### Title

Coupled Cluster Valence Bond Method: Efficient Computer Implementation and Application to Multiple Bond Dissociations and Strong Correlations in the Acenes.




### Journal

Journal of chemical theory and computation, 10(5)

### ISSN

1549-9618



### Authors

Small, David W
Lawler, Keith V
Head-Gordon, Martin



### Publication Date

2014-05-01

### DOI

10.1021/ct500112y

Peer reviewed

# Coupled Cluster Valence Bond method: Efficient computer implementation and application to multiple bond dissociations and strong correlations in the acenes



| | |
|---|---|
| Journal: | *Journal of Chemical Theory and Computation* |
| Manuscript ID: | ct-2014-00112y.R1 |
| Manuscript Type: | Article |
| Date Submitted by the Author: | 19-Apr-2014 |
| Complete List of Authors: | Small, David; UC Berkeley, Chemistry<br>Lawler, Keith; University of Nevada Las Vegas, Department of Chemistry<br>Head-Gordon, Martin; University of California, Berkeley, Chemistry |
| | |

# Coupled Cluster Valence Bond method: Efficient computer implementation and application to multiple bond dissociations and strong correlations in the acenes

David W. Small and Keith V. Lawler and Martin Head-Gordon

*Department of Chemistry, University of California, Berkeley, California 94720 and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720*

(Dated: April 19, 2014)

## Abstract

We describe an efficient implementation of the Coupled Cluster Valence Bond (CCVB) model. CCVB captures a certain essential part of the description of molecules with strong correlations (SC), which allows it to achieve correct energy profiles when covalent bonds are broken, while maintaining proper spin symmetry and size extensivity. To illustrate treatment of SC in bond-breaking, we examine the symmetric dissociation of the sulfur allotropes $S_6$ and $S_8$ into triplet S atoms. To show applicability to larger systems and to explore whether CCVB can capture aspects of SC that arise in extended $\pi$ systems, we report results for a series of acenes up to 12 fused benzene rings, with active spaces of up to 228 correlated electrons. The lowest-energy CCVB solutions found for two of the largest acenes show signatures consistent with multi-electron SC and partial delocalization.

## I. INTRODUCTION

The computational successes of Electronic Structure Theory (EST) for large molecular systems are longstanding. In particular, this is due to substantial algorithmic advancements in 2nd-order Moller-Plesset Perturbation Theory and numerous Kohn-Sham density functionals.[1–6] However, the accuracy of these methods is not uniform across diverse collections of systems of comparable size. The occurence of these qualitative errors is the essence of the concept of strong correlation (SC).

The definitive EST approaches to SC systems are the multireference (MR) methods.[7–14] Yet these models are computationally intractable for large systems,[15] so for the latter we are relegated to incorporating the extant tractable elements of the MR treatment into the more scalable methods.

Old and well-studied examples of this process are found in the strongly-orthogonal[16] geminal models.[17–31] These methods add MR correlation to the electron pairs of Hartree Fock (HF) without removing the latters' semi-independence, i.e. correlation is introduced within but not between pairs. The most demonstrated model in this category is the simplest one: Generalized Valence Bond Perfect Pairing (GVB-PP).[32–52]

We think GVB-PP connects well to the classic Lewis-structure picture of electronic structure, specifically if just one dot structure is used. Thus, for GVB-PP, the geminal pairs correspond to bonds and lone pairs. GVB-PP incorporates the minimal MR ingredients necessary for a qualitatively correct description of individual covalent bonds; according to GVB-PP theory, this constitutes the leading contribution to SC.

Of course, electron pairs will correlate, thus forming the basis for higher-order contributions to SC and leading to significant difficulties for traditional methods. A good example of this, and one particularly familiar to chemists, is the correlation between covalent bonds as they are stretched and broken, such as is found in double and triple bonds.

One effective way of understanding this type of correlation comes by observing that in the course of bond-breaking, electron configurations in which the pairs are individually converted to triplet spin and then coupled together into overall singlets become non-negligible. Since GVB-PP and restricted HF (RHF) omit these elements, their energy surfaces in the stretched-bond regime do not reflect the physically correct picture of coupled atoms or molecular fragments. In contrast, the correct *dissociation limit* energies can generally be

obtained via unrestricted HF (UHF) or DFT. This is because these models partially incorporate the requisite triplet character of the pairs. But what about the regime intermediate between equilibrium and full dissociation? The unrestricted models do not include the above-mentioned overall singlet coupling of the triplet pairs. Thus an unrestriction byproduct is the incorporation of higher-spin contaminants. In many cases, these exited states lie noticeably above the ground state, so the quality of the results is diluted in proportion to the extent of unrestriction. A clear example of this will be encountered later in this paper.

One way of dealing with the preceding problem is to project out the unwanted spin contaminants. This is spin-projected HF (SPHF), a method with a long history[53–66] and one that is enjoying a resurgence thanks to efforts in the Scuseria group.[67–72] State-of-the-art SPHF has several desirable properties, including variationality and computational efficiency. However, the method suffers from a lack of size extensivity. This problem is removed in the Coupled Cluster Valence Bond (CCVB) model.[73,74] CCVB was derived as a generalization of spin-projected UHF (SPUHF); indeed, any SPUHF wavefunction is a CCVB wavefunction (c.f. sections III and IV in ref. 73). In CCVB, size extensivity is obtained by replacing the variational energy expression of SPHF with a formally non-variational coupled-cluster-type approach. By doing this, the advantageous orbital invariance of SPHF is lost, as will be discussed in a different light below. We should note that we have yet to observe non-variational CCVB behavior in practice.

To summarize the above ideas, an essential part of the proper treatment of the bond-breaking SC phenomenon comprises spin couplings of (potentially numerous) collections of strongly correlated pairs. The CCVB approach for this is to spin couple pair clusters of varying sizes in a self consistent manner. This scheme is both efficient and qualitatively correct, and the result is a low-scaling method that respects spin symmetry, is size-extensive, and dissociates systems to the open-shell HF energies of the pertinent atoms or fragments.

CCVB leaves the GVB-PP pair division intact, and so it retains the single-Lewis-structure picture of the latter. But this also implies that our interpretation of this picture now allows for stretched bonds, as long as one can still reasonably parse the associated electrons into pairs. This would seem to indicate that the assortment of systems to which CCVB can be positively applied is reasonably diverse.

Of course, not all molecular systems are well described by a single Lewis structure, and, thus, the CCVB type of correlation is by no means the only source of SC. Indeed, the lack

of proper accomodation of multiple structures is a pervasive shortcoming of the methods based on GVB-PP,[75,76] and CCVB is no exception.[74] This is the oribital invariance problem mentioned above. Nevertheless, even for these sorts of systems, the CCVB-type correlation still represents an essential step along the path toward a robust treatment of SC. Furthermore, CCVB is able to partially incorporate multiple-structure effects, as we have shown previously for the cases of cyclic $H_4$ and $H_8$, and benzene.[73,74] Additional support for this will be given later in the present paper when we look at a series of acenes.

The purpose of this paper is to describe our efforts in translating the CCVB idea into an efficient computer program that can be applied to decently large systems. The latter will be demonstrated via example calculations after we describe the pertinent theoretical details in the following section.

## II. THEORY

We will begin with the assumption that a singlet state is being approximated. In a subsequent paper, we will show how the resulting theoretical development may be used to derive an approach for open-shell cases. In what follows, $n_e$ is the number of electrons and $n_b$ is the number of basis orbitals.

### A. CCVB wavefunction, energy, and amplitude equations for a given set of orbitals

#### *1. Basic equations*

The leading contribution to the CCVB wavefunction is the GVB-PP wavefunction. We index the latter's pairs from 1 to $n_o = \frac{1}{2}n_e$. The electrons are further classified as being either "core" or "active", which we will elaborate on shortly. We define $n_c$ ($n_p$) to be the number of core (active) electrons divided by two. Electron pairs 1 through $n_c$ are core, and the remaining pairs are active.

Each electron pair is described by a geminal wavefunction. These geminals are of the simplest possible type: the core pairs are treated the same as they are in RHF, and for the active electrons, each geminal is a 2 electrons in 2 orbitals, or (2,2), singlet wavefunction.

Thus, if the $k$th electron pair is active, with it are associated two spatial orbitals, which we label $\phi_k$ and $\phi_{\hat{k}}$. In our implementation, we define $\hat{k} = n_o + n_p - (k - n_c - 1)$. We thus have a set of orbitals, $\{\phi_q \mid q = 1 \ldots n_o + n_p\}$. This set is orthonormal and thus the orbitals have the "strong orthogonality" property: the orbitals for one pair are orthogonal to those for any other pair. The set is extended to a complete orthonormal basis for the molecular-orbital space, $\{\phi_q \mid q = 1 \ldots n_b\}$. The orbitals added for this extension do not directly affect the CCVB wavefunction, but they are integral to orbital optimization, which will be discussed later. The orbitals may be expressed relative to the atomic-orbital basis

$$\phi_q = \sum_{p=1}^{n_b} C_{pq} \zeta_p. \tag{1}$$

Is it convenient to work in second quantization, and therein we have operators for the associated spin orbitals: $\mathbf{a}^\dagger_{k_\alpha}$, $\mathbf{a}^\dagger_{k_\beta}$, $\mathbf{a}^\dagger_{\hat{k}_\alpha}$, and $\mathbf{a}^\dagger_{\hat{k}_\beta}$, along with their annihilation counterparts, where $\alpha$ and $\beta$ are the usual one-electron spin states. Each active geminal is created by an operator

$$\mathbf{g}^\dagger_{\mathbf{s},k} = \frac{1}{\sqrt{2(1+\cos^2(\theta_k))}} \left( 2\cos(\theta_k)\, \mathbf{a}^\dagger_{k_\alpha} \mathbf{a}^\dagger_{k_\beta} - \sin(\theta_k)\, \mathbf{a}^\dagger_{k_\alpha} \mathbf{a}^\dagger_{\hat{k}_\beta} - \sin(\theta_k)\, \mathbf{a}^\dagger_{\hat{k}_\alpha} \mathbf{a}^\dagger_{k_\beta} \right), \tag{2}$$

where $\theta_k$ is a parameter. For simplicity, we will endeavor to unify our treatment of core and active pairs, so for core-pair $i$, we may implicitly use the preceding equation by setting $\theta_i = 0$. In what follows, expressions and terms involving active-pair elements will often also be applicable using core elements, if the parts with hatted indices are ignored.

Equation (2) is derived from the valence-bond (VB) form of the PP singlet geminals. For each active pair, we define two VB orbitals:

$$\begin{aligned} \psi^+_k &= \phi_k \\ \psi^-_k &= \cos(\theta_k)\phi_k - \sin(\theta_k)\phi_{\hat{k}}. \end{aligned} \tag{3}$$

With these, we obtain a geminal, $\psi^+_k \psi^-_k (\alpha\beta - \beta\alpha)$, which is a scalar multiple of the one associated with equation (2). This may be compared to the more common natural-orbital form of the PP geminals, which, in the second-quantization perspective, would be like

$$c_1\, \mathbf{a}^\dagger_{k_\alpha} \mathbf{a}^\dagger_{k_\beta} + c_2\, \mathbf{a}^\dagger_{\hat{k}_\alpha} \mathbf{a}^\dagger_{\hat{k}_\beta}, \tag{4}$$

for appropriately chosen orbitals. The VB form has a certain interpretive advantage, and also it provides a connection to unrestricted Hartree Fock, whose wavefunction may be written as a product of strongly-orthogonal spin-contaminated geminals of the form $\psi^+_k \psi^-_k \alpha\beta$.[77]

We may now express the GVB-PP wavefunction:

$$|\Phi_0\rangle = \prod_{k=1}^{n_o} \mathbf{g}^\dagger_{\mathbf{s},k}|0\rangle. \tag{5}$$

For active-pair $k$, the associated (2,2) active space also contains triplet-spin geminals, and these are used in CCVB. They are given by

$$\begin{aligned}
\mathbf{g}^\dagger_{\mathbf{t_1},k} &= \frac{1}{\sqrt{2}}(-\mathbf{a}^\dagger_{k_\alpha}\mathbf{a}^\dagger_{\hat{k}_\beta} + \mathbf{a}^\dagger_{\hat{k}_\alpha}\mathbf{a}^\dagger_{k_\beta}) \\
\mathbf{g}^\dagger_{\mathbf{t_2},k} &= \mathbf{a}^\dagger_{k_\alpha}\mathbf{a}^\dagger_{\hat{k}_\alpha} \\
\mathbf{g}^\dagger_{\mathbf{t_3},k} &= \mathbf{a}^\dagger_{k_\beta}\mathbf{a}^\dagger_{\hat{k}_\beta}.
\end{aligned} \tag{6}$$

From the strong orthogonality, any two geminal operators belonging to different pairs commute, whether they are of the creation or of the annihilation type. We also have orthonormality relationships between the geminals of one pair,

$$\mathbf{g}_{\mathbf{s},k}\mathbf{g}^\dagger_{\mathbf{s},k}|0\rangle = 1, \tag{7}$$

$$\mathbf{g}_{\mathbf{t_i},k}\mathbf{g}^\dagger_{\mathbf{s},k}|0\rangle = \mathbf{g}_{\mathbf{s},k}\mathbf{g}^\dagger_{\mathbf{t_i},k}|0\rangle = 0, \quad \text{for every } i\ (1 \le i \le 3) \tag{8}$$

and

$$\mathbf{g}_{\mathbf{t_i},k}\mathbf{g}^\dagger_{\mathbf{t_j},k}|0\rangle = \delta_{ij}|0\rangle, \quad \text{for every } i,j\ (1 \le i,j \le 3). \tag{9}$$

The properties listed in this paragraph will allow straightforward computation of the relevant Hamiltonian ($\mathbf{H}$) matrix elements.

We now introduce the CCVB wavefunction:

$$\begin{aligned}
|\Psi\rangle = |\Phi_0\rangle + \sum_{k<l} t_{kl}|\Phi_{(kl)}\rangle + \sum_{k<l<m<n} \Big[ & t_{kl}t_{mn}|\Phi_{(kl)(mn)}\rangle \\
& + t_{km}t_{ln}|\Phi_{(km)(ln)}\rangle + t_{kn}t_{lm}|\Phi_{(kn)(lm)}\rangle \Big] + \cdots,
\end{aligned} \tag{10}$$

where the $t$'s are cluster amplitudes and the associated kets are certain electronic configurations that depend only on the $\theta_k$ and $C_{pq}$ variables. Below, we will give definitions for these kets as needed. Here, and in what follows, any subscript index associated with a cluster amplitude is active. We have shown only up to the terms that are quadratic in the amplitudes. The expansion in eq. (10) continues to all possible higher-order terms, but only the ones shown are needed for the CCVB amplitude equations.

At this point, we need to define the energy and establish a procedure to compute the amplitudes. For this, it is helpful to highlight CCVB's similarity to regular coupled cluster (CC) theory. In regular CC, each configuration is a linear combination of Slater determinants. In CCVB, instead of spin-orbital products, we use products of strongly orthogonal geminals. Like CC Doubles (CCD), CCVB is a cluster expansion relative to a configuration hierarchy: a reference $|\Phi_0\rangle$, doubly substituted configurations $|\Phi_{(kl)}\rangle$, quadruply substituted configurations $|\Phi_{(kl)(mn)}\rangle$, etc. Here, substitutions mean the removal of active singlet geminals from the reference and their replacement with triplet geminals from the same pairs in a singlet-coupled fashion. As with spin-orbital-based theories, non-zero Hamiltonian matrix elements are produced only when the relative substitution level between the left and right configurations is double or lower. The CCVB energy $E$ has the same form as that used in regular CC:

$$E = \langle \Phi_0 | \mathbf{H} | \Psi \rangle, \tag{11}$$

where $\mathbf{H}$ is the Hamiltonian operator. As is done in CCD, we project an eigenvalue equation with doubly substituted bras to obtain the CCVB amplitude equations

$$\langle \Phi_{(kl)} | \mathbf{H} | \Psi \rangle = E \langle \Phi_{(kl)} | \Psi \rangle. \tag{12}$$

It may be shown that

$$E = \langle \Phi_0 | \mathbf{H} | \Phi_0 \rangle + \sum_{k<l} t_{kl} \mu_{kl} \tag{13}$$

and that eq. (12) is equivalent to

$$\begin{aligned} 0 = \Omega_{kl} := \mu_{kl}(1 - t_{kl}^2) + t_{kl}\omega_{kl} + \\ \sum_{m \notin \{k,l\}} \Big[ t_{km}(\kappa_{lm} - t_{kl}\mu_{km}) + t_{lm}(\kappa_{km} - t_{kl}\mu_{lm}) \Big], \end{aligned} \tag{14}$$

where $\kappa_{lm} = \langle \Phi_{(kl)} | \mathbf{H} | \Phi_{(km)} \rangle$, $\mu_{mn} = \langle \Phi_0 | \mathbf{H} | \Phi_{(mn)} \rangle$, and $\omega_{kl} = \langle \Phi_{(kl)} | \mathbf{H} | \Phi_{(kl)} \rangle - \langle \Phi_0 | \mathbf{H} | \Phi_0 \rangle$. At this point, we have outlined the basic constituents of the CCVB approximation, and can now begin to discuss ways to calculate them.

*2. Computable expressions for the constituents*

From the definitions for the $\kappa$, $\mu$, and $\omega$ quantities, it is clear that we must explicitly define the doubly substituted configurations. We first note that

$$\mathbf{g}_{\mathbf{s},k}|\Phi_0\rangle = \prod_{\substack{m=1 \\ m\neq k}}^{n_o} \mathbf{g}^\dagger_{\mathbf{s},m}|0\rangle, \tag{15}$$

which follows from the orthonormality and commutation properties listed above. Further application of annihilation geminals will serve to remove their creation counterparts.

Next, we define an operator,

$$\mathbf{d}^\dagger_{\mathbf{s_2},kl} = \frac{1}{\sqrt{3}}\big(\mathbf{g}^\dagger_{\mathbf{t_1},k}\mathbf{g}^\dagger_{\mathbf{t_1},l} - \mathbf{g}^\dagger_{\mathbf{t_2},k}\mathbf{g}^\dagger_{\mathbf{t_3},l} - \mathbf{g}^\dagger_{\mathbf{t_3},k}\mathbf{g}^\dagger_{\mathbf{t_2},l}\big), \tag{16}$$

which may be used to define

$$|\Phi_{(kl)}\rangle = \mathbf{d}^\dagger_{\mathbf{s_2},kl}\mathbf{g}_{\mathbf{s},k}\mathbf{g}_{\mathbf{s},l}|\Phi_0\rangle. \tag{17}$$

The computation of the matrix elements may be simplified by using spin projection. Let

$$|\Phi_{[k]_2[l]_3}\rangle = \mathbf{g}^\dagger_{\mathbf{t_2},k}\mathbf{g}^\dagger_{\mathbf{t_3},l}\mathbf{g}_{\mathbf{s},k}\mathbf{g}_{\mathbf{s},l}|\Phi_0\rangle. \tag{18}$$

Like $|\Phi_0\rangle$, this is a product of geminals. In addition,

$$\mathbf{P}_\mathrm{S}\left(-\sqrt{3}|\Phi_{[k]_2[l]_3}\rangle\right) = |\Phi_{(kl)}\rangle, \tag{19}$$

where $\mathbf{P}_\mathrm{S}$ projects into the singlet subspace. Because $\mathbf{P}_\mathrm{S}$ commutes with $\mathbf{H}$ and $|\Phi_0\rangle$ is a singlet, we have

$$\mu_{kl} = -\sqrt{3}\langle\Phi_0|\mathbf{H}|\Phi_{[k]_2[l]_3}\rangle. \tag{20}$$

Here, the ket has triplets for the $k$th and $l$th geminals, which must be "blocked" by $\mathbf{H}$ if a non-zero result is to be produced. In second quantization we have

$$\mathbf{H} = \sum_{p,q} h_{pq}\,\mathbf{a}^\dagger_p\mathbf{a}_q + \frac{1}{4}\sum_{p,q,r,s}\langle pq||sr\rangle\,\mathbf{a}^\dagger_p\mathbf{a}^\dagger_q\mathbf{a}_r\mathbf{a}_s, \tag{21}$$

where the summations go over all spin-orbital indices. Therefore, in this case $\mathbf{H}$ must annihilate one pair-$k$ orbital and one pair-$l$ orbital, and replace these with an orbital from each pair. In other words,

$$\langle\Phi_0|\mathbf{H}|\Phi_{[k]_2[l]_3}\rangle = \sum_{p,r\in S_k}\sum_{q,s\in S_l}\langle 0|\mathbf{g}_{\mathbf{s},k}\mathbf{g}_{\mathbf{s},l}\,\mathbf{a}^\dagger_p\mathbf{a}_r\mathbf{a}^\dagger_q\mathbf{a}_s\,\mathbf{g}^\dagger_{\mathbf{t_2},k}\mathbf{g}^\dagger_{\mathbf{t_3},l}|0\rangle\langle pq||rs\rangle, \tag{22}$$

where $S_k = \{k_\alpha, \hat{k}_\alpha, k_\beta, \hat{k}_\beta\}$. To simplify and generalize, we introduce density matrices

$$P_{k;\,\mathbf{xy};\,pr} = \langle 0|\mathbf{g}_{\mathbf{x},k}\mathbf{a}_p^\dagger\mathbf{a}_r\mathbf{g}_{\mathbf{y},k}^\dagger|0\rangle, \tag{23}$$

where $\mathbf{x}, \mathbf{y} \in \{\mathbf{s}, \mathbf{t_1}, \mathbf{t_2}, \mathbf{t_3}\}$ and $p, r \in S_k$. This definition may be extended to include the core pairs by letting $S_m = \{m_\alpha, m_\beta\}$ for core index $m$. In that case, we have only one density matrix, $P_{m;\,\mathbf{ss};\,pr}$. The density-matrix elements for various choices of $\mathbf{xy}$ are given in the Supporting Information (SI).

Next we define

$$\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}} = \sum_{p,r\in S_k}\sum_{q,s\in S_l}\langle pq||rs\rangle P_{k;\,\mathbf{wx};\,pr}P_{l;\,\mathbf{yz};\,qs}, \tag{24}$$

and obtain

$$\mu_{kl} = -\sqrt{3}\,\sigma_{kl;\,\mathbf{st_2};\,\mathbf{st_3}}. \tag{25}$$

Similar reasoning produces

$$\begin{aligned}\kappa_{lm} &= -\sqrt{3}\langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{(km)}\rangle \\ &= \langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{[k]_2[m]_3}\rangle \\ &= \sigma_{lm;\,\mathbf{t_3s};\,\mathbf{st_3}}.\end{aligned} \tag{26}$$

We reiterate that the pair indices for the $\sigma$'s may be core or inactive, while the $\mu$'s and $\kappa$'s (and $\omega$'s) are only defined for active indices.

The $\sigma$'s have the following symmetries:

$$\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}} = \sigma_{lk;\,\mathbf{yz};\,\mathbf{wx}} \tag{27}$$

and

$$\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}} = \sigma_{kl;\,\mathbf{xw};\,\mathbf{zy}}, \tag{28}$$

the latter being the result of symmetry in the density matrices,

$$P_{k;\,\mathbf{wx};\,pr} = P_{k;\,\mathbf{xw};\,rp}. \tag{29}$$

This implies $\mu_{kl} = \mu_{lk}$ and $\kappa_{kl} = \kappa_{lk}$.

Turning to $\langle\Phi_0|\mathbf{H}|\Phi_0\rangle$, we note that this quantity is a sum of 1) interpair coulombic terms, which involve $\sigma$'s, 2) one-electron contributions, for which we define

$$\eta_{k;\,\mathbf{wx}} = \sum_{p,r\in S_k} h_{pr}P_{k;\,\mathbf{wx};\,pr}, \tag{30}$$

and 3) intrapair coulombic terms, for which we define

$$\rho_{k;\,\mathbf{w};\,\mathbf{x}} = \sum_{p,q}\sum_{r,s}\langle pq||rs\rangle f_{k;\,\mathbf{w};\,pq} f_{k;\,\mathbf{x};\,rs}, \tag{31}$$

where the $f_{k;\,\mathbf{w}}$ contain the expansion coefficients for the geminals according to

$$\begin{aligned}\mathbf{g}^{\dagger}_{\mathbf{w},k} = {} & f_{k;\,\mathbf{w};\,k_\alpha k_\beta}\,\mathbf{a}^{\dagger}_{k_\alpha}\mathbf{a}^{\dagger}_{k_\beta} + f_{k;\,\mathbf{w};\,k_\alpha \hat{k}_\beta}\,\mathbf{a}^{\dagger}_{k_\alpha}\mathbf{a}^{\dagger}_{\hat{k}_\beta} + f_{k;\,\mathbf{w};\,\hat{k}_\alpha k_\beta}\,\mathbf{a}^{\dagger}_{\hat{k}_\alpha}\mathbf{a}^{\dagger}_{k_\beta} \\ & + f_{k;\,\mathbf{w};\,\hat{k}_\alpha \hat{k}_\beta}\,\mathbf{a}^{\dagger}_{\hat{k}_\alpha}\mathbf{a}^{\dagger}_{\hat{k}_\beta} + f_{k;\,\mathbf{w};\,k_\alpha \hat{k}_\alpha}\,\mathbf{a}^{\dagger}_{k_\alpha}\mathbf{a}^{\dagger}_{\hat{k}_\alpha} + f_{k;\,\mathbf{w};\,k_\beta \hat{k}_\beta}\,\mathbf{a}^{\dagger}_{k_\beta}\mathbf{a}^{\dagger}_{\hat{k}_\beta}.\end{aligned} \tag{32}$$

Here $\mathbf{w}$ is one of $\mathbf{s}$, $\mathbf{t_1}$, $\mathbf{t_2}$, or $\mathbf{t_3}$ and

$$\begin{aligned} f_{k;\,\mathbf{s}} &= \frac{1}{\sqrt{2(1+\cos^2(\theta_k))}}\Big(2\cos(\theta_k), -\sin(\theta_k), -\sin(\theta_k), 0, 0, 0\Big), \\ f_{k;\,\mathbf{t_1}} &= 2^{-\frac{1}{2}}\Big(0, -1, 1, 0, 0, 0\Big), \\ f_{k;\,\mathbf{t_2}} &= \Big(0, 0, 0, 0, 1, 0\Big), \\ f_{k;\,\mathbf{t_3}} &= \Big(0, 0, 0, 0, 0, 1\Big), \end{aligned} \tag{33}$$

where the ordering here is that implied in eq. (32). For core pairs, we informally define $f_{i;\,\mathbf{s}} = (1, 0, 0, 0, 0, 0)$. In eq. (31), the sum over $p, q$ corresponds to the 6 pairs of indices found in $f_{k;\,\mathbf{w}}$, and similarly for the sum over $r, s$. Note that the $\eta$ and $\rho$ quantities are defined for both core and active pairs.

We then obtain

$$E_0 = \langle\Phi_0|\mathbf{H}|\Phi_0\rangle = \sum_{k=1}^{n_o}(\eta_{k;\,\mathbf{ss}} + \rho_{k;\,\mathbf{s};\,\mathbf{s}}) + \sum_{\substack{k,l=1\\k<l}}^{n_o}\sigma_{kl;\,\mathbf{ss};\,\mathbf{ss}}. \tag{34}$$

To compute $\omega_{kl}$, we begin with

$$\begin{aligned}\langle\Phi_{(kl)}|\mathbf{H}|\Phi_{(kl)}\rangle &= -\sqrt{3}\langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{(kl)}\rangle \\ &= -\sigma_{kl;\,\mathbf{t_2}\mathbf{t_1};\,\mathbf{t_3}\mathbf{t_1}} + \langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{[k]_2[l]_3}\rangle + \sigma_{kl;\,\mathbf{t_2}\mathbf{t_3};\,\mathbf{t_3}\mathbf{t_2}}.\end{aligned} \tag{35}$$

We note that $\sigma_{kl;\,\mathbf{t_2}\mathbf{t_3};\,\mathbf{t_3}\mathbf{t_2}} = 0$ because the relevant density matrices are zero. We can obtain an expression for $\langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{[k]_2[l]_3}\rangle$ that is similar to eq. (34), and indeed these two summations will involve many of the same terms, namely those involving only singlet geminals. This makes it relatively easy to compute $\langle\Phi_{[k]_2[l]_3}|\mathbf{H}|\Phi_{[k]_2[l]_3}\rangle - \langle\Phi_0|\mathbf{H}|\Phi_0\rangle$, and we

have

$$\begin{aligned}\omega_{kl} = &-\sigma_{kl;\,\mathbf{t_2t_1};\,\mathbf{t_3t_1}} \\ &+ \eta_{k;\,\mathbf{t_2t_2}} + \eta_{l;\,\mathbf{t_3t_3}} - \eta_{k;\,\mathbf{ss}} - \eta_{l;\,\mathbf{ss}} \\ &+ \rho_{k;\,\mathbf{t_2};\,\mathbf{t_2}} + \rho_{l;\,\mathbf{t_3};\,\mathbf{t_3}} - \rho_{k;\,\mathbf{s};\,\mathbf{s}} - \rho_{l;\,\mathbf{s};\,\mathbf{s}} \\ &+ \sigma_{kl;\,\mathbf{t_2t_2};\,\mathbf{t_3t_3}} - \sigma_{kl;\,\mathbf{ss};\,\mathbf{ss}} \\ &+ \sum_{\substack{m=1 \\ m\notin\{k,l\}}}^{n_o} \big(\sigma_{km;\,\mathbf{t_2t_2};\,\mathbf{ss}} - \sigma_{km;\,\mathbf{ss};\,\mathbf{ss}} \\ &+ \sigma_{lm;\,\mathbf{t_3t_3};\,\mathbf{ss}} - \sigma_{lm;\,\mathbf{ss};\,\mathbf{ss}}\big). \end{aligned} \tag{36}$$

Note that $\omega_{kl} = \omega_{lk}$.

### 3. *Efficient computation of the constituents*

We may now begin to discuss ways to efficiently compute the above quantities. It is apparent that all of these reduce to $\eta$, $\rho$, and $\sigma$ terms, and these involve the core-Hamiltonian matrix $G$, computed with respect to the $\{\phi_i\}$ set, and the two-electron integrals $\mathbb{M}_{pqrs} = (\phi_p\phi_q|\phi_r\phi_s)$. $G$ has only $n_b^2$ elements, and we store these in memory. Also, we do not need to compute the full $\mathbb{M}$ array. For each $\sigma$ term, the indices for the relevant $\mathbb{M}$ elements come from only two pairs, and for each $\rho$ term the associated indices come from only one pair. The two-pair integrals are of the Coulomb type, $\mathbb{M}_{k'k''l'l''}$, or exchange type, $\mathbb{M}_{k'l'l''k''}$, where $k', k'' \in \{k, \hat{k}\}$, etc. These integrals are only order $n_o^2$ in number, and we store these in memory. Note that this set contains only one integral from each grouping of equivalent integrals; for example $\mathbb{M}_{k\hat{k}ll}$ and $\mathbb{M}_{\hat{k}kll}$ are equal, therefore only one copy is stored. We define $\mathbb{I}_f = \{(p,q,r,s) \mid \mathbb{M}_{pqrs} \text{ required}\}$. The details of the associated integral transformation are discussed later on in this paper.

In our CCVB code, there is a function that takes in the arguments $k$, $l$, $\mathbf{wx}$, and $\mathbf{yz}$, and outputs $\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}}$. The code for this strongly resembles eq. (24), but there are a few adjustments included to ensure efficiency. First, we split $\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}}$ into Coulomb and exchange contributions:

$$\sigma^c_{kl;\,\mathbf{wx};\,\mathbf{yz}} = \sum_{p,r\in S_k} \sum_{q,s\in S_l} (pr|qs)\, P_{k;\,\mathbf{wx};\,pr} P_{l;\,\mathbf{yz};\,qs} \tag{37}$$

and

$$\sigma^{e}_{kl;\,\mathbf{wx};\,\mathbf{yz}} = -\sum_{p,r\in S_k}\sum_{q,s\in S_l}(ps|qr)\,P_{k;\,\mathbf{wx};\,pr}P_{l;\,\mathbf{yz};\,qs}. \tag{38}$$

For the exchange-integral part, we can get a four-fold reduction in the corresponding summation by noting that the 4-by-4 matrix $P_{k;\,\mathbf{wx}}$ is composed of four 2-by-2 blocks according to spin:

$$P_{k;\,\mathbf{wx}} = \begin{bmatrix} P^{\alpha\alpha}_{k;\,\mathbf{wx}} & P^{\alpha\beta}_{k;\,\mathbf{wx}} \\ P^{\beta\alpha}_{k;\,\mathbf{wx}} & P^{\beta\beta}_{k;\,\mathbf{wx}} \end{bmatrix}. \tag{39}$$

With this, we obtain

$$\sigma^{e}_{kl;\,\mathbf{wx};\,\mathbf{yz}} = -\sum_{\tau_1,\tau_2\in\{\alpha,\beta\}}\sum_{p\in S_k^{\tau_1}}\sum_{r\in S_k^{\tau_2}}\sum_{q\in S_l^{\tau_2}}\sum_{s\in S_l^{\tau_1}}(ps|qr)\,P^{\tau_1\tau_2}_{k;\,\mathbf{wx};\,pr}P^{\tau_2\tau_1}_{l;\,\mathbf{yz};\,qs}, \tag{40}$$

where $S^{\alpha}_{k} = \{k_\alpha, \hat{k}_\alpha\}$, etc.

The Coulomb-integral summation can be simplified by using

$$\mathbb{M}_{k'k''l'l''} = \mathbb{M}_{k''k'l'l''} = \mathbb{M}_{k''k'l''l'} = \mathbb{M}_{k'k''l''l'}. \tag{41}$$

Also, only the same-spin blocks of the density matrices will contribute to this Coulomb sum. Accordingly, we also use a Coulomb pair density matrix:

$$\begin{aligned} Q_{k;\,\mathbf{wx};\,kk} &= P_{k;\,\mathbf{wx};\,k_\alpha k_\alpha} + P_{k;\,\mathbf{wx};\,k_\beta k_\beta} \\ Q_{k;\,\mathbf{wx};\,k\hat{k}} &= P_{k;\,\mathbf{wx};\,k_\alpha \hat{k}_\alpha} + P_{k;\,\mathbf{wx};\,\hat{k}_\alpha k_\alpha} + P_{k;\,\mathbf{wx};\,k_\beta \hat{k}_\beta} + P_{k;\,\mathbf{wx};\,\hat{k}_\beta k_\beta} \\ Q_{k;\,\mathbf{wx};\,\hat{k}\hat{k}} &= P_{k;\,\mathbf{wx};\,\hat{k}_\alpha \hat{k}_\alpha} + P_{k;\,\mathbf{wx};\,\hat{k}_\beta \hat{k}_\beta}, \end{aligned} \tag{42}$$

with which we obtain

$$\sigma^{c}_{kl;\,\mathbf{wx};\,\mathbf{yz}} = \sum_{\chi_1\in M_k}\sum_{\chi_2\in M_l}(\chi_1|\chi_2)\,Q_{k;\,\mathbf{wx};\,\chi_1}Q_{l;\,\mathbf{yz};\,\chi_2}, \tag{43}$$

where $M_k = \{kk, k\hat{k}, \hat{k}\hat{k}\}$, etc.

Next, the two-electron integral array discussed above is stored such that all the integrals for the two relevant pairs occur sequentially in memory. Last, each $P$ density matrix and some $Q$ density matrices have several elements that are always 0. For example, the $\alpha\beta$ block of $P_{k;\,\mathbf{ss}}$, which contains e.g. $P_{k;\,\mathbf{ss};\,k_\alpha k_\beta}$, is 0 regardless of the value of $\theta_k$. None of these elements are stored, and the loops corresponding to the summations in eqns. (40) and (43) are reduced accordingly.

We use similar ideas to compute $\eta_{k;\,\mathbf{wx}}$ and $\rho_{k;\,\mathbf{w};\,\mathbf{x}}$. The three functions discussed in this subsection may be used to compute the energy (given a set of $t$ amplitudes) and the $\mu$ and $\kappa$ matrices in quadratic time. In principal, it is possible to compute the $\omega$ matrix in quadratic time, too. This may be accomplished if terms like $\sum_{m=1}^{n_o} \sigma_{km;\,\mathbf{ss};\,\mathbf{ss}}$ are computed first. However, because the $\omega$-matrix computation is not rate-limiting, we have used the naive cubic-scaling approach in our implementation.

## B. Varying the orbitals to optimize the CCVB energy

### *1. Establishing a Lagrangian*

We can now efficiently compute the terms that enter into the energy and amplitude equations, given a set of orbital parameters. We adopt the usual EST premise that the best values of these parameters are the ones that give the lowest value for $E$. In the more familiar case of a variational method, i.e. one where $E$ is a symmetric expectation value, we could proceed by optimizing $E$ directly. But this would not work in the present case, because we must ensure that the amplitude equations are solved for any set of orbitals being considered. This may be accomplished by substituting $E$ with a Lagrangian

$$\Lambda = E + \sum_{k<l} \lambda_{kl} \Omega_{kl}, \tag{44}$$

where the $\lambda_{kl}$ are the Lagrange multipliers. Following standard Lagrange-multiplier discourse, we see that optimizing $\Lambda$ is the same as optimizing $E$ subject to the constraints $\Omega_{kl} = 0$.

As it appears now, $\Lambda$ is a function of the $\lambda_{kl}$, $t_{kl}$, $C_{pq}$, and $\theta_k$ parameters. Of course, as we vary the $C_{pq}$, we must only consider values that correspond to orthonormal orbitals. To incorporate this constraint, we assume some given set of orthonormal orbitals defined by a matrix $D$, and then consider a rotation of these via an orthogonal transformation

$$C = De^{\Delta}, \tag{45}$$

where $\Delta$ is antisymmetric. We now have a set of effectively independent variables, i.e. the $\lambda_{kl}$, $t_{kl}$, $\Delta_{pq}$, and $\theta_k$ parameters, and their optimal values can be determined by searching for stationary points of $\Lambda$ iteratively. In principle, this can be done by treating all variables

in the same manner, but, again, this does not connect well with the standard procedures used to optimize variational methods, in which we would be looking for stationary points of $E$ with respect to the orbital variables only. To deal with this, we directly solve

$$\begin{aligned} \frac{\partial \Lambda}{\partial \lambda_{kl}} &= 0, \\ \frac{\partial \Lambda}{\partial t_{kl}} &= 0 \end{aligned} \tag{46}$$

for each set of orbitals considered. We proceed to vary the orbitals until stationarity with respect to their variables is achieved; this can be done with standard optimization procedures. The details of the eqn. (46) computation are as follows.

*2. Solving the amplitude equations*

We start by writing

$$\Omega_{kl} = \Omega_{kl;a}\, t_{kl}^2 + \Omega_{kl;b}\, t_{kl} + \Omega_{kl;c}, \tag{47}$$

where

$$\begin{aligned} \Omega_{kl;a} &= -\mu_{kl} \\ \Omega_{kl;b} &= \omega_{kl} - \sum_{m \notin \{k,l\}} (t_{km}\mu_{km} + t_{lm}\mu_{lm}) \\ \Omega_{kl;c} &= \mu_{kl} + \sum_{m \notin \{k,l\}} (t_{km}\kappa_{lm} + t_{lm}\kappa_{km}). \end{aligned} \tag{48}$$

We set up an iteration procedure as follows. If an appropriate set of previously computed amplitudes is available, we use these for an initial guess. These may come from a previous $\Lambda$ iteration, or from a separate CCVB calculation from which we are restarting. If these are not available, we begin by solving $\mu_{kl}(1 - t_{kl}^2) + t_{kl}\omega_{kl} = 0$ for each $t_{kl}$. Then, at each $t$-iteration step, we loop through all pairs of active $k$ and $l$, with $k < l$. For a given value of $k$ and $l$, we compute $\Omega_{kl;a}$, $\Omega_{kl;b}$, $\Omega_{kl;c}$ using the current set of $t$ amplitudes. We then solve the quadratic equation, eq. (47), for $t_{kl}$. We select the root that gives the lowest energy contribution, and move to the next value of $(k, l)$. The iterations are continued until convergence is obtained. Each iteration requires cubic time. Also, throughout the procedure, we store only one set of amplitudes. Thus, this process is much like the Gauss-Seidel method for systems of linear equations.

We then work on the partial derivatives of $\Lambda$ with respect to the $t_{kl}$. We use

$$\frac{\partial \Lambda}{\partial t_{kl}} = \mu_{kl} + \lambda_{kl}\,(2\,\Omega_{kl;a}\,t_{kl} + \Omega_{kl;b}) + \sum_{m \notin \{k,l\}} \big[\lambda_{km}\,(-\mu_{kl}\,t_{km} + \kappa_{lm}) + \lambda_{lm}\,(-\mu_{kl}\,t_{lm} + \kappa_{km})\big]. \tag{49}$$

These form a system of equations linear in the $\lambda_{kl}$, which we solve with Gauss-Seidel iterations, each one requiring cubic effort. We can now turn to the orbital variables.

### *3. Orbital-optimizing $\Lambda$: Geometric Direct Minimization*

Since we can now roughly view our energy minimzation problem as an optimization with respect to orbital variables only, we can make use of more familiar EST energy minimization techniques. In CCVB, we use Geometric Direct Minimization (GDM),[78] which is a Quasi-Newton method that is adapted to situations in which the variables form an orthogonal matrix. GDM is used in other GVB methods implemented in the Q-Chem program, and our usage is very similar. We follow the strategy used in ref. 79, where, as far as GDM is concerned, no distinction is made between the $\Delta$ and $\theta$ variables, i.e. they are passed into the GDM procedure as a uniform set of $\frac{1}{2}n_b(n_b-1)+n_p$ variables.

GDM works by taking certain input parameters and then taking a "step", i.e. returning a new set of variables that decrease the energy. The required input parameters are 1) a set of "current" values for the orbital variables, 2) the associated energy (for us, the value of $\Lambda$ with $\lambda_{kl}$ and $t_{kl}$ solved), 3) the energy gradient with respect to the orbital variables, and 4) the diagonal elements of the Hessian matrix for these variables.

There is freedom in choosing input 1), because we can make different choices of $D$ in eq. (45). We choose $D$ to correspond to the orbitals for the current optimization interation, i.e. the orbitals that give the lowest energy up to that point. In other words, as we successfully take a step and rotate the previous iteration's orbitals, we adjust $D$ simultaneously. With this, the current $\Delta$ elements are all 0, and, as we utilize below, the expressions for the associated partial derivatives are simplified greatly. Thus we can loosely view the $\Delta_{pq}$ optimization procedure as a continually reseting call to GDM, which continues until the null values of $\Delta$ have an associated null gradient. The $\theta$ variables, however, are not reset.

Of course, before the optimization commences, we must find an initial choice for $D$. Preferably, we use converged orbitals from a calculation at a nearby molecular geometry,

as would be appropriate if a potential energy surface (PES) were being computed. If this is not available, we generally try one of two options. For the first one, we run an RHF calculation, then localize the occupied orbitals, then find virtual orbitals that correspond well with the localized occupied orbitals in terms of exchange-integral overlaps, thus forming a set of correlating electron pairs, which may be converted to PP form. This procedure,[80] like its relatives,[45,46] is based on long experience with PP-type models: stripping the full treatment of electron correlation down to a small number of basic constituents and orbital-optimizing the resulting model tends to produce localized orbitals. Moreover, a certain degree of orbital localization is necessary here for size consistency and size extensivity. This RHF-based option constitutes the standard initial guess for other PP-type methods in Q-Chem.

For the second option, we run a UHF calculation, then rotate the occupied orbitals to form a set of strongly orthogonal pairs, as mentioned above in section II A 1 when we contrasted the VB and natural-orbital forms for PP. This can be useful for bond-breaking situations in which the pairs become significantly polarized.

We now discuss the computation of the remaining GDM inputs.

#### 4. *Computing* $\frac{\partial \Lambda}{\partial \Delta_{pq}}$

We want partial derivatives of $\Lambda$ with respect to the elements of $\Delta$, but these will be obtained by first computing the partial derivatives with respect to the $C_{pq}$ and then using the chain rule. To obtain the derivatives with respect to the $C_{pq}$, it is convenient to separate the terms containing these variables from the rest

$$\Lambda = \sum_{p,q=1}^{n_o+n_p} \Gamma_{pq} h_{pq} + \sum_{(p,q,r,s)\in\mathbb{I}_f} \Gamma_{pqrs} \mathbb{M}_{pqrs} \tag{50}$$

Here the $\Gamma$ terms depend explicitly only on the $\lambda_{kl}$, $t_{kl}$, and $\theta_k$ parameters, while the $h_{pq}$ and $\mathbb{M}_{pqrs}$ terms are dependent only on the $C_{pq}$.

In essence, the $\Gamma$ elements are those of a density matrix. Using eqn. (44) and the definitions for the terms contained therein, it would be possible to derive explicit expressions for $\Gamma$ elements. However, this would be tedious and not readily applicable to making variations in the CCVB equations and methodology. Instead, since we are computing $E$ and $\Lambda$ in terms of $\eta$, $\rho$, and $\sigma$ terms, we can compute the density matrix on the fly in the same way. Writing

out $\Lambda$ in terms of these three basic quantities, we obtain

$$\begin{aligned}\Lambda = &\sum_{i=1}^{n_o} \left(\eta_{i;\,\mathbf{ss}} + \rho_{i;\,\mathbf{s};\,\mathbf{s}}\right) + \sum_{\substack{i,j=1\\ i<j}}^{n_o} \sigma_{ij;\,\mathbf{ss};\,\mathbf{ss}} \\ &+ \sum_{i<j} \Bigg[ -\sqrt{3}\left(t_{ij} + \lambda_{ij}(1-t_{ij}^2)\right)\sigma_{ij;\,\mathbf{st_2};\,\mathbf{st_3}} \\ &\quad + \lambda_{ij}t_{ij}\Big(\eta_{i;\,\mathbf{t_2t_2}} + \eta_{j;\,\mathbf{t_2t_2}} - \eta_{i;\,\mathbf{ss}} - \eta_{j;\,\mathbf{ss}} + \rho_{i;\,\mathbf{t_2};\,\mathbf{t_2}} + \rho_{j;\,\mathbf{t_2};\,\mathbf{t_2}} - \rho_{i;\,\mathbf{s};\,\mathbf{s}} - \rho_{j;\,\mathbf{s};\,\mathbf{s}} \\ &\quad + \sigma_{ij;\,\mathbf{t_2t_2};\,\mathbf{t_3t_3}} - \sigma_{ij;\,\mathbf{ss};\,\mathbf{ss}} - \sigma_{ij;\,\mathbf{t_2t_1};\,\mathbf{t_3t_1}}\Big) \\ &\quad + \lambda_{ij}t_{ij}\sum_{\substack{m=1\\ m\notin\{i,j\}}}^{n_o} \left(\sigma_{im;\,\mathbf{t_2t_2};\,\mathbf{ss}} - \sigma_{im;\,\mathbf{ss};\,\mathbf{ss}} + \sigma_{mj;\,\mathbf{ss};\,\mathbf{t_2t_2}} - \sigma_{mj;\,\mathbf{ss};\,\mathbf{ss}}\right) \\ &\quad + \sum_{m\notin\{i,j\}} \Big(\lambda_{ij}t_{im}\sigma_{jm;\,\mathbf{t_3s};\,\mathbf{st_3}} + \sqrt{3}\lambda_{ij}t_{im}t_{ij}\sigma_{im;\,\mathbf{st_2};\,\mathbf{st_3}} \\ &\quad + \lambda_{ij}t_{jm}\sigma_{im;\,\mathbf{t_3s};\,\mathbf{st_3}} + \sqrt{3}\lambda_{ij}t_{jm}t_{ij}\sigma_{jm;\,\mathbf{st_2};\,\mathbf{st_3}}\Big)\Bigg], \end{aligned} \tag{51}$$

where all summations are over active indices unless indicated otherwise. We see that the $C_{pq}$ dependence lies entirely within the the $\eta$, $\rho$, and $\sigma$ terms. Suppose one of the resulting terms is $c\,\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}}$. Using the expansion in eqn. (24), we have a sum over two-electron integrals times coefficients. For each term in the summation, we simply add the coefficient to the $\Gamma$ element that corresponds to the integral. Once the summation is complete, we have compiled and stored the contribution of $c\,\sigma_{kl;\,\mathbf{wx};\,\mathbf{yz}}$ to the density matrix $\Gamma$. We can then move on to other $\sigma$ terms, and analagous procedures may be applied to the $\eta$ and $\rho$ terms. For these purposes, we have functions in our code that are highly similar to the $\eta$, $\rho$, and $\sigma$ functions. Computation of $\Gamma$ thus scales cubicly, the same as the computation of $\Lambda$ does. This strategy is readily usable for any CCVB model once we have established its amplitude equations.

We then compute the gradient of $\Lambda$ with respect to the $C_{pq}$. The form of eqn. (50) is quite general, so, apart from an accomodation for the reduced number of used $\mathbb{M}_{pqrs}$ elements, this gradient computation is similar to that done in other correlated methods that use orbital optimization.[81–87] It will therefore not be described in great detail here. We do note that differentiation of two-electron integrals with respect to an element of $C$ results in expressions involving integrals of the form $\mathbb{N}_{pqrs} = (\zeta_p\phi_q|\phi_r\phi_s)$, where $\zeta_p$ is an atomic orbital. As with the $\mathbb{M}$ array, we don't need the full set of integrals here, only the ones of the form $\mathbb{N}_{pqk'k''}$ and

$\mathbb{N}_{pk'k''q}$, where $k'$ and $k''$ come from the same pair. Here, the $q$ index is fully unconstrained and therefore includes inactive unoccupied indices. Integrals containing such indices aren't needed for the gradient, but are needed to compute the diagonal elements of the Hessian matrix. The number of retained $\mathbb{N}$ elements scales cubicly, and so these integrals are stored on disk. Note that these integrals are also transformed to produce the relevant $\mathbb{M}$ elements. Analogous to the two-electron case, we compute the full set of half transformed one-electron integrals, with the ones containing no inactive unoccupied indices being needed to form the gradient.

Our gradient computation is based on looping through all pertinent half-transformed one-electron integrals and stored $\mathbb{N}$ elements, and for each one adding its contributions to separate tallies for the various derivatives that can produce that integral. This requires cubic time. We can use the results to compute

$$\left[\frac{\partial \Lambda}{\partial \Delta_{pq}}\right] = \sum_{rs} \left[\frac{\partial \Lambda}{\partial C_{rs}}\right] \left[\frac{\partial C_{rs}}{\partial \Delta_{pq}}\right] = \sum_{r} \left( \left[\frac{\partial \Lambda}{\partial C_{rq}}\right] D_{rp} - \left[\frac{\partial \Lambda}{\partial C_{rp}}\right] D_{rq} \right), \tag{52}$$

where here, and in what follows,

$$[X] = X\big|_{C=D}, \text{ or } [X] = X\big|_{\Delta=0}, \tag{53}$$

depending on the context. This involves another cubic step.

### 5. *Computing* $\frac{\partial^2 \Lambda}{\partial \Delta_{pq}^2}$

For the Hessian diagonal, we begin with

$$\frac{\partial^2 \Lambda}{\partial \Delta_{pq}^2} = \mathbb{G}_1(\Lambda)_{pq} + \mathbb{G}_2(\Lambda)_{pq}, \tag{54}$$

where

$$\mathbb{G}_1(X)_{pq} = \sum_{rs} \left( \sum_{tu} \frac{\partial^2 X}{\partial C_{tu} \partial C_{rs}} \frac{\partial C_{tu}}{\partial \Delta_{pq}} \right) \frac{\partial C_{rs}}{\partial \Delta_{pq}}, \tag{55}$$

$$\mathbb{G}_2(X)_{pq} = \sum_{rs} \frac{\partial X}{\partial C_{rs}} \frac{\partial^2 C_{rs}}{\partial \Delta_{pq}^2}, \tag{56}$$

and $X$ is a variable quantity. Specializing to the $\Delta = 0$ case, we have

$$[\mathbb{G}_2(\Lambda)_{pq}] = -\sum_{r} \left( \left[\frac{\partial \Lambda}{\partial C_{rp}}\right] D_{rp} + \left[\frac{\partial \Lambda}{\partial C_{rq}}\right] D_{rq} \right), \tag{57}$$

and it is evident that this may be computed in passing while the gradient is being computed. This scales cubicly. For the other term, we have

$$[\mathbb{G}_1(X)_{pq}] = \mathbb{D}_1(X)_{pq} + \mathbb{D}_1(X)_{qp} - 2\,\mathbb{D}_2(X)_{pq}, \tag{58}$$

where

$$\mathbb{D}_1(X)_{pq} = \sum_{rt} \left[\frac{\partial^2 X}{\partial C_{tp} \partial C_{rp}}\right] D_{tq} D_{rq} \tag{59}$$

and

$$\mathbb{D}_2(X)_{pq} = \sum_{rt} \left[\frac{\partial^2 X}{\partial C_{tq} \partial C_{rp}}\right] D_{tp} D_{rq}. \tag{60}$$

When the partial derivatives in these expressions are applied to integrals, they serve to replace the molecular-orbital components with atomic-orbital components according to the given $\Delta$ indices. The summations, together with the $D_{tp}$ terms, simply serve to transform these atomic-orbital components back into molecular-orbital components. For example, we have

$$\mathbb{D}_2(\mathbb{M}_{lk\hat{k}\hat{l}})_{\hat{k}k} = \mathbb{M}_{l\hat{k}k\hat{l}} \tag{61}$$

and

$$\mathbb{D}_1(\mathbb{M}_{kkl\hat{l}})_{kp} = 2\,\mathbb{M}_{ppl\hat{l}}. \tag{62}$$

The latter can give a non-zero result for any value of $p$, and so we need to enlarge the set of computed $\mathbb{M}$ elements (but without changing $\mathbb{I}_f$ so the above expressions remain valid). Namely, we must include elements of the form $\mathbb{M}_{ppk'k''}$ and $\mathbb{M}_{pk'k''p}$, where $k', k'' \in \{k, \hat{k}\}$ and $p$ is inactive unoccupied. The number of additional integrals scales quadratically.

We have found it sufficient to approximate $[\mathbb{G}_1(\Lambda)_{pq}]$ by $[\mathbb{G}_1(E_0)_{pq}]$, i.e. the analogous term for the PP reference energy. Using eqn. (34), we have

$$\begin{aligned}
[\mathbb{G}_1(E_0)_{pq}] = &\sum_{k=1}^{n_o} \Big( \mathbb{D}_1(\eta_{k;\,\mathbf{ss}})_{pq} + \mathbb{D}_1(\eta_{k;\,\mathbf{ss}})_{qp} - 2\,\mathbb{D}_2(\eta_{k;\,\mathbf{ss}})_{pq} \Big) \\
&+ \sum_{k=1}^{n_o} \Big( \mathbb{D}_1(\rho_{k;\,\mathbf{s};\,\mathbf{s}})_{pq} + \mathbb{D}_1(\rho_{k;\,\mathbf{s};\,\mathbf{s}})_{qp} - 2\,\mathbb{D}_2(\rho_{k;\,\mathbf{s};\,\mathbf{s}})_{pq} \Big) \\
&+ \sum_{\substack{k,l=1 \\ k<l}}^{n_o} \Big( \mathbb{D}_1(\sigma_{kl;\,\mathbf{ss};\,\mathbf{ss}})_{pq} + \mathbb{D}_1(\sigma_{kl;\,\mathbf{ss};\,\mathbf{ss}})_{qp} - 2\,\mathbb{D}_2(\sigma_{kl;\,\mathbf{ss};\,\mathbf{ss}})_{pq} \Big).
\end{aligned} \tag{63}$$

In order for any of the terms on the RHS of this equation to be non-zero, either $p$ or $q$ must belong to one of the pairs associated with the corresponding $\eta$, $\rho$, or $\sigma$ term. For example,

if we take $\mathbb{D}_1(\eta_{k;\mathbf{ss}})_{pq}$, then $p$ must be in $\{k, \hat{k}\}$. Therefore, if both $p$ and $q$ are inactive unoccupied, the corresponding diagonal element of the Hessian is 0. Also, the two-index sum in eqn. (63) reduces to a one-index sum, while each one-index sum is reduced similarly.

In our code, there is a function that computes $\mathbb{D}_1(\sigma_{kl;\mathbf{ss};\mathbf{ss}})_{pq}$ for appropriate index combinations. This function is patterned after the one used to compute the various $\sigma$'s: we loop through integrals as in eqn. (24), and make substitutions as in eqns. (61) and (62). We use similar functions for the various combinations of $\mathbb{D}_1$ or $\mathbb{D}_2$ and the different function-argument types ($\eta$, $\rho$, or $\sigma$).

Our computation of the $\mathbb{G}_1$ constituent of the diagonal Hessian is summarized as follows. We loop through the relevant elements of $\Delta$, and for each one we loop through all the $\eta$, $\rho$, and $\sigma$ terms that make non-zero contributions via $\mathbb{D}_1$ or $\mathbb{D}_2$. This process scales cubicly.

At this point, we have discussed all the types of integrals required for CCVB, and we can now describe how they are computed. The one-electron integrals are computed in the standard way with cubic steps. For the two-electron integrals, note that each required integral involves (at least) two indices belonging to the same pair. We thus begin by computing all integrals with two same-pair indices and two AO indices. The same procedure is used in ref. 85 (see eqn. 11 there). As discussed in that paper, this computation scales quinticly for small molecules and asymptotically cubicly for large molecules. The number of these half-transformed integrals scales cubicly, and so we store them on disk. To obtain the necessary $\mathbb{N}$ elements, we must convert one of the two AO indices to MO, at quartic cost. We can then transform these elements into the required $\mathbb{M}$ elements, which scales cubicly.

### 6. *Computing* $\frac{\partial \Lambda}{\partial \theta_k}$ *and* $\frac{\partial^2 \Lambda}{\partial \theta_k^2}$

Returning to the optimization procedure, we also need to compute the gradient and Hessian diagonal of $\Lambda$ with respect to the $\theta_k$. For this, we again turn to eqn. (51), where we see that the derivatives sought boil down to derivatives of the $\eta$, $\rho$, and $\sigma$ terms. The associated summation involves no more than three indices, and for each $\theta_k$, only a quadratic number of terms can contribute to the pertinent first or second derivative. With this, our computation of the $\theta$ gradient and Hessian diagonal scales cubicly. This can be reduced to quadratic: each $\sigma$ term involves only two indices and the third-sum index can in principle be contracted, but we have not explored this because this part of the CCVB calculation is

not rate limiting.

To obtain derivatives of the $\eta$'s, $\rho$'s, and $\sigma$'s with respect to the $\theta_k$, we only need to evaluate the same derivatives of the $f_{k;\mathbf{w};pq}$ and the $P_{k;\mathbf{wx};pr}$. We denote these by $f_{k;\mathbf{w}';pq}$, $P_{k;\mathbf{wx}';pr}$, $f_{k;\mathbf{w}'';pq}$, and $P_{k;\mathbf{wx}'';pr}$ where one or two primes indicates the first or second derivative, respectively. The differentiated density matrices are shown explicitly in the SI. We then have

$$\frac{\partial \sigma_{kl;\mathbf{wx};\mathbf{yz}}}{\partial \theta_k} = \sigma_{kl;\mathbf{wx}';\mathbf{yz}}, \tag{64}$$

$$\frac{\partial^2 \sigma_{kl;\mathbf{wx};\mathbf{yz}}}{\partial \theta_k \partial \theta_l} = \sigma_{kl;\mathbf{wx}';\mathbf{yz}'}, \tag{65}$$

etc., where the primes here indicate that we are to use the appropriate differentiated density matrices. In other words, to differentiate a $\sigma$ term, we can simply replace a density matrix in eqn. (24) with its differentiated counterpart. The same idea holds for the $\eta$'s and $\rho$'s.

We have completed our description of the CCVB algorithm, and we give a summary of the optimization cycle, along with example timing data to be discussed later, in Table I. It should be noted that the scalings in this table are defined with respect to molecule size. For a fixed molecule, i.e. fixed core and active-space sizes, Steps 2 through 6 do not depend on the basis-set size.

## III. APPLICATIONS

For what follows, all HF, CCVB, and CC calculations were done in Q-Chem,[88] and all Configuration Interaction (CI) calculations were done in GAMESS.[89] All PES plots were made with gnuplot[90] and all orbital plots were made with gOpenMol.[91,92]

### A. Symmetric dissociations of $S_6$ and $S_8$

In this subsection, we will consider the singlet ground states of $S_6$ and $S_8$ in the $D_{3d}$ and $D_{4d}$ point-group symmetries, respectively. Within these geometric contraints, the coordinates are fully determined by three parameters: a bond angle, a dihedral angle, and a nearest-neighbor bond length. We are interested in bond-breaking PES's; for this, we fix the angles at the optimized values given in ref. 93, and vary the bond length until the two molecules dissociate symmetrically into 6 and 8 triplet S atoms, respectively. This geometry setup is illustrated in Fig. 1, and example z-matrices for these PES's are given in the SI.

These PES's entail the breaking of 6 and 8 covalent bonds, and these bonds are coupled: for each sulfur atom, two electrons (one electron from each of the atom's two associated bonds) must reorganize into triplet alignment. These examples are thus very much in line with the bond-breaking SC discussion found in the Introduction.

For simplicity, and to enable comparison with CI, we will employ the minimal treatment necessary for qualitative correctness for these PES's: 6 and 8 active CCVB pairs, i.e. (12,12) and (16,16) active spaces. Here, the CCVB or CI energy will approach 6 or 8 times the S-atom restricted open-shell HF (ROHF) energy in the dissociation limit.

We used the UHF-based initial guess discussed in section II B 3 for the most stretched geometry and then sequentially read-in the orbitals from adjacent geometries for the remaining computations. Calculations were done in the 6-31G* basis.

### *1.* $S_6$

CCVB, UHF, and CI results are shown in Fig. 2. The CI results were obtained using the CCVB orbitals, i.e. we performed full CI (FCI) calculations within the (12,12) active space defined by the CCVB orbitals. Our main observation at this point is that CCVB recovers a great deal of the static correlation for this system, while UHF is clearly missing a significant portion of the correct result.

The UHF computations were not done in the standard way. We used the fact that UHF may be parameterized by a set of strongly orthogonal $\alpha$ and $\beta$ orbitals, as discussed in the Theory section. Since CCVB also uses strongly orthogonal orbitals, we have simply modified the CCVB code to be able perform UHF calculations in this alternative way. Similar to what is done in CCVB, we can set up core and active spaces. For UHF, this means spin polarization is allowed only within the active space, with the remaining electrons receiving an RHF-type treatment. If all pairs are allowed to polarize, the result is equivalent to regular UHF. Here, we only allowed spin polarization in the bonding (12,12) active space, i.e. we are unrestricting in 6 active pairs. While this technically raises the energy, the PES is qualitatively unchanged.

The reason we use this UHF strategy is to enable comparisons between CCVB, post-HF methods, and CI. The key to this is basing all calculations on the (12,12) active space. This is fairly straightforward in post-HF methods; for example, in CC we simply constrain all

amplitude indices to belong to the active space. Hence, we have computed active-space CC singles doubles with perturbative triples (CCSD(T)) energies using the paired-UHF orbitals. The error associated with this calculation is quantified by comparing these energies to that obtained from active-space FCI based on the paired UHF orbitals, CI(UHF). On the other hand, we must use active-space FCI based on the CCVB orbitals, CI(CCVB), to quantify the CCVB error. These subtleties are necessary to ensure fair comparisons. Even so, although the CI(CCVB) energy is generally below that of CI(UHF), they are generally close to each other, as seen in Fig. S-1 in the SI.

The results for the preceding ideas are found in Fig. 3. In this plot, we are showing the region intermediate between equilibrium and dissociation. This is because UHF begins to spin polarize between 2.3 and 2.4 Å, i.e. UHF and RHF are equivalent for shorter bond lengths, so no active-space comparison is practical there. In the intermediate region between 2.6 and 3.2 Å, we see that CCVB is actually outperforming CCSD(T). This exemplifies our assertions made in the Introduction: the unrestricted methods are hampered by spin contaminants in the intermediate, or recoupling, regime before complete dissociation.

#### *2.* $S_8$

Beyond the increase in active-space size, the calculations for $S_8$ are essentially the same as those done for $S_6$, except that here FCI is excluded due to time considerations. The general PES's are shown in Fig. S-2 in the SI, while in Fig. 4, we show the difference between the CCVB and active-space CCSD(T) energies. These results simply provide more support for the observations made above in the $S_6$ case: CCVB again outperforms CCSD(T) in the intermediate regime between 2.6 and 3.2 Å.

### B. Acenes

In this subsection, we consider the acenes, i.e. linearly fused benzene rings. These molecules have been of significant recent interest in EST because SC gradually emerges as $n$, the number of rings, is increased.[71,94–105]

A CCVB calculation reveals the presence of SC in essentially two ways: First, $\theta_k$ values near $\pi/2$ indicate SC within a pair, with the limiting value corresponding to orthogonality

between the two involved VB orbitals. Second, $t_{kl}$ values near $1/\sqrt{3}$ indicate the presence of strong inter-pair couplings associated with 4 (or more) electron SC. Both of these limits are approached in the dissociation of coupled bonds. Hence, the largest angle and amplitude values of each calculation will be among the main results of this subsection.

We have done CCVB calculations for $n = 4$ up to $n = 12$. All C-C and C-H bond lengths were set at 1.4 and 1.08 Å, respectively, all valence electrons were correlated, and, following ref. 96, calculations were done with the STO-3G basis set.

For the above sulfur rings, it was relatively easy to deduce the way CCVB would treat the essential correlations: going around the rings, UHF has a simple spin-alternation pattern, which provided an excellent starting point. On the contrary, extensive conjugation in the expansive $\pi$ spaces makes such conjecturing for the acene systems much less straightforward. In systems like this, where the form of the active orbitals is initially unclear, using multiple initial guesses with CCVB (or any method based on GVB-PP) is advisable. However, in general, different initial-guess procedures can lead to different energy stationary points, or "solutions". The existence of multiple low-lying solutions is a very common finding for PP-based methods and it is related to their lack of orbital invariance. This makes the general problem of finding, or, more precisely, certifying, the CCVB global minimum very hard.

We experimented with both the RHF and UHF-based schemes discussed in section II B 3. In our acenes calculations, the RHF-based guess always produced the best solutions, in that the corresponding energies were always far lower than those of the UHF-based guess. This is not unexpected: UHF orbitals, as rotated into VB form, are usually highly delocalized, even for medium and large molecules. But, as stated in II B 3, optimal GVB-PP (and CCVB) solutions exhibit a certain degree of orbital localization, so the UHF-based guess is poor for larger systems. Hence, we will only report the results for the RHF-based guess. For this, we separately tried two localization schemes: Pipek-Mezey (PM)[106] and Boys.[107] The results are summarized in Tables II and III.

Before we study this data, we note that the relatively large size of 12-acene makes for a good opportunity to assess the efficiency of our CCVB implementation. For this molecule, on a single AMD Opteron 6376 Processor (2.3 GHz), one iteration, i.e. one orbital-optimization cycle, takes about 17 minutes and the total number of iterations for each PM-guess calculation is given in Table II. In Table I, we have included data for the time consumed by each step in the iteration, as a percentage of the total time for one iteration. For this, we

averaged over a 10-iteration interval. We have included data for the 6-31G basis, for which one iteration takes about 86 minutes. In the interest of time, we did not allow the latter calculation to complete, and so we will not examine its results outside the timings context.

Among our first observations in looking at Table I is that our $\theta$-derivatives calculation (and, to lesser extents, steps 3 and 6) is not highly efficient: it requires 37% of the total time for STO-3G. This is not quite an accurate representation of the fundamental complexity of this step, i.e. there remains room for significant optimization. But, we were not motivated to do this because, as mentioned previously, the time required for this step (and the others from 2 through 6) does not depend on the basis-set size, so it will be highly overshadowed by the integral transformation as we move to decent basis sets. This effect is already evident in the 6-31G data, where the $\theta$ derivatives require only 7% of the total time.

Calculation of the $t$ and $\lambda$ elements is exceedingly fast. Indeed, using a very tight convergence criterion of $10^{-10}$ in Euclidean-norm error (not root-mean-square), the STO-3G $t$ and $\lambda$ calculations required only 12 iterations each, while for 6-31G they each required 9 iterations. There are 6,441 amplitudes involved in each of these computations.

Looking at the acenes results more generally, we see that the PM guess appears to produce solutions that do not indicate SC. Generally, each orbital for these solutions is localized to some benzene ring. Localization is typical of GVB-PP orbitals, and it underlies the use of orbital localization procedures for the initial guess. The orbitals for the two pairs giving the largest amplitude in 12-acene are plotted in Figs. S-3 and S-4 in the SI. One of these two pairs corresponds to the maximal $\theta_k$.

For the most part, the Boys guess produced results similar to those of PM, except for $n = 10$ and $n = 12$. In those cases, GDM was able to find a path moving away from orbital localization towards a lower energy solution, indeed lower than the PM one. The corresponding maximal $\theta_k$ and $t_{kl}$ indicate some level of SC. The orbitals corresponding to the largest amplitude for 12-acene are shown in Fig. 5. We show the GVB-PP NO form of these orbitals in Fig. 6. In either form, the orbitals clearly display some delocalization. Interestingly, the NO-form orbitals adhere to $D_{2h}$ point-group symmetry.

To the extent that these features of the Boys-guess solutions for $n = 10$ and $n = 12$ are consistent with those of the global minima, whether we have found the latter or not, it is evident that CCVB is detecting SC and associated delocalization in the acenes, in line with the results of previous studies.

## IV. DISCUSSION AND CONCLUSIONS

In this paper, we have analyzed an efficient implementation of the CCVB method. The computation is dominated by an at-worst quintic-scaling step, and this allows CCVB to be used for relatively large molecular systems.

We applied CCVB to multiple-bond dissociations in $S_6$ and $S_8$. There we found that CCVB gives a SC description superior to that of traditional methods.

We applied CCVB to small and medium sized acenes. For two of these molecules, the best available CCVB results implied SC. It is likely that there exist similar solutions for the other comparably sized acenes, i.e. solutions lower in energy that the ones reported above, with SC implications and orbital delocalization. But, even for the two noted molecules, it is difficult to be sure that we have isolated the global energy minimum. This is a challenge common to all orbital-optimization methods, and especially for the orbital-variant ones. It would be possible to obtain a more conclusive picture by, for example, taking a large acene, doing CCVB in the $\pi$ space with an exhaustive set of initial guesses and, possibly, an annealing procedure or the like. But, because the acenes are not the focal point of this work, this lies well outside the present scope. Instead, we point out that for the two noted molecules, there exist localized-orbital solutions that are fairly close in energy to the lower delocalized ones. The localized solutions are similar to what we might expect in GVB-PP; indeed the observed delocalization would be very unusual in GVB-PP. Therefore, the leading contribution to the CCVB energy, that of the reference, is strongly favoring localization. That the optimization can pull away from this in favor of delocalization suggests a significant driving force: strong interpair coupling. Thus, we think it reasonable to conclude that CCVB is giving a legitimate indication of SC in these systems.

The examples considered in this work support our claim that CCVB represents an authentic piece of the SC picture. The method is not without its limitations, however, and this was evident most in the acenes, systems with multiple important resonance structures. The need for improvements in this situation is a basis for further work in the CCVB realm.

## ACKNOWLEDGMENTS

This work was supported by the Director, Office of Energy Research, Office of Basic En-

ergy Sciences, Chemical Sciences Division of the U.S. Department of Energy under contract no. DE- AC0376SF00098. M.H.-G. is a part-owner of Q-CHEM Inc.

## SUPPORTING INFORMATION

Density matrices, example z-matrices and extra PES plots for the $S_6$ and $S_8$ calculations, and extra orbital plots for 12-acene. This material is available free of charge via the Internet at http://pubs.acs.org.

[1] Cremer, D. *WIREs Comput. Mol. Sci.* **2011**, *1*, 509–530.

[2] Saebø, S. In *Linear-Scaling Techniques in Computational Chemistry and Physics*; Zalesny, R., Papadopoulos, M. G., Mezey, P. G., Leszczynski, J., Eds.; Challenges and Advances in Computational Chemistry and Physics; Springer Netherlands, 2011; Vol. 13; pp 65–82.

[3] Rubensson, E. H.; Rudberg, E.; Salek, P. In *Linear-Scaling Techniques in Computational Chemistry and Physics*; Zalesny, R., Papadopoulos, M. G., Mezey, P. G., Leszczynski, J., Eds.; Challenges and Advances in Computational Chemistry and Physics; Springer Netherlands, 2011; Vol. 13; pp 263–300.

[4] Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.

[5] Bowler, D. R.; Miyazaki, T. *Rep. Prog. Phys.* **2012**, *75*, 036503.

[6] Kussmann, J.; Beer, M.; Ochsenfeld, C. *WIREs Comput. Mol. Sci.* **2013**, *3*, 614–636.

[7] Olsen, J.; Yeager, D. L.; Jørgensen, P. *Adv. Chem. Phys.* **1983**, *54*, 1–176.

[8] Roos, B. In *Methods in Computational Molecular Physics*; Diercksen, G., Wilson, S., Eds.; D.Reidel Publishing, Dordrecht, Netherlands, 1983; pp 161–187.

[9] Shepard, R. *Adv. Chem. Phys.* **1987**, *69*, 63–200.

[10] Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233–266.

[11] Hirao, K., Ed. *Recent Advances in Multireference Methods*; World Scientific, Singapore, 1999.

[12] Sherrill, C. D.; H. F. Schaefer III, *Adv. Quantum Chem.* **1999**, *34*, 143–269.

[13] Roos, B. O. In *Theory and Applications of Computational Chemistry*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; pp 725–764.

[14] Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. *Chem. Rev.* **2012**, *112*, 108–181.

[15] This used to be true even for medium sized molecules, but the outlook is changing, see e.g. ref.s 108–111.

[16] Interesting alternatives to orthogonal geminals are currently being explored, see e.g. ref.s 112–115.

[17] Hurley, A. C.; Lennard-Jones, J.; Pople, J. A. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1953**, *220*, 446–455.

[18] Parks, J. M.; Parr, R. G. *J. Chem. Phys.* **1958**, *28*, 335–345.

[19] Arai, T. *J. Chem. Phys.* **1960**, *33*, 95–98.

[20] Kutzelnigg, W. *J. Chem. Phys.* **1964**, *40*, 3640–3647.

[21] Kapuy, E. *J. Chem. Phys.* **1966**, *44*, 956–962.

[22] Surján, P. R. *Top. Curr. Chem.* **1999**, *203*, 63–88.

[23] Surjan, P. R.; Szabados, A.; Jeszenszki, P.; Zoboki, T. *J. Math. Chem.* **2012**, *50*, 534–551.

[24] Rosta, E.; Surjan, P. R. *J. Chem. Phys.* **2002**, *116*, 878–890.

[25] Rassolov, V. A. *J. Chem. Phys.* **2002**, *117*, 5978–5987.

[26] Rassolov, V. A.; Xu, F.; Garashchuk, S. *J. Chem. Phys.* **2004**, *120*, 10385–10394.

[27] Li, S.; Ma, J.; Jiang, Y. *J. Chem. Phys.* **2003**, *118*, 5736–5745.

[28] Ma, J.; Li, S. H.; Li, W. *J. Comput. Chem.* **2006**, *27*, 39–47.

[29] Tarumi, M.; Kobayashi, M.; Nakai, H. *Int. J. Quantum Chem.* **2013**, *113*, 239–244.

[30] Pernal, K.; Cioslowski, J. *Ann. Phys. (Berlin)* **2004**, *13*, 194–200.

[31] Pernal, K. *Comp. Theor. Chem.* **2013**, *1003*, 127 – 129.

[32] F.B. Bobrowicz and W.A. Goddard, in *Methods of Electronic Structure Theory 3*, edited by H.F. Schaefer (Plenum, New York, 1977), p. 79.

[33] Cullen, J. *Chem. Phys.* **1996**, *202*, 217–229.

[34] Moss, B. J.; Bobrowicz, F. W.; W. A. Goddard III, *J. Chem. Phys.* **1975**, *63*, 4632–4639.

[35] Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard III, W. A.; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875–11882.

[36] Hunt, W. J.; Hay, P. J.; W. A. Goddard III, *J. Chem. Phys.* **1972**, *57*, 738–748.

[37] Hay, P. J.; Hunt, W. J.; W. A. Goddard III, *Chem. Phys. Lett.* **1972**, *13*, 30–35.

[38] Hay, P. J.; Hunt, W. J.; Goddard, W. A. *J. Am. Chem. Soc.* **1972**, *94*, 8293–8301.

[39] Dykstra, C. E. *J. Chem. Phys.* **1980**, *72*, 2928–2935.

[40] Carter, E. A.; W. A. Goddard III, *J. Chem. Phys.* **1988**, *88*, 1752–1763.

[41] Hartke, B.; Carter, E. A. *J. Chem. Phys.* **1992**, *97*, 6569–6578.

[42] Gibson, D. A.; Ionova, I. V.; Carter, E. A. *Chem. Phys. Lett.* **1995**, *240*, 261–267.

[43] G. R . De Maré,; Evleth, E. M.; Poirier, R. A.; Collin, G. J. *Can. J. Chem.* **1994**, *72*, 1230–1237.

[44] Langlois, J.; Muller, R.; Coley, T.; Goddard III, W.; Ringnalda, M.; Won, Y.; Friesner, R. *J. Chem. Phys.* **1990**, *92*, 7488–7497.

[45] Langlois, J.-M.; Yamasaki, T.; Muller, R. P.; Goddard, W. A. I. *J. Phys. Chem.* **1994**, *98*, 13498–13505.

[46] Wang, Y.; Poirier, R. A. *J. Comput. Chem.* **1996**, *17*, 313–325.

[47] Murphy, R. B.; Pollard, W. T.; Friesner, R. A. *J. Chem. Phys.* **1997**, *106*, 5073–5084.

[48] Sejpal, M.; Messmer, R. P. *Chem. Phys.* **2001**, *270*, 237 – 243.

[49] Sodt, A.; Beran, G. J. O.; Jung, Y.; Austin, B.; Head-Gordon, M. *J. Chem. Theory Comput.* **2006**, *2*, 300–305.

[50] Beran, G. J. O.; Head-Gordon, M.; Gwaltney, S. R. *J. Chem. Phys.* **2006**, *124*, 114107.

[51] Anderson, A. G.; Goddard, W. A. *J. Chem. Phys.* **2010**, *132*, 164110.

[52] Fracchia, F.; Filippi, C.; Amovilli, C. *J. Chem. Theory Comput.* **2012**, *8*, 1943–1951.

[53] Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1509–1520.

[54] Hardisson, A.; Harriman, J. E. *J. Chem. Phys.* **1967**, *46*, 3639–3648.

[55] Bunge, C. F. *Phys. Rev.* **1967**, *154*, 70–79.

[56] Kaldor, U. *J. Chem. Phys.* **1968**, *48*, 835–837.

[57] Martino, F.; Ladik, J. *Phys. Rev. A* **1971**, *3*, 862–865.

[58] Mayer, I.; Ladik, J.; Biczó, G. *Int. J. Quantum Chem.* **1973**, *7*, 583–608.

[59] Nakatsuji, H. *J. Chem. Phys.* **1973**, *59*, 2586–2595.

[60] Mestechkin, M. M. *Int. J. Quantum Chem.* **1973**, *7*, 425–457.

[61] Phillips, D. H.; Schug, J. C. *J. Chem. Phys.* **1974**, *61*, 1031–1039.

[62] Yamaguchi, K.; Yoshioka, Y.; Takatsuka, T.; Fueno, T. *Theor. Chim. Acta* **1978**, *48*, 185–206.

[63] Mayer, I. *Adv. Quantum Chem.* **1980**, *12*, 189–262.

[64] Karadakov, P. *Int. J. Quantum Chem.* **1986**, *30*, 239–264.

[65] Schlegel, H. B. *J. Phys. Chem.* **1988**, *92*, 3075–3078.

[66] Byrman, C. P.; Lenthe, J. H.; Verbeek, J. *Theor. Chim. Acta* **1993**, *86*, 129–136.

[67] Scuseria, G. E.; Jiménez-Hoyos, C. a.; Henderson, T. M.; Samanta, K.; Ellis, J. K. *J. Chem. Phys.* **2011**, *135*, 124108.

[68] Jiménez-Hoyos, C. a.; Henderson, T. M.; Tsuchimochi, T.; Scuseria, G. E. *J. Chem. Phys.* **2012**, *136*, 164109.

[69] Samanta, K.; Jiménez-Hoyos, C. A.; Scuseria, G. E. *J. Chem. Theory Comput.* **2012**, *8*, 4944–4949.

[70] Rivero, P.; Jiménez-Hoyos, C. a.; Scuseria, G. E. *J. Phys. Chem. A* **2013**, *117*, 8073–8080.

[71] Rivero, P.; Jiménez-Hoyos, C. A.; Scuseria, G. E. *J. Phys. Chem. B* **2013**, *117*, 12750–12758.

[72] Stein, T.; Jiménez-Hoyos, C. A.; Scuseria, G. E. *J. Phys. Chem. A* **2014**, in press.

[73] Small, D. W.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *130*, 084103.

[74] Small, D. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19285–19297.

[75] Beran, G. J. O.; Head-Gordon, M. *Mol. Phys.* **2006**, *104*, 1191–1206.

[76] Lawler, K. V.; Beran, G. J. O.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 024107.

[77] Amos, A. T.; Hall, G. G. *Proc. R. Soc. Lond. A* **1961**, *263*, 483–493.

[78] Van Voorhis, T.; Head-Gordon, M. *Mol. Phys.* **2002**, *100*, 1713–1721.

[79] Lawler, K. V.; Small, D. W.; Head-Gordon, M. *J. Phys. Chem. A* **2010**, *114*, 2930–2938.

[80] Sano, T. *J. Mol. Struct. (Theochem)* **2000**, *528*, 177 – 191.

[81] Scuseria, G. E.; Schaefer III, H. F. *Chem. Phys. Lett.* **1987**, *142*, 354–358.

[82] Koch, H.; Jensen, H. J. A.; Jørgensen, P.; Helgaker, T.; Scuseria, G. E.; Schaefer, H. F. *J. Chem. Phys.* **1990**, *92*, 4924–4940.

[83] Sherrill, C. D.; Krylov, A. I.; Byrd, E. F. C.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 4171–4181.

[84] Pedersen, T. B.; Koch, H.; Hättig, C. *J. Chem. Phys.* **1999**, *110*, 8318–8327.

[85] Van Voorhis, T.; Head-Gordon, M. *J. Chem. Phys.* **2002**, *117*, 9190–9201.

[86] Bozkaya, U.; Turney, J. M.; Yamaguchi, Y.; Schaefer, H. F.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *135*, 104103.

[87] Boguslawski, K.; Tecmer, P.; Ayers, P.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Towards an efficient description of strongly correlated electrons with mean-field cost. 2014, arXiv:1401.8019. arXiv.org e-Print archive. http://arxiv.org/abs/1401.8019 (accessed Apr 9, 2014).

[88] Shao, Y. et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

[89] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[90] Williams, T.; Kelley, C.; many others, Gnuplot 4.6: an interactive plotting program. `http://www.gnuplot.info`, 2012.

[91] Laaksonen, L. *J. Mol. Graph.* **1992**, *10*, 33–34.

[92] Bergman, D. L.; Laaksonen, L.; Laaksonen, A. *J. Mol. Graph. Model.* **1997**, *15*, 301–306.

[93] Raghavachari, K.; Rohlfing, C. M.; Binkley, J. S. *J. Chem. Phys.* **1990**, *93*, 5862–5874.

[94] Bendikov, M.; Duong, H. M.; Starkey, K.; Houk, K. N.; Carter, E. A.; Wudl, F. *J. Am. Chem. Soc.* **2004**, *126*, 7416–7417.

[95] Kadantsev, E. S.; Stott, M. J.; Rubio, A. *J. Chem. Phys.* **2006**, *124*, 134901.

[96] Hachmann, J.; Dorando, J. J.; Avilés, M.; Chan, G. K.-L. *J. Chem. Phys.* **2007**, *127*, 134309.

[97] Gidofalvi, G.; Mazziotti, D. A. *J. Chem. Phys.* **2008**, *129*, 134108.

[98] Jiang, D.-e.; Dai, S. *J. Phys. Chem. A* **2008**, *112*, 332–335.

[99] Casanova, D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9779–9790.

[100] Qu, Z.; Zhang, D.; Liu, C.; Jiang, Y. *J. Phys. Chem. A* **2009**, *113*, 7909–7914.

[101] Bettinger, H. F. *Pure Appl. Chem.* **2010**, *82*, 905 – 915.

[102] Ess, D. H.; Johnson, E. R.; Hu, X.; Yang, W. *J. Phys. Chem. A* **2011**, *115*, 76–83.

[103] Pelzer, K.; Greenman, L.; Gidofalvi, G.; Mazziotti, D. A. *J. Phys. Chem. A* **2011**, *115*, 5632–5640.

[104] Mizukami, W.; Kurashige, Y.; Yanai, T. *J. Chem. Theory Comput.* **2013**, *9*, 401–407.

[105] Plasser, F.; Pašalić, H.; Gerzabek, M. H.; Libisch, F.; Reiter, R.; Burgdörfer, J.; Müller, T.; Shepard, R.; Lischka, H. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 2581–2584.

[106] Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[107] Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[108] Chan, G. K.-L.; Sharma, S. *Annu. Rev. Phys. Chem.* **2011**, *62*, 465–481.

[109] Ghosh, D.; Hachmann, J.; Yanai, T.; Chan, G. K.-L. *J. Chem. Phys.* **2008**, *128*, 144117.

[110] Kurashige, Y.; Yanai, T. *J. Chem. Phys.* **2011**, *135*, 094104.

[111] Liu, F.; Kurashige, Y.; Yanai, T.; Morokuma, K. *J. Chem. Theory Comput.* **2013**, *9*, 4462–4469.

[112] Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. *J. Chem. Theory Comput.* **2013**, *9*, 1394–1401.

[113] Johnson, P. A.; Ayers, P. W.; Limacher, P. A.; Baerdemacker, S. D.; Neck, D. V.; Bultinck, P. *Comp. Theor. Chem.* **2013**, *1003*, 101 – 113.

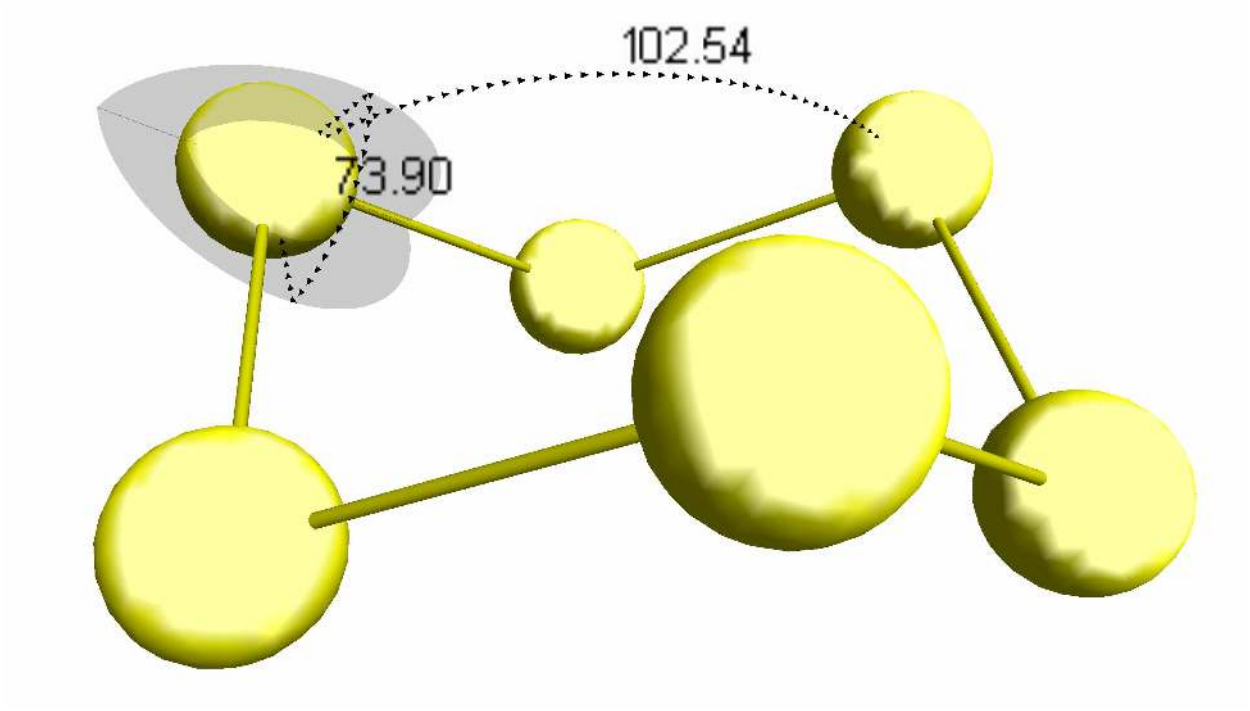
[114] Neuscamman, E. *Phys. Rev. Lett.* **2012**, *109*, 203001.

[115] Neuscamman, E. *J. Chem. Phys.* **2013**, *139*, 181101.

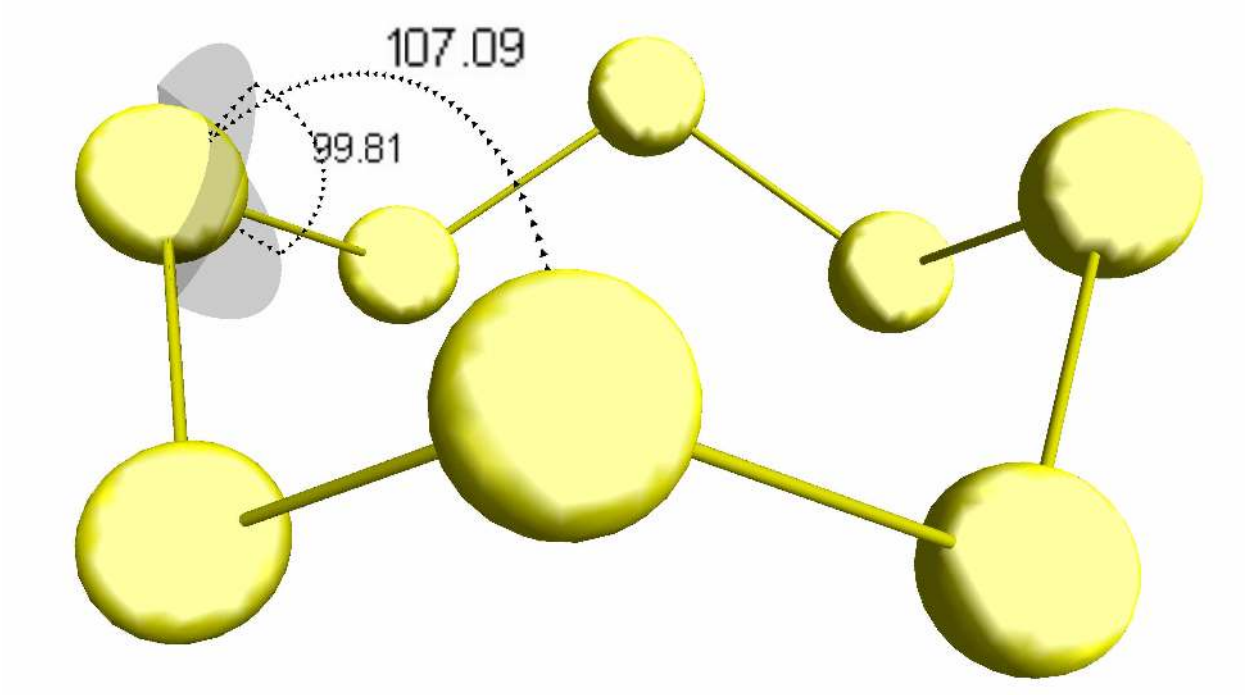
[116] Bode, B. M.; Gordon, M. S. *J. Mol. Graph. Model.* **1998**, *16*, 133–138.

## LIST OF FIGURES

## FIGURES




(a) $S_6$.




(b) $S_8$.

FIG. 1

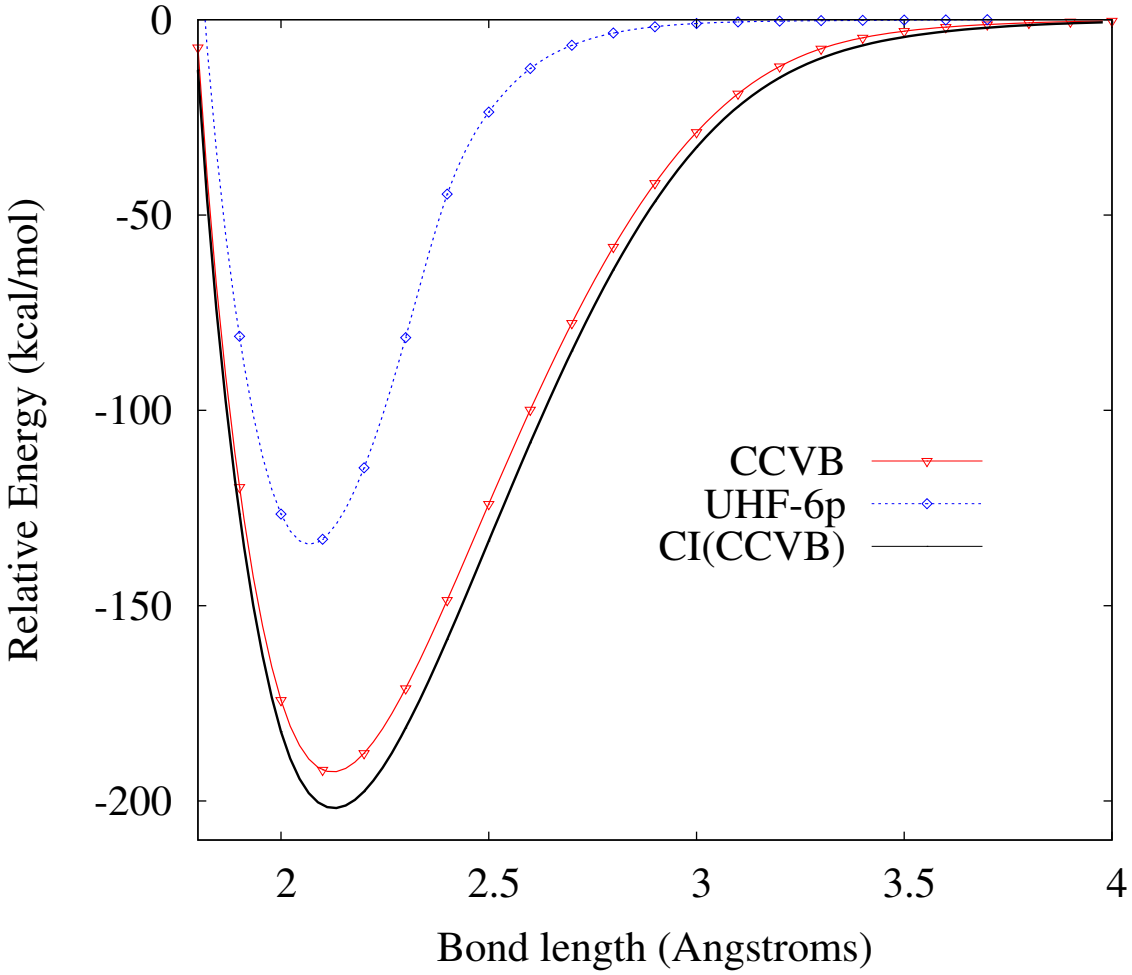

0
-50
-100
-150
-200
Relative Energy (kcal/mol)
2
2.5
3
3.5
4
Bond length (Angstroms)
CCVB
UHF-6p
CI(CCVB)

FIG. 2

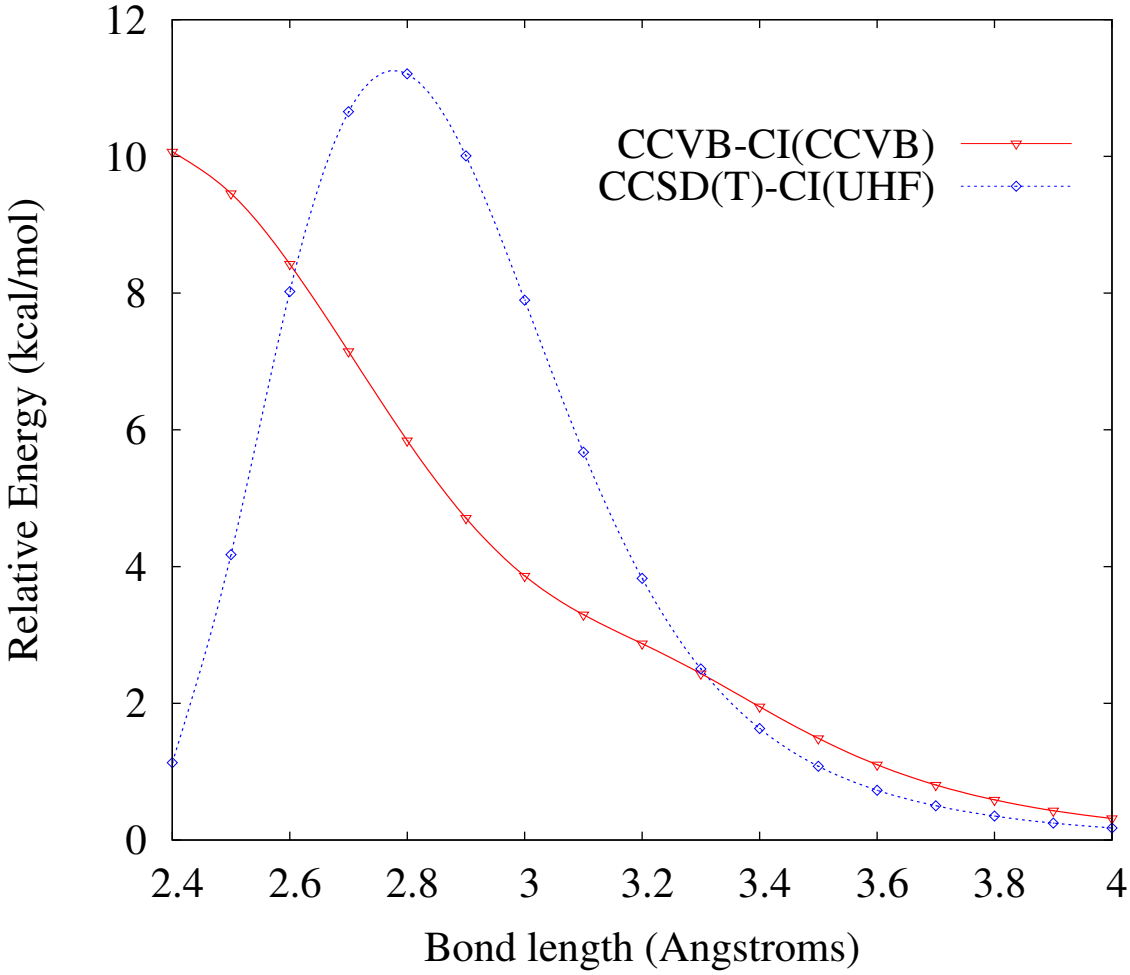

Relative Energy (kcal/mol)
Bond length (Angstroms)
CCVB-CI(CCVB)
CCSD(T)-CI(UHF)
0
2
4
6
8
10
12
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4

FIG. 3

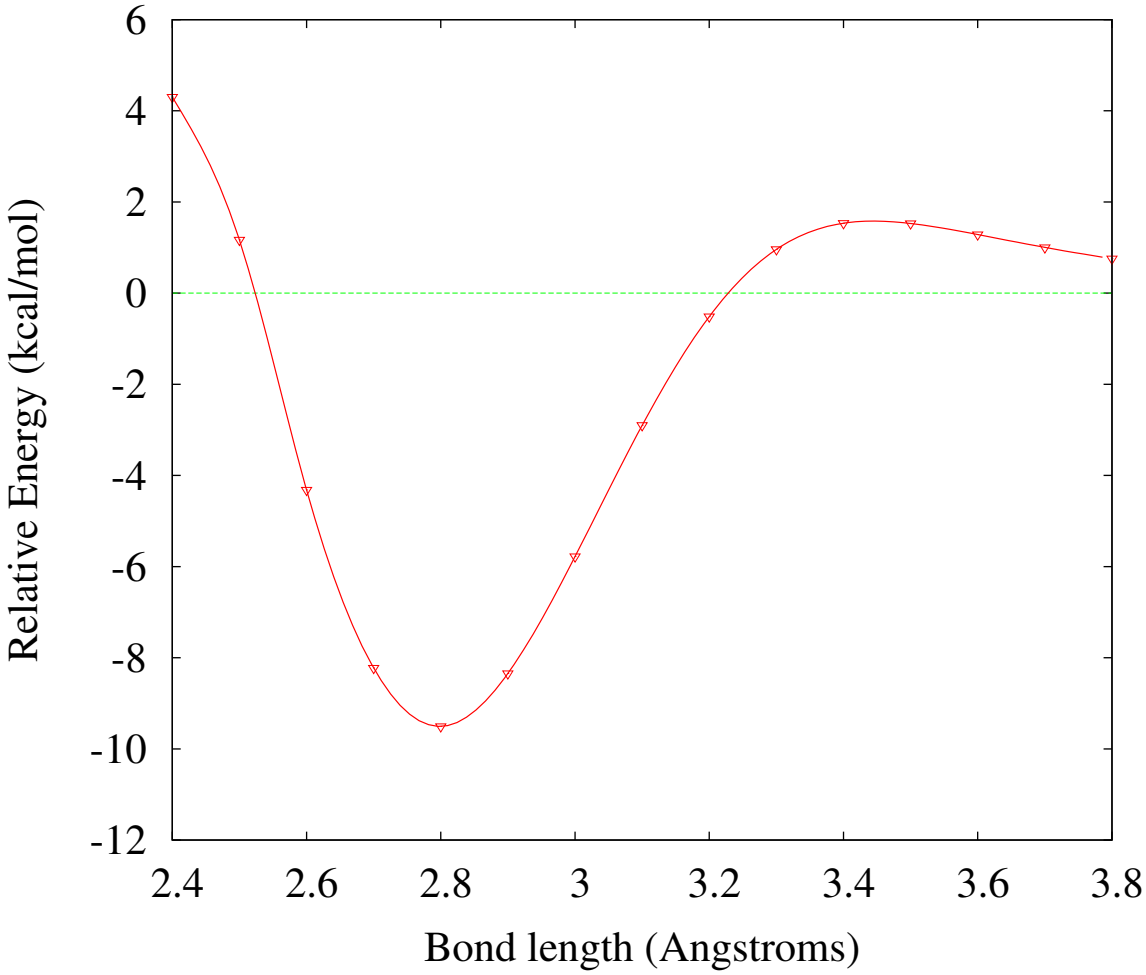
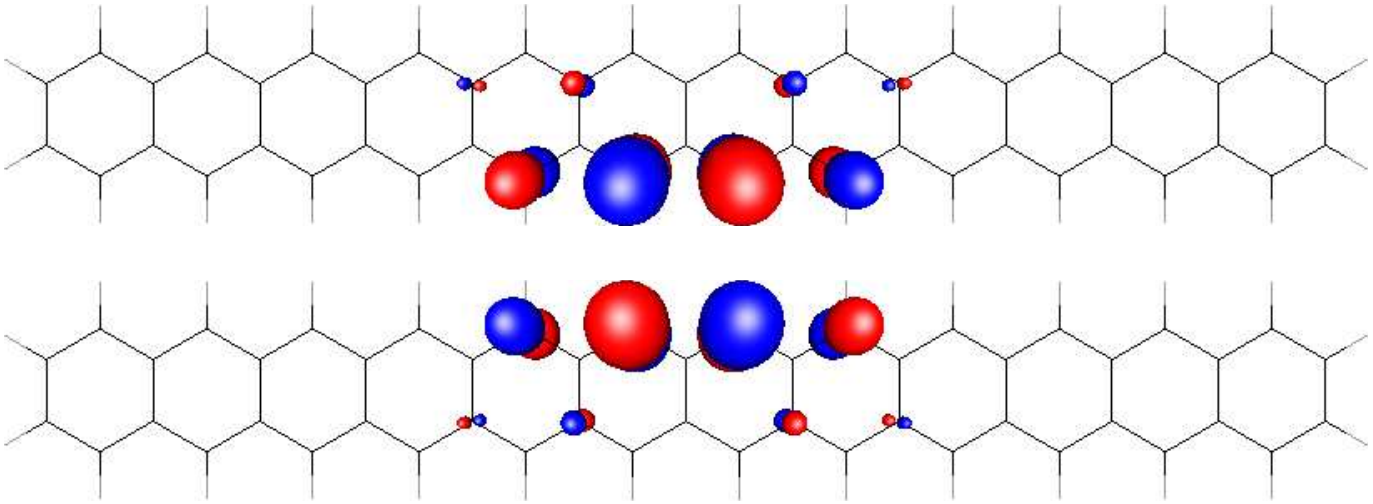
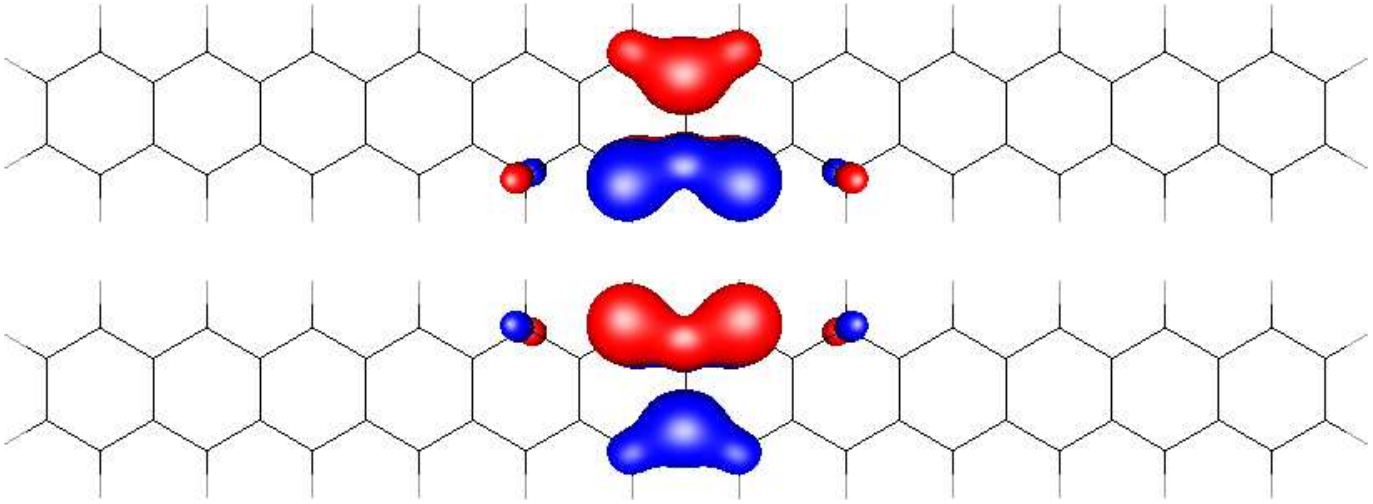

Relative Energy (kcal/mol)
6
4
2
0
-2
-4
-6
-8
-10
-12
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
Bond length (Angstroms)

FIG. 4

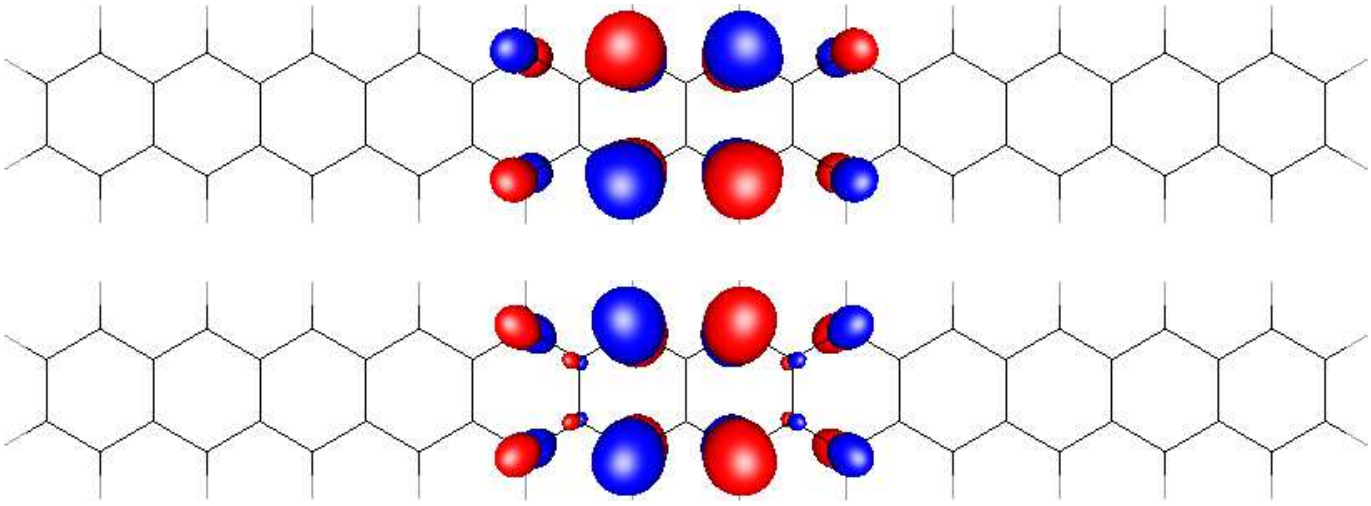


(a) Pair 1 orbitals. $\theta_k = 1.40556$, the largest value observed.

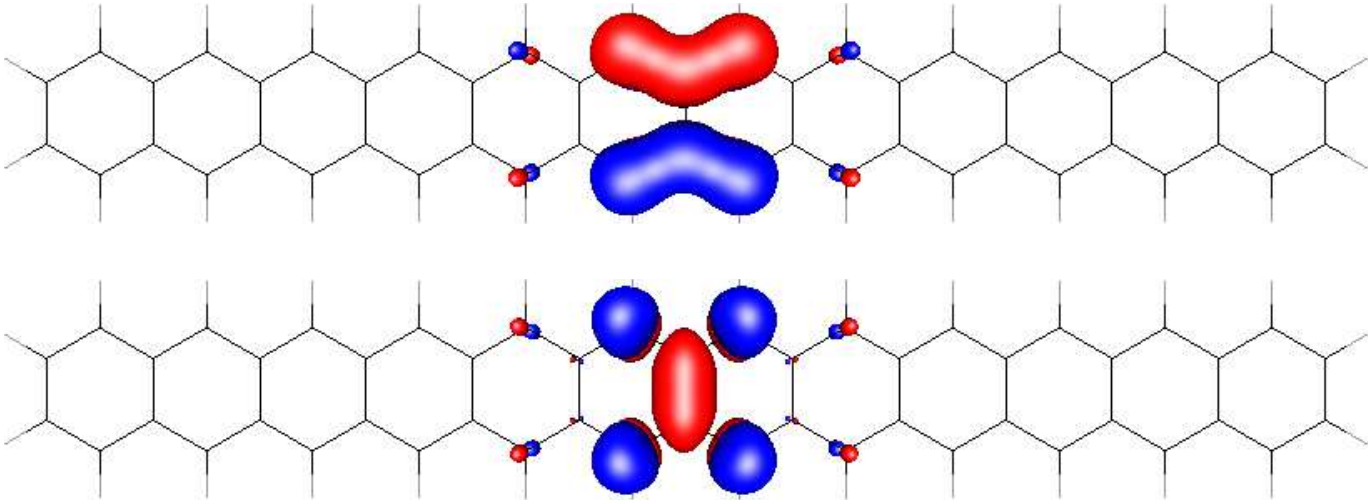


(b) Pair 2 orbitals. $\theta_k = 0.82807$.

FIG. 5



(a) Pair 1 orbitals.



(b) Pair 2 orbitals.

FIG. 6

## TABLES

TABLE I: CCVB optimization-cycle steps, their scalings, and example timings for 12-acene.

| Step | Scaling | avg. time % (STO-3G) | avg. time % (6-31G) |
|---|---|---|---|
| 1. integral transformation | $n_e^4$-$n_e^5$ | 46.67 | 88.76 |
| 2. $E_0$ | $n_e^2$ | 0.01 | 0.00 |
| 3. $\mu$, $\kappa$, $\omega$ | $n_e^3$ | 3.92 | 0.72 |
| 4. $t$, $\lambda$, $E$ | $n_e^3$ | 0.02 | 0.00 |
| 5. $\theta$ derivatives | $n_e^3$ | 37.10 | 6.96 |
| 6. $\Gamma$ | $n_e^3$ | 5.19 | 0.97 |
| 7. $\Delta$ gradient | $n_e^3$ | 0.08 | 0.04 |
| 8. $\Delta$ diagonal Hessian | $n_e^3$ | 6.85 | 2.38 |
| 9. GDM | $n_e^3$ | 0.16 | 0.17 |

TABLE II: CCVB results for acenes: Pipek-Mezey guess

| n | active e$^-$ | RHF Energy [a] | CCVB Energy [a] | Largest $\theta_k$ | Largest $\|t_{kl}\|$ | # iter. [b] |
|---|---|---|---|---|---|---|
| 4 | 84 | -680.2181369 | -680.9109851 | 0.92755 | 0.148141 | 131 |
| 5 | 102 | -830.9865879 | -831.8230161 | 0.94512 | 0.138115 | 75 |
| 6 | 120 | -981.7543923 | -982.7376693 | 0.94639 | 0.144413 | 81 |
| 7 | 138 | -1132.5219100 | -1133.6515428 | 0.95376 | 0.148738 | 102 |
| 8 | 156 | -1283.2892986 | -1284.5546786 | 0.95422 | 0.141423 | 87 |
| 9 | 174 | -1434.0566283 | -1435.4683416 | 0.96136 | 0.145324 | 102 |
| 10 | 192 | -1584.8239312 | -1586.3817953 | 0.96424 | 0.146976 | 109 |
| 11 | 210 | -1735.5912218 | -1737.2951455 | 0.96557 | 0.147757 | 94 |
| 12 | 228 | -1886.3585067 | -1888.2084464 | 0.96624 | 0.148141 | 105 |

[a] in Hartree

[b] conv. crit.: $10^{-7}$ a.u. grad. r.m.s.

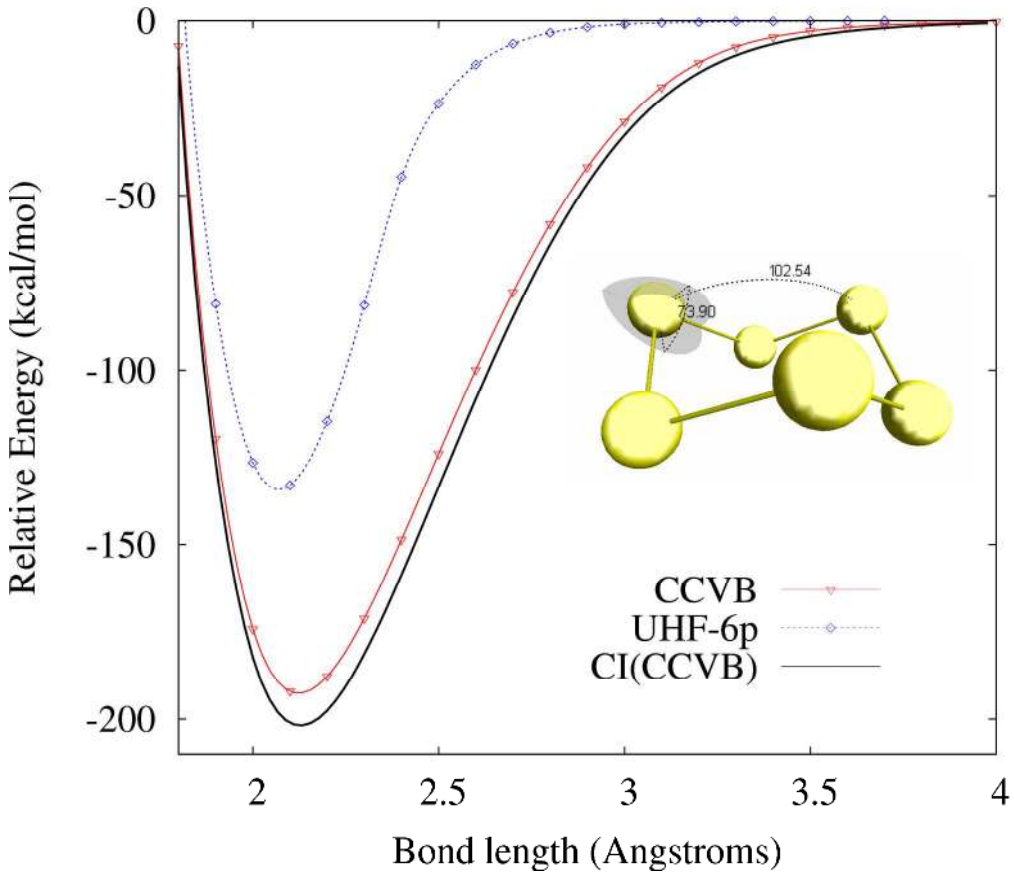
TABLE III: CCVB results for acenes: Boys guess

| n | active e$^-$ | Energy [a] | Largest $\theta_k$ | Largest $\|t_{kl}\|$ | $E$(Boys)-$E$(PM) [b] |
|---|---|---|---|---|---|
| 4 | 84 | -680.9109851 | 0.92755 | 0.148141 | 0 |
| 5 | 102 | -831.8230161 | 0.94512 | 0.138115 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| 6 | 120 | -982.7349707 | 0.94797 | 0.154769 | 1.69 |
| 7 | 138 | -1133.6461946 | 0.95243 | 0.145717 | 3.36 |
| 8 | 156 | -1284.5546786 | 0.95422 | 0.141423 | 0 |
| 9 | 174 | -1435.4673776 | 0.96265 | 0.149656 | 0.60 |
| 10 | 192 | -1586.3936550 | 1.39244 | 0.328789 | -7.44 |
| 11 | 210 | -1737.2891977 | 0.96687 | 0.147101 | 3.73 |
| 12 | 228 | -1888.2231568 | 1.40556 | 0.331675 | -9.23 |

[a] in Hartree

[b] in kcal/mol




177x153mm (300 x 300 DPI)