical reactions. Only recently it has finally been confirmed experimentally that by selective vibration excitation one can accelerate certain chemical reactions [52]. Such an acceleration due to throwing one of the reactants out of equilibrium can be a geometry-related effect (e.g., more favorable relative position of fragments of reactants taking part in the reaction) but it can also be an electronic-structure effect related to a bond twist or stretch.

Take, e.g., the reaction of ethylene fluorination. As the simplest example of organic molecule fluorination, it is interesting both for theorists and experimentalists. Despite its simplicity, there is a large discrepancy between the experimentally and theoretically determined reaction barrier heights [53, 54]. The experimentally observed barrier is lower than those obtained by state-of-the-art theoretical approaches, and it was hypothesized that the thermal vibrations of F_2 molecule may promote reaction [53].

We studied a T-shaped structure of the $C_2H_4 \cdots F_2$ complex at different intermolecular distances and F_2 bond lengths (see Fig. 8). One can immediately see that as F_2 bond is stretched, the vdW minimum deepens. A maximal attraction is achieved when the F_2 bond is stretched to c.a. 4.80 a.u. and the distance between the monomers is only 3.20 a.u. While in real systems vibrations do not cause the bonds to be stretched this much, even close to the equilibrium along this stretching mode, the interaction energy grows significantly enough to have impact on the reaction barrier. Additionally, we see that the optimal intermolecular distance diminishes along this stretching mode. While those observations do not allow one to determine the height of the fluorination barrier energy, they do highlight the importance of accurate description of non-covalent interactions, while the molecules are attempting a reaction. Since even highly sophisticated coupled-cluster methods frequently fail at this task [22], it is unsurprising that the computed value of a reaction barrier may be inaccurate.

We have already observed a similar behavior (i.e., attraction growth accompanying a bond stretching) for other vdW complexes in Ref. [22], where we explained it by the increased polarizability of the electrons in the stretched bond region and an electron density buildup in the region between monomers. This causes a rise in both

Fig. 8 Interaction energy map for the $C_2H_4 \cdots F_2$ complex. d indicates a distance between a center of mass of the ethylene molecule with the fixed geometry ($R_{CC} = 2.53$ a.u., $R_{CH} = 2.05$ a.u., $\theta_{HCH} = 117^\circ$) and the position of the F nucleus. R_{F-F} denotes the varied bond length. R_{eq} and d_{eq} correspond to the geometry of the lowest energy, whereas R_{min} and d_{min} to that corresponding to the lowest interaction energy



Pauli repulsion and dispersion interaction, which explains the existence of a minimum in the interaction energy along the F_2 bond stretch mode.

3 Conclusions

We have shown that the Embedding Extended Random Phase Approximation GVB method produces results on par with CCSD(T) for dispersion-dominated van der Waals complexes and is similar in accuracy to SAPT(CCD) when it comes to hydrogen-bonded systems. The method is particularly useful when it comes to non-covalently bonded complexes involving molecules out of their equilibrium geometries, as it is able to simultaneously capture both the energetic effects of bond stretching and twisting and more subtle van der Waals interactions.

To showcase this advantage, we have employed EERPA-GVB to study two unusual vdW systems, for which single-reference methods like MP2 break down. The first, a T-shaped ethylene dimer where one of the $C=C$ bonds was twisted, was compared to its classic, flat counterpart, which highlighted the importance of the $\pi-\sigma$ interaction and the particular role that π electrons play in dispersion interactions.

Role of π electrons was shown to be equally prominent in another studied system: a $C_2H_4 \cdots F_2$ complex. We have demonstrated the significance of a lone pairs- π

interaction, by not only comparing the interaction energy for complexes of twisted and flat ethylene with fluorine, but also by decomposing the inter-monomer correlation energy expression into contributions from interactions between pairs of geminals. Such energy decomposition is also possible for any other system, and since geminals are usually localized on bonds and atoms, it is an excellent and intuitive interpretive tool. It could be employed, e.g., to investigate also $\pi-\pi$, $\pi-\sigma$ and other types of interactions.

Finally, we have shown that stretching the F_2 bond in the same ethylene-fluorine complex causes a significant deepening of the vdW minimum, which is a result of a rise in the dispersion interaction. The enhanced attraction between the molecules may facilitate the ethylene fluorination reaction when the fluorine molecule is thermally excited to a stretching vibrational mode.

We conclude that EERPA-GVB is a useful tool to study molecular interaction qualitatively and quantitatively when bonds stretching, breaking or twisting is involved. This area is largely unexplored due to the lack of theoretical methods of both sufficient accuracy and modest computational cost. EERPA-GVB fills this gap in the computational chemistry toolbox.

Finally, it is worth mentioning that the concept of embedding a group of electrons of one monomer in a field created by electrons in another monomer, exploited in the EERPA correlation correction, can be applied to smaller

localized entities, i.e., to geminals. Such an approach would extend the applicability of EERPA to any system and is under development in our group.

Acknowledgements This work was supported by the National Science Center of Poland under Grant Nos. 2016/23/B/ST4/02848 and 2017/26/D/ST4/00780.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Woods L, Dalvit D, Tkatchenko A, Rodriguez-Lopez P, Rodriguez A, Podgornik R (2016) *Rev Mod Phys* 88(4):045003/1
- Wheeler SE, Seguin TJ, Guan Y, Doney AC (2016) *Acc Chem Res* 49(5):1061
- Pollino JM, Weck M (2005) *Chem Soc Rev* 34(3):193
- Chang XY, Ehlich R, Hudson AJ, Piecuch P, Polanyi JC (1997) *Faraday Discuss* 108:411
- Fabrizio A, Corminboeuf C (2018) *J Chem Phys Lett* 9(3):464
- Heindl AH, Wende RC, Wegner HA (2018) *Beilstein J Org Chem* 14(1):1238
- Liakos DG, Neese F (2015) *J Chem Theory Comput* 11(9):4054
- Piecuch P, Włoch M (2005) *J Chem Phys* 123(22):224105
- Hohenberg P, Kohn W (1964) *Phys Rev* 136(3B):B864
- Grimme S, Hansen A, Brandenburg JG, Bannwarth C (2016) *Chem Rev* 116(9):5105
- Jeziorski B, Moszynski R, Szalewicz K (1994) *Chem Rev* 94(7):1887
- Jansen G (2013) *WIREs Comput Mol Sci* 4(2):127
- Mo Y, Gao J, Peyerimhoff SD (2000) *J Chem Phys* 112(13):5530
- Khaliullin RZ, Cobar EA, Lochan RC, Bell AT, Head-Gordon M (2007) *J Phys Chem A* 111(36):8753
- Parrish RM, Gonthier JF, Corminboeuf C, Sherrill CD (2015) *J Chem Phys* 143(5):051103
- Pastorczyk E, Prlj A, Gonthier JF, Corminboeuf C (2015) *J Chem Phys* 143(22):224107
- Skouteris D, Manolopoulos DE, Bian W, Werner HJ, Lai LH, Liu K (1999) *Science* 286(5445):1713
- Xie T, Wang D, Bowman JM, Manolopoulos DE (2002) *J Chem Phys* 116(17):7461
- Shen Z, Ma H, Zhang C, Fu M, Wu Y, Bian W, Cao J (2017) *Nat Commun* 8:1
- Jankunas J, Sneha M, Zare RN, Bouakline F, Althorpe SC, Herraiz-Aguilar D, Aoiz FJ (2014) *Proc Natl Acad Sci USA* 111(1):15
- Li J, Jiang B, Guo H (2013) *Chem Sci* 4:629
- Pastorczyk E, Shen J, Hapka M, Piecuch P, Pernal K (2017) *J Chem Theory Comput* 13:5404
- Zoboki T, Szabados A, Surján P (2013) *J Chem Theory Comput* 9:2602
- Rassolov VA (2002) *J Chem Phys* 117(13):5978
- McWeeny R (1959) *Proc R Soc A* 253:242
- Bobrowicz FW, Goddard WA (1977) In: Schaefer HF III (ed) *Modern Theoretical Chemistry: Methods of Electronic Structure Theory*. Plenum, New York, pp 79–127
- Hurley AC, Lennard-Jones JE, Pople JA (1953) *Proc R Soc A* 220:446
- Kutzelnigg W (1964) *J Chem Phys* 40:3640
- Arai T (1960) *J Chem Phys* 33:95
- Chatterjee K, Pastorczyk E, Jawulski K, Pernal K (2016) *J Chem Phys* 144:244111/1
- Surján PR (1999) *Topics in current chemistry*, vol 203. Springer, Berlin, pp 63–88
- Surján P, Szabados A, Jeszenszki P, Zoboki T (2012) *J Math Chem* 50:534
- Jeszenszki P, Rassolov V, Surján P, Szabados A (2015) *Mol Phys* 113:249
- Pastorczyk E, Pernal K (2015) *Phys Chem Chem Phys* 17:8622
- McLachlan AD, Ball MA (1964) *Rev Mod Phys* 36:844
- Gambacurta D, Catara F (2008) *Phys Rev B* 77:205434
- Chatterjee K, Pernal K (2012) *J Chem Phys* 137:204109/1
- Chatterjee K, Pernal K (2015) *Theor Chem Acc* 134:118
- Pernal K (2014) *J Chem Theory Comput* 10:4332
- Pernal K (2018) *Phys Rev Lett* 120:013001
- Pernal K (2016) *Phys Chem Chem Phys* 18:21111
- Kendall RA, Dunning Jr TH, Harrison RJ (1992) *J Chem Phys* 96(9):6796
- Aidas K, Angeli C, Bak KL, Bakken V, Bast R, Boman L, Christiansen O, Cimiraglia R, Coriani S, Dahle P, Dalskov EK, Ekström U, Enevoldsen T, Eriksen JJ, Ettenhuber P, Fernández B, Ferrighi L, Fliegl H, Frediani L, Hald K, Halkier A, Hättig C, Heiberg H, Helgaker T, Hennum AC, Hettema H, Hjertenæs E, Høst S, Høyvik IM, Iozzi MF, Jansik B, Jensen HJA, Jonsson D, Jørgensen P, Kauczor J, Kirpekar S, Kjergaard T, Klopper W, Knecht S, Kobayashi R, Koch H, Kongsted J, Krapp A, Kristensen K, Ligabue A, Lutnæs OB, Melo JI, Mikkelsen KV, Myhre RH, Neiss C, Nielsen CB, Norman P, Olsen J, Olsen JMH, Osted A, Packer MJ, Pawłowski F, Pedersen TB, Provasi PF, Reine S, Rinkevicius Z, Ruden TA, Ruud K, Rybakin VV, Sałek P, Samson CCM, de Merás AS, Saue T, Sauer SPA, Schimmelpfennig B, Sneskov K, Steindal AH, Sylvester-Hvid KO, Taylor PR, Teale AM, Tellgren EI, Tew DP, Thorvaldsen AJ, Thøgersen L, Vahtras O, Watson MA, Wilson DJD, Ziolkowski M, Ågren H (2014) *WIREs Comput Mol Sci* 4(3):269
- Hohenstein EG, Sherrill CD (2010) *J Chem Phys* 133(1):014101
- Parrish RM, Burns LA, Smith DGA, Simmonett AC, DePrince AE, Hohenstein EG, Bozkaya U, Sokolov AY, Di Remigio R, Richard RM, Gonthier JF, James AM, McAlexander HR, Kumar A, Saitow M, Wang X, Pritchard BP, Verma P, Schaefer HF, Patkowski K, King RA, Valeev EF, Evangelista FA, Turney JM, Crawford TD, Sherrill CD (2017) *J Chem Theory Comput* 13(7):3185
- Boys SF, Bernardi F (1970) *Mol Phys* 19(4):553
- Zhao Y, Truhlar DG (2005) *J Phys Chem A* 109(25):5656
- Gdanitz RJ (2000) *J Chem Phys* 113(13):5145
- Dulieu O, Osterwalder A (2017) *Cold chemistry: molecular scattering and reactivity near absolute zero*, vol 11. Royal Society of Chemistry, Cambridge
- Cao Y, Wong MW (2014) *J Mol Model* 20(4):2185
- Hermann J, Alfe D, Tkatchenko A (2017) *Nat Commun* 8:14052
- Stensitzki T, Yang Y, Kozich V, Ahmed AA, Kössl F, Kühn O, Heyne K (2018) *Nat Chem* 10:126
- Lu YJ, Xie T, Fang JW, Shao HC, Lin JJ (2008) *J Chem Phys* 128(18):184302
- Feng H, Allen WD (2010) *J Chem Phys* 132(9):094304