Vita Summary

1 Name and Surename

Name:	Katharina Boguslawski
Name at birth:	Katarzyna Bogusławska

2 Education

2009–2012	Dr. Sc. ETH Zurich, Department of Chemistry and Applied Biosciences, ETH Zurich, Switzerland; title of thesis: *Dissecting the quantum chemical spin state problem — Entropy measures, matrix product states and reconstruction algorithms*; supervisor: Prof. Dr. Markus Reiher; date of defence: 17.10.2012
2008–2009	M.Sc. Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Switzerland; title of thesis: *In search of accurate spin density distributions*; supervisor: Prof. Dr. Markus Reiher
2005–2008	B.Sc. Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Switzerland

3 Employment

07/2016–present	Assistant Professor, Faculty of Chemistry, Nicolaus Copernicus University in Torun, Poland (04/2017–12/2017: Career break for maternity leave)
04/2016–06/2016	Assistant Professor, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University in Torun, Poland
10/2015–03/2016	Research Assistant, Faculty of Physics, Astronomy and Infor-

Katarzyna Bogusławska

	matics, Nicolaus Copernicus University in Torun, Poland
07/2013–07/2015	Postdoctoral Fellow, Department of Chemistry and Chemical Biology, McMaster University, Canada
11/2012–06/2013	Postdoctoral Fellow, Laboratory of Physical Chemistry, ETH Zurich, Switzerland

4 Scientific achievements

(a) Title of scientific achievement

Series of publications:
"Modeling strong and weak correlation using electron-pair states"

(b) Authors, titles, year of publication, name of the journal

Additional information on the contributions of all authors are given in separate Appendices, including statements of the other coauthors.

[H1] **K. Boguslawski**, P. Tecmer, *"Benchmark of dynamic electron correlation models for seniority-zero wavefunctions and their application to thermochemistry"*, *J. Chem. Theory Comput.* 13 **2017**, 5966-5983 (**corresponding author**).
My estimated contribution to the paper: 80%

[H2] (a) **K. Boguslawski**, *"Targeting excited states in all-trans polyenes with electron-pair states"*, *J. Chem. Phys.* 145 **2016**, 234105 (**corresponding author**).
My estimated contribution to the paper: 100%
(b) **K. Boguslawski**, *"Erratum: "Targeting excited states in all-trans polyenes with electron-pair states" [J. Chem. Phys. 145, 234105 (2016)]"*, *J. Chem. Phys.* 147 **2017**, 139901 (**corresponding author**).
My estimated contribution to the paper: 100%

[H3] **K. Boguslawski**, P. Tecmer, Örs Legeza, *"Analysis of two-orbital correlations in wave functions restricted to electron-pair states"*, *Phys. Rev. B* 94 **2016**, 155126 (**corresponding author**).
My estimated contribution to the paper: 80%

[H4] **K. Boguslawski**, P. W. Ayers, *"Linearized Coupled Cluster Correction on the Antisymmetric Product of 1-reference orbital Geminals"*, *J. Chem. Theory Comput.* 11 **2015**, 5252-5261 (**corresponding author**).
My estimated contribution to the paper: 90%

[H5] (a) **K. Boguslawski**, P. Tecmer, *"Orbital entanglement in quantum chemistry"*, *Int. J. Quantum Chem.* 115 **2015**, 1289-1295 (**invited perspective, corresponding author**).
My estimated contribution to the paper: 50%
(b) **K. Boguslawski**, P. Tecmer, *"Erratum: Orbital entanglement in quantum chemistry"*, *Int. J. Quantum Chem.* **2017**, e25455 (doi:10.1002/qua.25455)

Katarzyna Boguslawski

(**corresponding author**).
My estimated contribution to the paper: 50%

[H6] **K. Boguslawski**, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, D. Van Neck, *"Nonvariational orbital optimization techniques for the AP1roG wave function"*, *J. Chem. Theory Comput.* 10 **2014**, 4873-4882 (**corresponding author**).
My estimated contribution to the paper: 80%

[H7] **K. Boguslawski**, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, D. Van Neck, *"Projected seniority-two orbital optimization of the antisymmetric product of one-reference orbital geminal"*, *J. Chem. Phys.* 140 **2014**, 214114.
My estimated contribution to the paper: 70%

[H8] **K. Boguslawski**, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, D. Van Neck, *"Efficient description of strongly correlated electrons with mean-field cost"*, *Phys. Rev. B* 89 **2014**, 201106(R).
My estimated contribution to the paper: 80%

(c) Discussion of scientific goals, obtained results, and their significance

c.1 Introduction and motivation

Over the last decades, theoretical approaches have been successfully used to determine molecular properties and to provide a fundamental understanding of chemical reactivity and reaction mechanisms. Specifically, theoretical methods are particularly instructive when experimental studies of thermodynamics, kinetics, complexation, and reaction mechanisms are complicated due to, for instance, toxicity, radioactivity, and instability of chemical compounds.[1,2] Most importantly, a reliable quantum-mechanical treatment must address the fact that electrons do not move independently, but in a *correlated* fashion.
However, the accurate modeling of the correlated motion of electrons remains an open problem in theoretical chemistry. This difficulty originates from the different contributions that govern the correlated motion of electrons,[3] commonly referred to as *strong* and *weak correlation*. An exact determination of these effects can be obtained from the Full-Configuration-Interaction (FCI) method. Unfortunately, FCI is only computationally feasible for small systems[4] with up to approximately 20 electrons. Its computational limitations led to the development of various approximate methods, which form hierarchies of approximations to FCI. In quantum chemistry, two main ansätze can be distinguished: density functional theory (DFT) and wave function-based methods. Although DFT is superior to wave function methods in terms of efficiency and cost, it is unreliable for systems with strong electron correlations like transition metal and actinide compounds[5,6] or bond-breaking processes. To treat strong electron correlation, multiconfigurational wave function-based methods can be applied. The most popular and well-established approaches are the Complete-Active-Space Self-Consistent-Field (CASSCF) ansatz[7] and its extensions,[8,9] the Multi-Reference

Katarzyna Boguslawski

Configuration-Interaction Singles and Doubles (MR-CISD) approach,[10] and Multi-Reference Coupled Cluster methods.[11] The former methods are usually employed in combination with a subsequent perturbative treatment of weak correlation.[12–15] However, standard multi-reference methods are computationally very expensive and their computational cost scales exponentially with the size of the system, an effect known as the *curse of dimension*. This resource bottleneck limits their application to small molecular system, typically small molecular building blocks of realistic materials.[1,10,16–21,21–32]

To overcome the exponential-scaling wall of standard *ab initio* methods, unconventional electron correlation approaches have been introduced into quantum chemistry; most of these approaches are based on compact parameterizations of the many-electron wave function. One example is the Density-Matrix Renormalization-Group (DMRG)[33–38] algorithm, which is a cheaper alternative to CASSCF methods. DMRG has been already successfully applied in transition metal chemistry,[39–46] where it vanquished notorious failures of conventional approaches. Recent research suggests that DMRG represents a promising alternative to study actinide chemistry.[47] Although DMRG is much less computationally demanding than standard quantum mechanical models for strong correlation, it is still very expensive, so only small building blocks of large actinide compounds can be investigated using the DMRG algorithm.

c.2 Electronic wave functions based on electron pairs

Alternatively, strongly correlated materials can be efficiently modeled using approaches based on non-interacting electron pairs. The electronic wave function is then constructed from electron-pair states,[48] called *geminals*. An advantage of this approach is that intra-pair electron correlations are taken into account from the beginning, and in many cases this captures the dominant fraction of strong—and to a smaller extent, weak—correlation.[49–52] Examples for geminal-based approaches are the Antisymmetric Product of Strongly orthogonal Geminals (APSG),[53–58] the Antisymmetrized Geminal Power[59–61] (which is a special case of projected Hartree–Fock–Bogoliubov[62]), the Antisymmetric Product of Interacting Geminals[48,54,63–73] (APIG), Generalized Valence Bond[74–78] (GVB), and the Antisymmetric Product of 1 reference orbital Geminals (AP1roG).[52] These new geminal-based methods are distinguished from other multi-reference methods by their comparatively negligible resource requirements, which are comparable to DFT methods. This makes geminal-based approaches ideal candidates for modeling strongly-correlated materials.

A promising family of geminal-based wave function ansätze allows us to approximate the APIG wave function efficiently, resulting in a computationally tractable model. One example is AP1roG that represents an efficient parameterization of the Doubly-Occupied (DO) CI wave function, requiring only mean-field computational cost. In contrast, traditional DOCI implementations suffer from factorial scaling. The AP1roG wave function ansatz can be written in terms of one-particle functions as a fully general pair-coupled-cluster-doubles (pCCD) wave function,[79] i.e.,

$$|\mathrm{AP1roG}\rangle = \exp\left(\sum_{i=1}^{P}\sum_{a=P+1}^{K} c_i^a a_a^\dagger a_{\bar{a}}^\dagger a_{\bar{i}} a_i\right)|\Phi_0\rangle, \tag{1}$$

Katarzyna Boguslawski

where $a_p^\dagger$, $a_{\bar{p}}^\dagger$ and a_p, $a_{\bar{p}}$ are the electron creation and annihilation operators for α-(p) and β-electrons ($\bar{p}$), and $|\Phi_0\rangle$ is some independent-particle wave function (usually, but not restricted to the Hartree–Fock (HF) determinant). In the above equation, we used the standard notation where indices i and a correspond to occupied and virtual orbitals with respect to $|\Phi_0\rangle$. P denotes the number of electron pairs ($P = N/2$ with N being the total number of electrons) and K is the number of one-particle functions. $\{c_i^a\}$ are the geminal coefficients and link the geminal wave function with the underlying one-particle basis functions.[80] Especially, the geminal coefficient matrix encodes the orbital-pairing scheme in the geminal wave function. We should emphasize that, in the AP1roG ansatz, all virtual orbitals are allowed to contribute to each geminal. Thus, unlike APSG and GVB-PP, this approach does not require the orbital-pairing scheme to be optimized.[81]

c.3 Ensuring size-consistency

Unfortunately, the AP1roG wave function ansatz in eq. (1) is not size-consistent[82] and hence does not provide reliable potential energy surfaces. To ensure size-consistency, we have to optimize the one-particle basis functions. This can be done in a fully variational manner **[H8]**, analogous to orbital-optimized coupled cluster[83] (OCC), or using approximate seniority-based projection schemes **[H6,H7]**. In the following, we will briefly summarize the main steps and ideas of our optimization procedures introduced in **[H6-H8]**.

c.3.1 Variational orbital optimization

In **[H8]**, we presented the first variational orbital optimization of the AP1roG wave function using a Lagrange formulation. The orbitals are then chosen to minimize the AP1roG energy expression subject to the constraint that the wave function amplitude (or geminal coefficient) equations are satisfied. The energy Lagrangian takes thus the form

$$\mathcal{L} = \langle\Phi_0|e^{-\boldsymbol{\kappa}}\hat{H}e^{\boldsymbol{\kappa}}|\mathrm{AP1roG}\rangle + \sum_{i,a}\lambda_i^a\big(\langle\Phi_{i\bar{i}}^{a\bar{a}}|e^{-\boldsymbol{\kappa}}\hat{H}e^{\boldsymbol{\kappa}}|\mathrm{AP1roG}\rangle - Ec_i^a\big), \tag{2}$$

with $\{\lambda_i^a\}$ being the Lagrange multipliers. In the above equation, $\boldsymbol{\kappa}$ is the generator of orbital rotations

$$\boldsymbol{\kappa} = \sum_{p>q}\kappa_{pq}(a_p^\dagger a_q - a_q^\dagger a_p), \tag{3}$$

where (κ_{pq}) is a skew-symmetric matrix and transforms into a new orthogonal basis with a transformation $U = e^{\boldsymbol{\kappa}}$ and $e^{-\boldsymbol{\kappa}}\hat{H}e^{\boldsymbol{\kappa}}$ is the Hamiltonian in the rotated basis.[83] $|\Phi_{i\bar{i}}^{a\bar{a}}\rangle$ denotes a single-pair-excited determinant with respect to the reference determinant $|\Phi_0\rangle$. The variational orbital gradient $\boldsymbol{g}$ is obtained as the partial derivative of the Lagrange energy functional with respect to the orbital rotation coefficients $\{\kappa_{pq}\}$

Katarzyna Boguslawski

evaluated at $\boldsymbol{\kappa} = 0$ **[H6,H7]**,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \kappa_{pq}}\Big|_{\boldsymbol{\kappa}=0} =& \langle \Phi_0 + \sum_{i,a} \lambda_i^a \Phi_{i\bar{i}}^{a\bar{a}} | \Big[(a_p^\dagger a_q - a_q^\dagger a_p), \hat{H} \Big] | \mathrm{AP1roG} \rangle \\ & - \langle \Phi_0 | \Big[(a_p^\dagger a_q - a_q^\dagger a_p), \hat{H} \Big] | \mathrm{AP1roG} \rangle \sum_{ia} \lambda_i^a c_i^a. \end{aligned} \tag{4}$$

In contrast to conventional OCC theory, occupied–occupied and virtual–virtual orbital rotations are non-redundant and have to be considered in the orbital gradient equations. Thus, the indices p and q run over all occupied and virtual orbitals. The Lagrange multipliers $\{\lambda_i^a\}$ require the solution of an additional set of equations defined by $\partial \mathcal{L}/\partial c_i^a$, again evaluated at $\boldsymbol{\kappa} = 0$, which leads to a set of equations for the Lagrange multipliers, analogous to the Λ-equations in CC theory,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial c_i^a}\Big|_{\boldsymbol{\kappa}=0} =& \langle \Phi_0 | \hat{H} a_a^\dagger a_{\bar{a}}^\dagger a_{\bar{i}} a_i | \mathrm{AP1roG} \rangle + \sum_{jb} \lambda_j^b \langle \Phi_{j\bar{j}}^{b\bar{b}} | \hat{H} a_a^\dagger a_{\bar{a}}^\dagger a_{\bar{i}} a_i | \mathrm{AP1roG} \rangle \\ & - E\lambda_i^a - \langle i\bar{i} | a\bar{a} \rangle \sum_{jb} \lambda_j^b c_j^b = 0. \end{aligned} \tag{5}$$

The requirement that the derivative of $\mathcal{L}$ with respect to the Lagrange multipliers $\{\lambda_i^a\}$ is stationary results in the standard set of equations for the geminal coefficients

$$\frac{\partial \mathcal{L}}{\partial \lambda_i^a}\Big|_{\boldsymbol{\kappa}=0} = \langle \Phi_{i\bar{i}}^{a\bar{a}} | \hat{H} | \mathrm{AP1roG} \rangle - E c_i^a = 0. \tag{6}$$

To obtain the unitary transformation matrix U, we first expand the energy Lagrangian as a function of $\boldsymbol{\kappa}$ up to second order

$$\mathcal{L}^{(2)}(\boldsymbol{\kappa}) = \mathcal{L}^{(0)} + \boldsymbol{\kappa}^\dagger \boldsymbol{g} + \frac{1}{2} \boldsymbol{\kappa}^\dagger \boldsymbol{A} \boldsymbol{\kappa}, \tag{7}$$

where $\boldsymbol{A}$ is the molecular orbital Hessian and $\boldsymbol{g}$ is the orbital gradient whose elements are defined in eq. (4). Thus, minimizing the Lagrangian with respect to $\{\kappa_{pq}\}$ leads to the well-known equation for the orbital rotation coefficients

$$\boldsymbol{\kappa} = -\boldsymbol{A}\boldsymbol{g}. \tag{8}$$

After the orbital gradient and Hessian are determined, the matrix representation of κ can be evaluated from the above equation. The transformation matrix is then approximated as $U \approx 1 + \boldsymbol{\kappa} + \frac{1}{2}\boldsymbol{\kappa}^\dagger \boldsymbol{\kappa}$ and orthogonalized.

c.3.2 Non-variational orbital optimization schemes

We also developed different (approximate) orbital optimization techniques that exploit the generalized Brillouin theorem[84] of a seniority-zero wave function for the orbital parameters.[85,86] The corresponding algorithms are discussed in **[H6,H7]**. In contrast to the variational orbital optimization, the non-variational orbital optimization schemes do not require the solution of the Λ-equations of AP1roG. Instead, our

Katarzyna Boguslawska

methods are based on the assumption that the seniority-zero-plus-two ($\Omega = 0, 2$) sectors and the seniority-zero ($\Omega = 0$) sector can be decoupled (the seniority number is defined as the number of unpaired electrons in a Slater determinant). The optimal set of orbitals is chosen such that the decoupling condition is satisfied. To understand the relation between orbital optimization and decoupling of the $\Omega = 0, 2$ and $\Omega = 0$ sectors, we assume that $\Psi_{\text{oo}}^{(0)}$ is a CI expansion comprising only closed-shell configurations, *i.e.*, a seniority-zero wave function, constructed from a set of optimized orbitals. Further, consider $\Psi_{\text{oo}}^{(0,2)}$ to be CI expansion containing both closed-shell Slater determinants and Slater determinants with exactly two unpaired electrons, *i.e.*, a CI expansion restricted to the seniority-zero-plus-two sectors, again with optimized orbitals. Assuming that the optimal seniority-zero-plus-two solution is very close to the optimal seniority-zero solution, we can write

$$\Psi_{\text{oo}}^{(0,2)} = \Psi_{\text{oo}}^{(0)} + \Psi_{\text{oo}}^{(2)} = \big(1 + \sum_{p \neq q} t_{pq} a_p^\dagger a_q\big) \Psi_{\text{oo}}^{(0)}, \tag{9}$$

where $\{t_{pq}\}$ are some expansion coefficients. The seniority-zero-plus-two sectors and the seniority-zero sector can be decoupled if CI expansions restricted to the $\Omega = 0, 2$ and $\Omega = 0$ sector have the same energy expectation value. Thus, the following equation must hold

$$\langle \Psi_{\text{oo}}^{(0,2)} | \hat{H} | \Psi_{\text{oo}}^{(0,2)} \rangle - \langle \Psi_{\text{oo}}^{(0)} | \hat{H} | \Psi_{\text{oo}}^{(0)} \rangle = 0. \tag{10}$$

The above condition is satisfied to first order in $\{t_{pq}\}$ if (t_{pq}) is a skew-symmetric matrix. In this case, the seniority-two contribution of $\Psi_{\text{oo}}^{(0,2)}$ (the second term on the right hand side of eq. (9)) can be written as

$$\Psi_{\text{oo}}^{(2)} = \sum_{p > q} t_{pq} (a_p^\dagger a_q - a_q^\dagger a_p) \Psi_{\text{oo}}^{(0)}. \tag{11}$$

Using the above equation, the decoupling condition eq. (10) can be straightforwardly simplified. Keeping only terms up to first order in $\{t_{pq}\}$, we obtain the approximate decoupling condition

$$\langle \Psi_{\text{oo}}^{(2)} | \hat{H} | \Psi_{\text{oo}}^{(0)} \rangle + \langle \Psi_{\text{oo}}^{(0)} | \hat{H} | \Psi_{\text{oo}}^{(2)} \rangle + \mathcal{O}(\boldsymbol{t}^2) = 0.$$

Substituting eq. (11) in the above equation, the approximate decoupling condition can be reformulated as

$$\langle \Psi_{\text{oo}}^{(0)} | \Big[\big(a_q^\dagger a_p - a_p^\dagger a_q\big), \hat{H} \Big] | \Psi_{\text{oo}}^{(0)} \rangle = 0 \quad \forall p > q, \tag{12}$$

which is equivalent to the generalized Brillouin theorem[84] of a seniority-zero wave function for the orbital parameters.[85,86]

Therefore, the decoupling of the seniority-zero-plus-two sectors and the seniority-zero sector is equivalent to satisfying the orbital-dependent part of the generalized Brillouin theorem of an $\Omega = 0$ wave function to first order. In cases where higher order terms in $\{t_{pq}\}$ become important, the simplification of eq. (10) to satisfy eq. (12)

is not valid and a seniority-zero-plus-two wave function will result in a lower energy expectation value than a CI expansion constrained to the seniority-zero sector.[87] In the following, eq (12) will be used as a starting point to derive different approximate orbital-optimization schemes.

Choosing the optimal set of orbitals such that eq. (12) is satisfied scales factorially with system size. We can reduce the computational cost by making two assumptions as outlined in Ref. 52. First, we assume that the AP1roG model is a decent approximation to a seniority-zero wave function, *i.e.*, $|\text{AP1roG}\rangle \approx |\Psi^{(0)}\rangle$. The good performance of AP1roG in approximating the doubly occupied CI (DOCI) wave function (a wave function restricted to the seniority-zero sector) has been shown in many numerical examples.[52,82,88,89] Then, the rotated set of orbitals can be obtained by solving

$$\langle \text{AP1roG}|\Big[\big(a_q^\dagger a_p - a_p^\dagger a_q\big), \hat{H}\Big]|\text{AP1roG}\rangle = 0 \quad \forall p > q. \tag{13}$$

Yet, the cost of solving the above equation still grows factorially with system size and we are compelled to make further simplifications for the sake of computational tractability. Our second, more pragmatic assumption is, therefore, to restrict the excitation manifold to all singly-pair excited determinants with respect to the reference determinant $|\Phi_0\rangle$. Thereby, the optimal set of orbitals can be determined in a computationally efficient way. In doing so, eq. (13) reduces to

$$\langle \Phi_0 + \sum_{i,a} c_i^a \Phi_{i\bar{i}}^{a\bar{a}}|\Big[\big(a_q^\dagger a_p - a_p^\dagger a_q\big), \big(e^{-\boldsymbol{\kappa}}\hat{H}e^{\boldsymbol{\kappa}}\big)\Big]|\text{AP1roG}\rangle = 0 \quad \forall p > q, \tag{14}$$

where we have written the Hamiltonian explicitly in the rotated basis. Eq. (14) represents our starting point for different orbital-optimization procedures that aim at decoupling the seniority-zero and seniority-zero-plus-two sectors.

In our first approach, no further approximations are made and the $\{\kappa_{pq}\}$ are optimized such that eq. (14) is fulfilled. Since the action of $a_p^\dagger a_q$ on the bra-state (or equivalently on the ket-state) of eq. (14) generates a seniority-two wave function, the optimized set of orbitals is obtained by requiring that the projection of the seniority-two sector on the AP1roG wave function vanishes, *i.e.*, the Hamiltonian does not connect the seniority-two and seniority-zero sectors in the case of an optimal basis. Thus, we will call eq. (14) the projected-seniority-two condition using the commutator formulation (PS2c) and the left hand side of eq. (14) the PS2c orbital gradient $\boldsymbol{g}^{(\text{PS2c})}(\boldsymbol{\kappa})$. Note that the PS2c method bears similarity to the variational orbital optimization scheme. While in the variational orbital optimization procedure we need to solve for $\{\lambda_i^a\}$ to construct all important intermediates (one- and two-particle response density matrices), only the geminal coefficients are required in the PS2c approach.

In our second (non-variational) orbital-optimization procedure, we impose the stronger condition that each term of the commutator vanishes separately,

$$\langle \Phi_0 + \sum_{i,a} c_i^a \Phi_{i\bar{i}}^{a\bar{a}}|\big(a_q^\dagger a_p - a_p^\dagger a_q\big)\big(e^{-\boldsymbol{\kappa}}\hat{H}e^{\boldsymbol{\kappa}}\big)|\text{AP1roG}\rangle = 0 \quad \forall p > q. \tag{15}$$

Note that the above equation is still antisymmetric in its indices p, q. We will refer to eq. (15) as the antisymmetric PS2 condition (PS2a) and the left hand side of eq. (15)

Katarzyna Boguslawski

as the PS2a orbital gradient $\boldsymbol{g}^{(\mathrm{PS2a})}(\boldsymbol{\kappa})$.
In our most stringent approximation, we require that the first term of eq. (15) equals zero individually and the projected-seniority-two condition simplifies to

$$\langle \Phi_0 + \sum_{i,a} c_i^a \Phi_{i\bar{i}}^{a\bar{a}} | a_q^\dagger a_p (e^{-\boldsymbol{\kappa}} \hat{H} e^{\boldsymbol{\kappa}}) | \mathrm{AP1roG} \rangle = 0 \quad \forall p > q. \tag{16}$$

This choice was motivated by the analogy to the Brillouin theorem of HF theory. Note that in the case of HF theory, the above relation is known as the Brillouin theorem for a closed-shell Slater determinant. The above equation corresponds to the PS2 condition and its left hand side is the PS2 orbital gradient $\boldsymbol{g}^{(\mathrm{PS2})}(\boldsymbol{\kappa})$. For all non-variational orbital optimization techniques, the matrix representation of $\boldsymbol{\kappa}$, and thus the unitary transformation matrix, is determined from eq. (8) using the corresponding orbital gradient $\boldsymbol{g}^{(\mathrm{PS2x})}(\boldsymbol{\kappa} = 0)$ evaluated for the current set of orbitals and the (diagonal approximation of the) orbital Hessian $\boldsymbol{A}_{pq,pq} = \frac{\partial \boldsymbol{g}_{pq}}{\partial \kappa_{pq}}\big|_{\boldsymbol{\kappa}=0}$.

c.3.3 Assessing the performance of orbital-optimized AP1roG

The accuracy of the variational and non-variational orbital optimization methods has been benchmarked against a set of challenging multi-reference systems. Specifically, we have investigated the dissociation process of hydrogen chains containing up to 50 hydrogen atoms **[H7,H8]** and second-row diatomics (LiF, C_2) **[H7]**, the Be–H_2 insertion reaction, the automerization process of cyclobutadiene, and the energetic stability of the monocyclic and bicyclic forms of the pyridyne molecule **[H6]**. As an

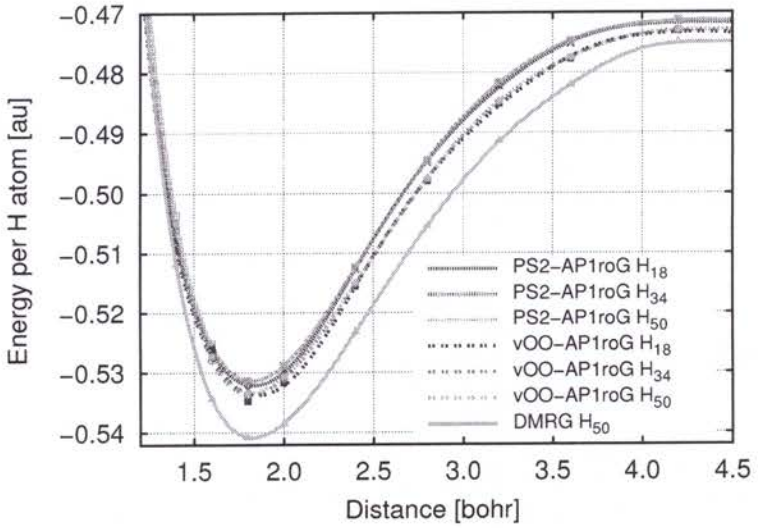


Figure 1: Fitted total energy per hydrogen atom curves for hydrogen chains of different lengths using the PS2 and variational orbital optimization procedures. The DMRG reference data has been taken from ref. 90.

example, we present the dissociation process of linear hydrogen chains (see Figure 1), the barrier heights of the automerization process of cyclobutadiene (see Table 1), and the energetic stability of the monocyclic and bicyclic forms of pyridyne (summarized in Table 2).

The symmetric dissociation of linear hydrogen chains up to 50 hydrogen atoms is a commonly-used molecular model for strongly-correlated systems and remains a remarkably difficult task for conventional quantum-chemistry approaches.[90–93] Furthermore, studying the (symmetric) dissociation of hydrogen chains allows us to numerically assess to what extent the PS2 optimization procedure recovers size-consistency. Figure 1 shows the fitted DMRG reference energy-per-atom curve for the symmetric stretching problem of H_{50}. The overall agreement of PS2-/voo(variational orbital optimization)-AP1roG with DMRG is good. Larger deviations from the DMRG reference are found around the equilibrium distance and originate from dynamic correlation effects that cannot be captured by AP1roG, but can be included *a posteriori* (*vide infra*). For stretched internuclear distances, the PS2-/voo-AP1roG energy-per-atom curves are closely parallel to the DMRG reference.

Furthermore, our numerical studies on cyclobutadiene and pyridyne reveal that the accuracy of PS2a- and PS2c-AP1roG is similar to voo-AP1roG in most cases. In particular, the PS2c orbital optimization scheme yields total electronic energies and energy splittings that are similar to the variational orbital optimization counterpart. As expected, the PS2a-AP1roG approach gives total electronic energies that are higher than PS2c-AP1roG and voo-AP1roG, because of additional constraints in the orbital gradient that are not present in PS2c-AP1roG and voo-AP1roG. For all investigated (multi-reference) molecules, both PS2a- and PS2c-optimized molecular orbitals are essentially indistinguishable from voo-AP1roG molecular orbitals. Furthermore, all three orbital optimization methods produce symmetry-broken solutions (localized, hybrid molecular orbitals). As an example, the optimized valence molecular orbitals

Table 1: Barrier heights of the automerization process of cyclobutadiene. Differences with respect to the multi-reference Mk-MRCCSD(T) results are given in parentheses. GS: ground state; TS: transition state.

Method	Total energy [Hartree]		Barrier height	
	GS	TS	[Hartree]	[kcal/mol]
PS2-AP1roG	−153.719 321	−153.690 619	0.028 702	18.0 (+10.2)
PS2a-AP1roG	−153.884 005	−153.850 491	0.033 514	21.0 (+13.2)
PS2c-AP1roG	−153.886 993	−153.854 496	0.032 497	20.4 (+12.6)
voo-AP1roG	−153.887 097	−153.854 631	0.032 466	20.4 (+12.6)
MP2	−153.643 539	−153.592 092	0.051 447	32.3 (+24.5)
CAS(4,4)SCF	−153.713 999	−153.630 231	0.083 768	52.6 (+44.8)
NEVPT2/CAS(4,4)	−154.190 998	−154.125 766	0.065 232	40.9 (+33.1)
CAS(20,16)SCF	−153.814 502	−153.758 254	0.056 248	35.3 (+27.5)
NEVPT2/CAS(20,16)	−154.167 433	−154.101 856	0.065 577	41.2 (+33.4)
Mk-MRCCSD(T)[94]	–	–	–	7.8
Experiment[95]	–	–	–	1.6–10

Katarzyna Bystrek

Table 2: Energetic stability of the monocyclic and bicyclic forms of the pyridyne molecule. Differences with respect to the Mk-MRCCSD(T) reference value are given in parentheses.

Method[a]	Total energy [Hartree]		ΔE(bicyclic,monocyclic)	
	monocyclic	bicyclic	[Hartree]	[kcal/mol]
PS2a-AP1roG	−245.806 197	−245.797 390	0.008 807	5.5 (−3.3)
PS2c-AP1roG	−245.822 958	−245.798 196	0.024 763	15.5 (+6.7)
voo-AP1roG	−245.823 951	−245.799 810	0.024 140	15.2 (+6.4)
MP2	−246.414 371	−246.396 546	0.017 825	11.2 (+2.4)
CAS(8,8)SCF	−245.500 591	−245.467 299	0.033 291	20.9 (+12.1)
NEVPT2/CAS(8,8)	−246.403 581	−246.393 769	0.009 812	6.2 (−2.6)
CCSD[96]	–	–	–	−3.6 (−12.4)
tailored CCSD[11]	–	–	–	6.8 (−2.0)
tailored CCSD(T)[11]	–	–	–	9.0 (+0.2)
Mk-MRCCSD(T)[96]	–	–	–	8.8

[a] In AP1roG, MP2, CASSCF and NEVPT2 calculations we used the same geometries as in Ref. 97

of the ground and transition state of cyclobutadiene are displayed in Figure 2.

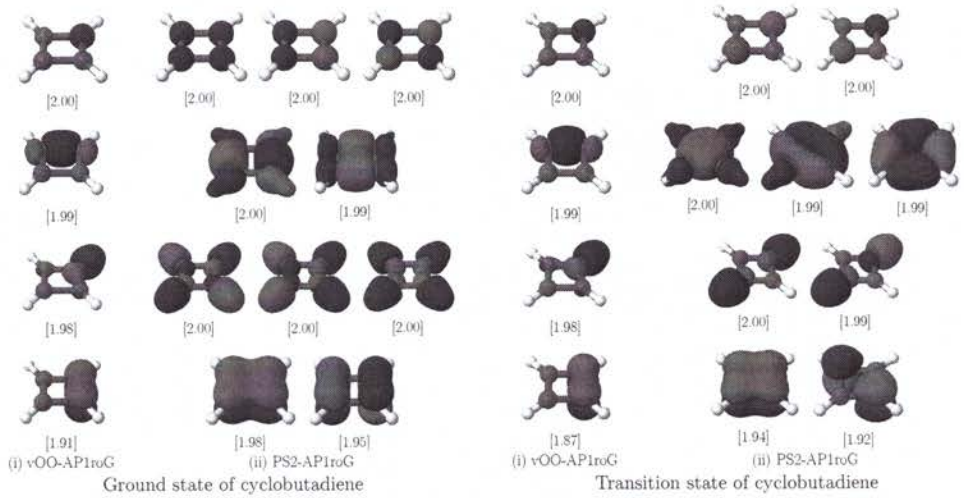


Figure 2: Valence occupied voo- and PS2-AP1roG molecular orbitals for the rectangular and square (transition state) geometries of cyclobutadiene. The voo-, PS2c- and PS2a-AP1roG optimization techniques yield qualitatively the same molecular orbitals.

In contrast to PS2c-, PS2a-, and voo-AP1roG, the PS2 optimization and its convergence behavior are sensitive to the initial guess orbitals. Specifically, we observed

Katarzyna Boguslaw

that PS2-AP1roG tends to alleviate symmetry-breaking and leads to more delocalized molecular orbitals if localized orbitals are used as initial guess. However, the PS2 orbital optimization scheme suffers from convergence difficulties for larger molecular systems, which can be related to the severe approximations made in the PS2 orbital gradient. By contrast, the convergence of PS2a- and PS2c-AP1roG is more robust. While the PS2 orbital optimization leads to faster and smoother convergence when (delocalized) canonical orbitals are used as initial guess, PS2a-, PS2c-, and voo-AP1roG converge faster when localized orbitals are provided as starting orbitals. Although PS2- and PS2a-AP1roG require parts of the three-particle reduced density matrix for the calculation of the orbital gradient, all presented AP1roG orbital-optimization schemes are limited by the four-index transformation of the two-electron repulsion integrals that scales as $\mathcal{O}(N^5)$. We should note that a level shift had to be applied to the orbital Hessian in the PS2a-approach so that smooth convergence could be achieved. This, however, destroys the quadratic convergence. Furthermore, convergence difficulties of all non-variational orbital optimization schemes might be attributed to our pragmatic assumption of restricting the projection manifold to singly-pair excited determinants. This problem could be alleviated by including also higher pair excitations. In practical applications, we, therefore, recommend voo-AP1roG because of its stable convergence and its robustness against the initial guess orbitals for larger molecular systems (see also **[H6,H7]** for additional examples). The non-variational orbital-optimization techniques, however, can be used to dodge local minima if convergence difficulties are encountered in the variational orbital-optimization scheme.

Finally, our studies on multi-reference problems support that orbital-optimized AP1roG can be considered as an alternative to standard quantum chemistry methods. For the systems investigated in **[H6-H8]**, PS2a-, PS2c- and voo-AP1roG approaches provide lower total electronic energies than CASSCF and yield energy splittings that are considerably closer to NEVPT2 or multi-reference CC reference data.

c.4 Assessing the accuracy of geminal-based wave functions using concepts of quantum information theory

In the previous section, we scrutinized that AP1roG combined with an orbital optimization protocol can reliably model strongly-correlated systems, even in molecules with multiple degenerate single-particle states. However, most of the analysis presented so far was based on energetic arguments or one-body correlation functions like occupation numbers, neglecting any in-depth analysis of the AP1roG wave function. In our follow-up publication **[H3]**, we scrutinized how reliably AP1roG can approximate the (exact) electronic wave function for one-dimensional systems where quantum fluctuations have a more pronounced role. Specifically in **[H3]**, we use concepts of quantum information theory to assess the accuracy of electronic wave functions in terms of orbital entanglement and orbital-pair correlations.[43,98–100,100–103] As we outline in our perspective article **[H5]**, these measures are particularly instructive to dissect electron correlation effects,[43,99] elucidate chemical reactions,[101,104–109] detect changes in the electronic wave function,[110–112] and define active orbital spaces.[113,114] In this section, we will utilize orbital correlations to dissect electron correlation effects

Katayna Boguslawski

captured by the AP1roG model. Our orbital entanglement and correlation analysis exploits two entropic measures, the so-called single-orbital entropy and orbital-pair mutual information, to quantify the entanglement of and the correlation between orbitals.
The entanglement entropy of orbital i, also called single-orbital entropy, can be calculated from the eigenvalues of the one-orbital reduced density matrix $\omega_{\alpha;i}$,[115]

$$s_i = -\sum_{\alpha=1}^{4} \omega_{\alpha;i} \ln \omega_{\alpha;i}. \tag{17}$$

The single-orbital entropy is thus the von Neumann entropy of the reduced density matrix of the orbital of interest whose elements can be calculated from the one- and two-particle reduced density matrices **[H5]**, γ_q^p and Γ_{rs}^{pq}, where for a given wave function $|\Psi\rangle$

$$\gamma_q^p = \frac{\langle\Psi|a_p^\dagger a_q|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{18}$$

and

$$\Gamma_{rs}^{pq} = \frac{\langle\Psi|a_p^\dagger a_q^\dagger a_s a_r|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{19}$$

or from generalized correlation functions.[100,101] The one-orbital reduced density matrix ρ_i is spanned by the basis states of the one-orbital Fock space ($|—\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, $|\updownarrow\rangle$) and is thus a 4×4 matrix. Similarly, the entanglement of two orbitals is quantified by the two-orbital entropy $s_{i,j}$,

$$s_{i,j} = -\sum_{\alpha=1}^{16} \omega_{\alpha;i,j} \ln \omega_{\alpha;i,j}, \tag{20}$$

where $\omega_{\alpha;i,j}$ are the eigenvalues of the two-orbital reduced density matrix $\rho_{i,j}$, which is defined in terms of basis states of a two-orbital Fock space (16 possible states in the case of spatial orbitals). In contrast to ρ_i, the matrix elements of $\rho_{i,j}$ can be written in terms of the elements of the 1-, 2-, 3-, and 4-particle reduced density matrices, γ_q^p, Γ_{rs}^{pq}, Γ_{stu}^{pqr}, and Γ_{tuvw}^{pqrs}, with

$$\Gamma_{stu}^{pqr} = \frac{\langle\Psi|a_p^\dagger a_q^\dagger a_r^\dagger a_u a_t a_s|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{21}$$

and

$$\Gamma_{tuvw}^{pqrs} = \frac{\langle\Psi|a_p^\dagger a_q^\dagger a_r^\dagger a_s^\dagger a_w a_v a_u a_t|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{22}$$

Given s_i and $s_{i,j}$, we can quantify the correlations between two orbitals i and j by the orbital-pair mutual information,[98,115,116]

$$I_{i|j} = s_i + s_j - s_{i,j}, \tag{23}$$

which includes correlations of both classical and quantum origin. It is generally

Katarzyna Boguslawska

accepted that the mutual information measure pairwise correlations. In the following, we will briefly discuss the simplifications that can be made when dealing with seniority-zero wave functions, which allow us to evaluate the one- and two-orbital reduced density matrices with no additional cost compared to the optimization of the AP1roG wave function **[H3,H5]**.

c.4.1 Correlation functions for seniority-zero wave functions

If the electronic wave function is a CI-expansion with pair-excited Slater determinants only, that is a seniority-zero wave function, ρ_i and $\rho_{i,j}$ have a particular simple form. Restricting the wave function expansion to either doubly-occupied or unoccupied orbitals, the 4×4 matrix of ρ_i reduces to a 2×2 matrix, while the 16×16 matrix representation of $\rho_{i,j}$ becomes a 4×4 matrix. Furthermore, for seniority-zero wave functions, we can use the relations $\gamma^p_p = \gamma^{\bar{p}}_{\bar{p}} = \Gamma^{p\bar{p}}_{p\bar{p}}$ and $\Gamma^{p\bar{q}}_{p\bar{q}} = {}^4\Gamma^{p\bar{p}q\bar{q}}_{p\bar{p}q\bar{q}}$ [117] ($\bar{p}$ is equivalent to $p_\downarrow$) so that only the 1- and 2-particle reduced density matrcies are required to determine ρ_i and $\rho_{i,j}$. Specifically, we have

$$\rho_i = \begin{pmatrix} 1-\gamma^i_i & 0 \\ 0 & \gamma^i_i \end{pmatrix} \tag{24}$$

for the seniority-zero one-orbital RDM expressed in the basis $\{-,\#\}$, and

$$\rho_{j,i} = \begin{pmatrix} 1-\gamma^i_i-\gamma^j_j+\Gamma^{j\bar{i}}_{j\bar{i}} & 0 & 0 & 0 \\ 0 & \gamma^i_i-\Gamma^{j\bar{i}}_{j\bar{i}} & -\Gamma^{i\bar{i}}_{j\bar{j}} & 0 \\ 0 & -\Gamma^{j\bar{j}}_{i\bar{i}} & \gamma^j_j-\Gamma^{j\bar{i}}_{j\bar{i}} & 0 \\ 0 & 0 & 0 & \Gamma^{j\bar{i}}_{j\bar{i}} \end{pmatrix} \tag{25}$$

for the seniority-zero two-orbital RDM expressed in the basis $\{--,-\#,\#-,\#\#\}$. We should note that, for a seniority-zero wave function, the maximum value of s_i is $\ln 2$ (in contrast to $\ln 4$ for the general one-orbital RDM).

For AP1roG, the response 1- and 2-particle RDMs are used to construct ρ_i and $\rho_{i,j}$ and are defined as

$$\gamma^p_q = \langle\Phi_0|(1+\Lambda)e^{-\hat{T}_p}a^\dagger_p a_q e^{\hat{T}_p}|\Phi_0\rangle \tag{26}$$

and

$$\Gamma^{pq}_{rs} = \langle\Phi_0|(1+\Lambda)e^{-\hat{T}_p}a^\dagger_p a^\dagger_q a_s a_r e^{\hat{T}_p}|\Phi_0\rangle, \tag{27}$$

where $\Lambda = \sum_{ia}\lambda^i_a a^\dagger_i a^\dagger_{\bar{i}} a_{\bar{a}} a_a$ is a de-excitation operator with Lagrange multipliers λ^a_i obtained by solving eq. (5). Furthermore, due to the special structure of the wave function, the only non-zero elements are γ^p_p, Γ^{pq}_{pq}, and Γ^{qq}_{pp}. We should note that the response density matrices are not Hermitian and, in general, we have $\Gamma^{qq}_{pp} \neq \Gamma^{pp}_{qq}$. The deviation from Hermiticity of the response density matrices is an artefact of the truncation of the full cluster operator and disappears if the full cluster operator is taken in the Coupled Cluster ansatz. As the AP1roG method uses, however, a truncated cluster operator, we cannot exclude non-symmetric two-particle response density matrices. Furthermore, if the response density matrices are not symmetric and are thus not N-representable, the resulting eigenvalues of $\rho_{p,q}$ might result in

Katarzyna Boguslaw

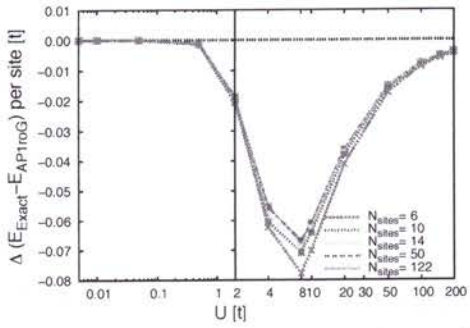


Figure 3: Deviation of the voo-AP1roG total energies from exact values for different strengths of the repulsive on-site interaction for the 1-D Hubbard model with periodic boundary conditions for $N_{sites} = 6, 10, 14, 50, 122$. The exact values for small U ($U < 0.001t$) for $N_{sites} = 50, 122$ could not be converged.

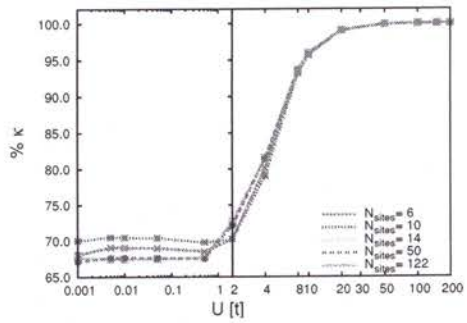


Figure 4: Percentage of the correlation energy $\%\kappa$ for different strengths of the repulsive on-site interaction in the half-filled 1-D Hubbard model with periodic boundary conditions for $N_{sites} = 6, 10, 14, 50, 122$ captured by voo-AP1roG. The exact values for small U ($U < 0.001t$) for $N_{sites} = 50, 122$ could not be converged.

negative values for orbital pair p, q. However, we haven't observed any problems with N-representability of the response density matrices if the orbital basis is optimized within the AP1roG method. Only minor N-representability issues have been observed when using canonical Hartree–Fock orbitals in the strong correlation regime with negative eigenvalues of order 10^{-3} or much smaller. Since negative eigenvalues are unphysical, we have discarded them when calculating the correlation functions.

c.4.2 Orbital-pair correlations in strongly-correlated systems

In the following, we will briefly discuss the correlations between the one-particle functions that are used to construct the geminals. Specifically, we will scrutinize how accurately the AP1roG model can reproduce orbital-pair correlations and orbital-entanglement in the one-dimensional Hubbard model with periodic boundary conditions,

$$\hat{H}_{\text{Hub}} = -t \sum_{\substack{j \\ \sigma\in\{\uparrow,\downarrow\}}} \left(a^\dagger_{(j+1)\sigma} a_{j\sigma} + a^\dagger_{j\sigma} a_{(j+1)\sigma} \right) + U \sum_j n_{j\uparrow} n_{j\downarrow}, \tag{28}$$

where the first term describes nearest-neighbor hopping, while the second term represents the repulsive on-site interaction. $n_{j\sigma} = a^\dagger_{j\sigma} a_{j\sigma}$ is the local number operator. It is well-known that the one-dimensional half-filled Hubbard model for $U = 0t$ is gapless, where all four local basis states ($|{-}\rangle$,$|{\uparrow}\rangle$,$|{\downarrow}\rangle$,$|{\uparrow\downarrow}\rangle$) have equal weights $\frac{1}{4}$ and hence the site entropy $s_i = \ln(4)$. For $U > 0t$, the charge gap opens and the weight of the unoccupied and doubly-filled basis states decrease. In the large $U/t \to \infty$ limit, only the $|{\uparrow}\rangle$ and $|{\downarrow}\rangle$ states have weights of 0.5 with $s_i = \ln(2)$ as the model becomes equivalent to the spin-$1/2$ Heisenberg model and the ground state is an antiferromag-

Katayun Boguslawski

netic state. Therefore, a wave function restricted to electron-pair states ($|—\rangle$ and $|\#\rangle$) cannot properly describe the correlations in both the large U/t limit and, to a smaller extent, for small U/t using the local on-site basis. To properly model such wave functions, we have to change the basis, which allows us to describe correlations of the one-dimensional half-filled Hubbard model with only unoccupied and doubly-filled basis states. Such a basis can be obtained self-consistently within the AP1roG method as described above. Note that the correlation and entanglement measures are basis dependent and thus the one-site(orbital) and two-orbital correlations within the optimized AP1roG basis will differ from those in the local on-site basis. In order to assess the accuracy of AP1roG in describing orbital-pair correlations of the one-dimensional half-filled Hubbard model, we performed DMRG calculations using the optimized AP1roG basis. As an example, we will only investigate orbital correlations in the one-dimensional Hubbard model with 14 sites. Additional numerical results can be found in **[H3]**.

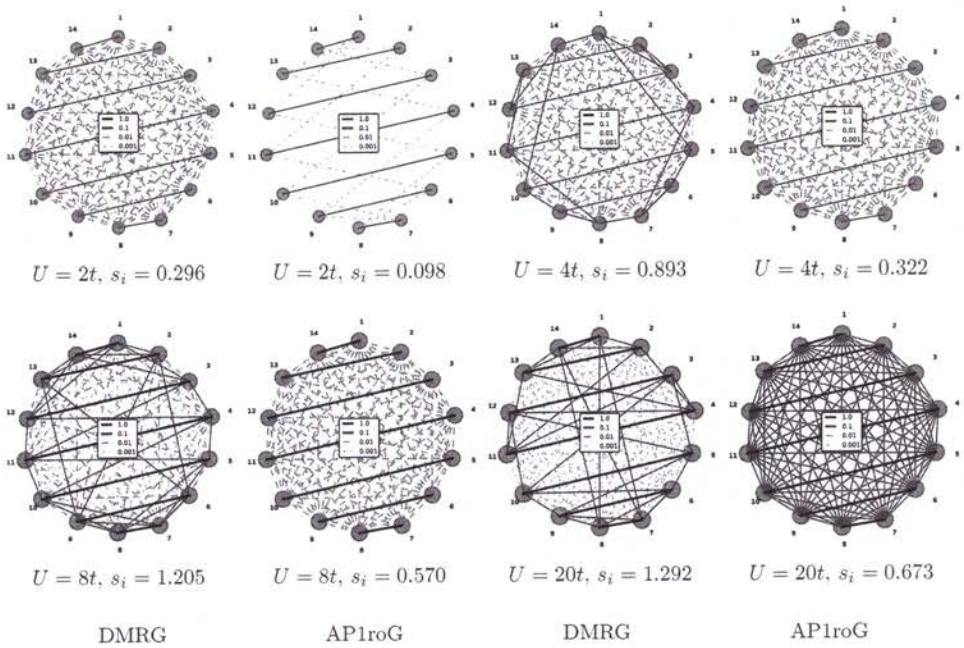
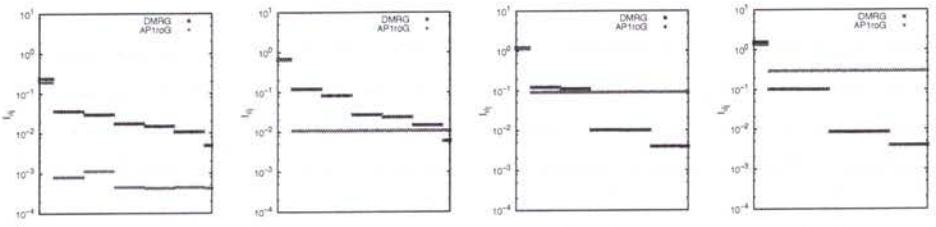


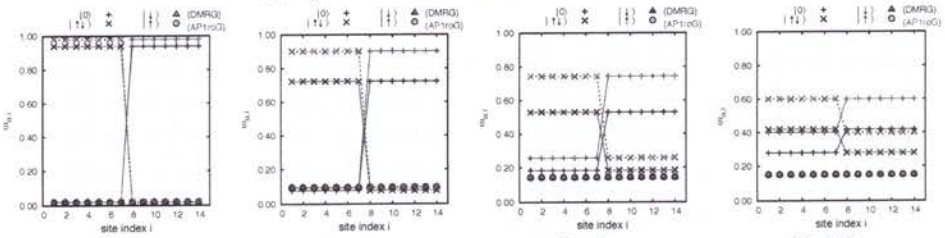
Figure 5: Orbital-pair mutual information for the half-filled 1-D Hubbard model with periodic boundary conditions, 14 sites, and different on-site interaction strengths for the optimized AP1roG basis. The single-orbital entropy is site-independent and given below each figure. The strength of the orbital-pair correlations for both the DMRG (left panel) and AP1roG (right panel) correlation diagrams are color-coded: black lines indicate strong correlations, while green lines indicate weak correlations.

Before scrutinizing orbital-pair correlations, we will first compare total and correlation energies obtained by the variational orbital optimized (voo-)AP1roG model with reference data obtained from the solution of the Lieb-Wu equations[118] (see Figure 3 for

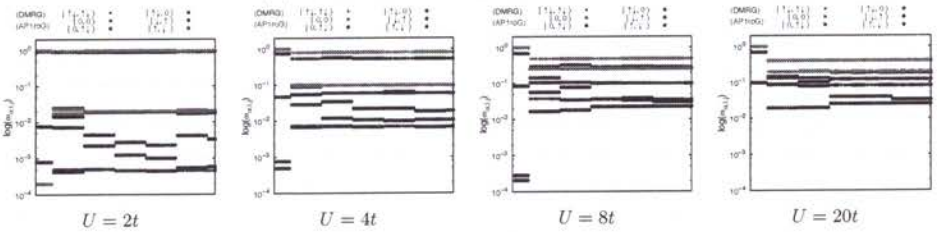
$N_{sites} = 6, 10, 14, 50, 122$) **[H8]**. voo-AP1roG can reproduce the exact total energies in the limit of zero and infinite (repulsive) on-site interaction. The largest deviations from the exact solution (up to $0.075t$ per site) are found for the intermediate region of the on-site interaction, that is, for $2t < U < 50t$. Figure 4 shows the percentage of the correlation energy captured by voo-AP1roG calculated as $\%\kappa = \frac{E^{\text{voo-AP1roG}} - E^{\text{HF}}}{E^{\text{exact}} - E^{\text{HF}}} \cdot 100$. In the limit of zero and infinite U, the voo-AP1roG model becomes exact; for $U = 0$ the wave function can be exactly described by a single Slater determinant and thus



(a) $I_{i|j}$ for both DMRG (blue) and AP1roG (red)



(b) $\omega_{\alpha;i}$ for both DMRG (blue) and AP1roG (red)



(c) $\omega_{\alpha;i,j}$ for both DMRG (blue) and AP1roG (red)

Figure 6: Decaying values of the mutual information for the half-filled Hubbard model with 14 sites using the optimized AP1roG orbital basis (a). $I_{i|j}$ is sorted with respect to the DMRG reference values so that each value of $I_{i|j}$ is shown for the same orbital pair i and j in both DMRG and AP1roG calculations. Eigenvalues of the (b) one-orbital reduced density matrix and (c) two-orbital reduced density matrix for the half-filled Hubbard model with 14 sites obtained in DMRG and AP1roG calculations using the optimized AP1roG orbital basis. The eigenvalues of $\rho_{i,j}$ for each pair i, j are ordered as in (a). Red lines and symbols indicate AP1roG data, while blue lines and symbols mark the corresponding DMRG results.

the correlation energy approaches zero, while for $U \to \infty$, the quantum state can be represented by the perfect pairing wave function. For growing (repulsive) U, the percentage of the correlation energy covered by voo-AP1roG increases gradually (see Figure 4). Featuring mean-field-like scaling, voo-AP1roG can recover about 71% of the correlation energy in the weak interaction regime, about 80% for intermediate interaction strengths, and approximately 93% in the case of strong on-site interaction for all chain lengths studied.

Figure 5 shows the orbital-pair mutual information and the single-orbital entropy obtained from DMRG (left panel) and AP1roG (right panel), respectively, for different strengths of U for the optimized AP1roG basis. For all investigated values of U, AP1roG can reproduce the most important orbital correlations (*cf.* the black/blue lines in Figure 5). Weaker orbital correlations ($I_{i|j} \leq 10^{-2}$) are, however, underestimated for small U/t if the wave functions is restricted to the seniority-zero sector. For increasing repulsive on-site interactions $U \geq 4t$, AP1roG gradually overestimates orbital-pair correlations compared to the DMRG reference distribution (*cf.* increasing number of red/blue lines). The observed overcorrelation of AP1roG for increasing U/t is also visible in the decaying values of $I_{i|j}$ shown in Figuree 6(a). Thus, restricting the wave function to the seniority-zero sector results in an overestimation of the medium-sized and weak orbital-pair correlations.

If the orbitals, and thus the reference determinant, are optimized, AP1roG can accurately describe the largest orbital-pair correlations in all investigated systems, misses, however, a large fraction of the weaker orbital-pair correlations. In the strong correlation limit, AP1roG considerably overestimates intermediate and weaker orbital-pair correlations and results in a prolonged plateau of the mutal information. This overcorrelation can be explained by the eigenvalue spectra of ρ_i (one-orbital reduced density matrix) and $\rho_{i,j}$ (two-orbital reduced density matrix), which are used to determine the orbital-based correlation functions. While in the weak correlation limit, the eigenvalues corresponding to singly-occupied states are (orders of magnitudes) smaller than those corresponding to doubly-occupied or empty states, their weights gradually increase when we approach the strong correlation regime. Specifically, in the strong correlation limit, singly-occupied states become important and need to be included in the wave function model to accurately describe the spectrum of ρ_i and $\rho_{i,j}$ even if the orbital basis is optimized (see Figure 6).

We should emphasize that if the one-particle functions are not optimized and the Hartree–Fock determinant is taken as reference determinant in the AP1roG ansatz, all orbital-pair correlations are smaller than the DMRG reference values. Thus, overcorrelation is introduced through the optimization of the AP1roG basis, which minimizes the total electronic energy, but simultaneously deteriorates orbital-pair correlations (see also **[H3]** for a complete picture). Finally, for molecular geometries around the equilibrium structure, both $I_{i|j}$ and the eigenvalue spectra of ρ_i and $\rho_{i,j}$ suggest that AP1roG provides accurate zero-order wave functions (with and without orbital optimization) where the missing orbital-pair correlations could be accurately modeled using *a posteriori* approaches for weak electron correlation (see again **[H3]** for a complete picture). To conclude, our orbital correlation analysis confirms that the accuracy and reliability of a quantum chemical model cannot be assessed by comparing electronic energies only (total, relative, or correlation energies). To remedy this

Katayun Boguslawski

issue, quantum entanglement and correlation measures can provide deeper insights into and understanding of electronic structures.

c.5 Capturing dynamic correlation beyond electron-pair states

As all other geminal-based models, AP1roG misses a large fraction of weak (dynamic) electron correlation effects. To address this problem and account for weak electron correlation effects in the geminal reference wave function, various *a posteriori* corrections have been proposed. These include models based on single- and multi-reference Perturbation Theory,[53,57,73,119,120] Extended Random Phase Approximation,[56,58] (Linearized) Coupled Cluster theory,[55,78] and Density Functional Theory.[121,122] In the case of AP1roG, dynamic correlation was included using Perturbation Theory,[119] single-reference Coupled-Cluster theory,[87] and Density Functional Theory.[121,122] In our work, we focused on two different corrections schemes: Dynamical correlation was included using (i) Perturbation Theory **[H1]** and (ii) a Linearized Coupled Cluster correction **[H4]**. In general, the dynamic energy corrections we propose allow us to approach chemical accuracy in many molecular systems. The resulting AP1roG-based methods outperform most conventional Coupled Cluster approximations, like CCSD, LCCSD, and BCCD, featuring a similar or lower computational scaling at the same time. In the following, we will briefly introduce the dynamic energy corrections we developed. All methods below, including numerical tests, are presented in **[H1,H4]**.

c.5.1 Second-order perturbation theory corrections

One drawback of multi-reference Perturbation Theory is the arbitrariness of the theoretical model. For example, there are different choices for the zero-order Hamiltonian $\hat{H}_0$, the dual state $\langle\tilde{\Psi}|$ in the projector, and the choice of the projection space.[123] Poor choices may lead to technical difficulties and unphysical solutions. Specifically for the AP1roG wave function, two different PT models have been proposed that allow us to describe electron correlation effects beyond electron pairs.[119] While these approaches provide reliable spectroscopic constants for some first-row diatomic molecules, their performance deteriorates when moving to heavy-element containing compounds like actinide species. To scrutinize the performance of a perturbation theory correction for molecules where both strong and weak correlation is important, we developed various perturbation theory models that feature different selections for the zero-order Hamiltonian $\hat{H}_0$, the dual state $\langle\tilde{\Psi}|$, and the projection space, while the zero-order wave function is restricted to the AP1roG reference function of eq. (1), $|\Psi^{(0)}\rangle = |\text{AP1roG}\rangle$. Furthermore, in the derivation of all PT models, we will use the quantum chemical Hamiltonian in its normal product form *shifted* by the correlation energy of AP1roG $E^{(0)}_{\text{corr}}$ (the shift in energy is indicated by "/"),

$$\begin{aligned}\hat{H}'_N &= \hat{H} - \langle\Phi_0|\hat{H}|\Phi_0\rangle - E^{(0)}_{\text{corr}} \\ &= (\hat{H}_0 - \langle\Phi_0|\hat{H}_0|\Phi_0\rangle) + (\hat{V} - \langle\Phi_0|\hat{V}|\Phi_0\rangle - E^{(0)}_{\text{corr}}) \\ &= (\hat{H}_0)_N + \hat{V}'_N, \end{aligned} \tag{29}$$

Katarzyna Boguslawski

where the quantum chemical Hamiltonian $\hat{H}$ is divided into a zero-order contribution $\hat{H}_0$ and a perturbation $\hat{V}$. It is convenient to rewrite $\hat{H}$ into a sum of a one- ($\hat{H}_1$) and a two-electron ($\hat{H}_2$, here indicated by $\hat{W}$) part,

$$\hat{H} = \hat{H}_1 + \hat{W} = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} \langle pq|rs\rangle a_p^\dagger a_q^\dagger a_s a_r. \tag{30}$$

In the above equation, h_{pq} and $\langle pq|rs\rangle$ are the one- and two-electron (written in physicists' notation) integrals, respectively, determined for the one-particle basis functions p, q, r, s. We should note that, in our work, we restrict $\hat{H}_0$ to be a one-body operator so that the normal-product form of the perturbation $\hat{V} = \hat{H} - \hat{H}_0$ can be written as an operator shifted by the AP1roG correlation energy, $\hat{V}'_N = \hat{V}_N - E^{(0)}_{\text{corr}}$ (again indicated by "′"), as we have $E^{(0)}_{\text{corr}} = \langle \Phi_0 | \hat{W}_N | \text{AP1roG}\rangle$. Introducing a shifted perturbation operator $\hat{V}'_N$ is equivalent to neglecting contractions (or diagrams) that correspond to the AP1roG correlation energy in the PT equations, which will be indicated by the $'$ in the sum of the $\hat{W}'_N$ operator.
As in conventional Rayleigh–Schrödinger PT, the exact wave function can be written as an order-by-order expansion, $|\Psi\rangle = |\text{AP1roG}\rangle + \lambda|\Psi^{(1)}\rangle + \ldots$, where λ is the order parameter. The first-order correction to the wave function is expanded in a set of Slater determinants $|\Phi_p\rangle$,

$$|\Psi^{(1)}\rangle = \sum_p t_p |\Phi_p\rangle, \tag{31}$$

and forced to be orthogonal to the zero-order wave function, here, $|\text{AP1roG}\rangle$,

$$\langle \Psi^{(1)} | \text{AP1roG}\rangle = 0. \tag{32}$$

This orthogonality constraint restricts the choice of the projection space used for the expansion of $|\Psi^{(1)}\rangle$ (and higher orders). By construction, all pair-excitations with respect to $|\Phi_0\rangle$ have to be excluded as they don't satisfy $\langle \Phi_p | \text{AP1roG}\rangle = 0$. Similarly, introducing an order parameter λ in the Hamiltonian $\hat{H} = \hat{H}_0 + \lambda\hat{V}$ and equating coefficients of powers of λ, we obtain the zero-, first-, and higher-order PT equations. Specifically for the first-order correction to the wave function, we have to solve

$$\begin{aligned} (\hat{H}_0)_N |\Psi^{(1)}\rangle + \hat{V}'_N |\text{AP1roG}\rangle &= 0 \\ \sum_p t_p (\hat{H}_0)_N |\Phi_p\rangle + \hat{V}'_N |\text{AP1roG}\rangle &= 0. \end{aligned} \tag{33}$$

Since we have introduced a shifted normal-product Hamiltonian, the zero- and first-order energy corrections vanish,

$$E^{(0)} + E^{(1)} = \frac{1}{\langle \tilde{\Psi} | \text{AP1roG}\rangle} \langle \tilde{\Psi} | \hat{V}'_N | \text{AP1roG}\rangle = 0, \tag{34}$$

where $\langle \tilde{\Psi}|$ is the dual of the unperturbed state $|\text{AP1roG}\rangle$. Specific choices for $\langle \tilde{\Psi}|$ will be considered below. The first non-zero correction to the energy is of second-order and

Katarzyna Boguslawski

can be calculated from the first-order wave function and the shifted normal-product perturbation Hamiltonian,

$$E^{(2)} = \frac{\langle\tilde{\Psi}|\hat{V}'_N|\Psi^{(1)}\rangle}{\langle\tilde{\Psi}|\mathrm{AP1roG}\rangle}. \tag{35}$$

Before we focus on possible choices of $(\hat{H}_0)_N$ and $\hat{V}'_N$ as well as $\langle\tilde{\Psi}|$, we will define our projection space used in the expansion of $|\Psi^{(1)}\rangle$ in eq. (31). Following previous PT models, the projection space will contain all possible excitations with respect to a reference determinant. This reference determinant is, however, not arbitrary, but restricted to the reference determined of AP1roG, $|\Phi_0\rangle$ of eq. (1). We should emphasize that $|\Phi_0\rangle$ is not equivalent to the Hartree–Fock determinant as in conventional CC theory, but adjusted during the optimization of the AP1roG wave function. Choosing $|\Phi_0\rangle$ as reference determinant, the first-order correction can be written as

$$|\Psi^{(1)}\rangle = \hat{T}|\Phi_0\rangle, \tag{36}$$

where $\hat{T}$ is some excitation operator that substitutes electrons from the occupied to the virtual space with respect to $|\Phi_0\rangle$. Furthermore, we will restrict $\hat{T}$ to contain double excitations without electron pairs, $\hat{T} = \hat{T}'_2$, as well as single and double excitations, $\hat{T} = \hat{T}_1 + \hat{T}'_2$. If only double excitations are included, the excitation operator is specified as

$$\hat{T}'_2 = \frac{1}{2}\sum_{ij}^{\mathrm{occ}}\sum_{ab}^{\mathrm{virt}}{}' t_{ij}^{ab}\hat{E}_{ai}\hat{E}_{bj}, \tag{37}$$

where $\hat{E}_{ai} = a_a^\dagger a_i + a_{\bar{a}}^\dagger a_{\bar{i}}$ is the singlet excitation operator and the perturbation amplitudes t_{ij}^{ab} are symmetric with respect to pair-exchange, *i.e.*, $t_{ij}^{ab} = t_{ji}^{ba}$. Similar to our notation of the shifted normal-product Hamiltonian, the prime in the above summation indicates that pair-excited determinants are excluded in the excitation operator, *i.e.*, $t_{i\bar{i}}^{a\bar{a}} = 0$. Exclusion of pair-excitations fulfils the orthogonality condition eq. (32). If the excitation operator contains both single and double excitations, the single excitations can be accounted for by adding

$$\hat{T}_1 = \sum_{i}^{\mathrm{occ}}\sum_{a}^{\mathrm{virt}} t_i^a \hat{E}_{ai} \tag{38}$$

to the double excitation operator. Choosing eq. (31) as the first-order correction to the wave function with excitation operators as defined in eqs. (37) and (38), the corresponding perturbation amplitudes are then obtained by solving

$$\sum_p t_p\langle\overline{\Phi_q}|(\hat{H}_0)_N|\Phi_p\rangle + \langle\overline{\Phi_q}|\hat{V}'_N|\mathrm{AP1roG}\rangle = 0. \tag{39}$$

Note that the above equations depend on the partition of the Hamiltonian $\hat{H}$ into the zero-order and the perturbation part as well as the projection manifold, but not on the choice of the dual state $\langle\tilde{\Psi}|$. In the following, we will consider different partitionings of $\hat{H}$ as well as two choices for $\langle\tilde{\Psi}|$.

Katarzyna Boguslaw

In all Perturbation Theory models, the $\hat{H}_0$ Hamiltonian was constructed from the Fock operator,

$$\hat{F} = \sum_{pq} \left(h_{pq} + \sum_{i}^{\text{occ}} (\langle pi||qi\rangle + \langle pi|qi\rangle) \right) a_p^\dagger a_q = \sum_{pq} f_{pq} a_p^\dagger a_q, \tag{40}$$

where $\langle pi||qi\rangle$ are the two-electron integrals in physicists' notation containing the Coulomb $\langle pi|qi\rangle$ and exchange $\langle pi|iq\rangle$ terms, and included either only the diagonal or both the diagonal and off-diagonal inactive Fock operator.

Choosing a diagonal $(\hat{H}_0)_N$ Hamiltonian. If $(\hat{H}_0)_N$ is restricted to the diagonal of the inactive Fock operator, the zero-order Hamiltonian reads

$$(\hat{H}_0)_N = \hat{F}_N^{\text{d}} = \sum_{p} f_{pp} \{a_p^\dagger a_p\}, \tag{41}$$

while the perturbation becomes (cf. eq. (29))

$$\hat{V}'_N = \hat{F}_N^{\text{o}} + \hat{W}'_N = \sum_{p \neq q} f_{pq} \{a_p^\dagger a_q\} + \frac{1}{2} {\sum_{pqrs}}' \langle pq|rs\rangle \{a_p^\dagger a_q^\dagger a_s a_r\}. \tag{42}$$

Eqs. (41) and (42) are then substituted into eq. (39) to solve for the PT amplitudes. Note that, in the case of a diagonal zero-order Hamiltonian, the PT amplitudes are obtained from a set of uncoupled equations.
Now, we will consider two different choices for the dual state. First, consider $\langle\tilde{\Psi}|$ to be restricted to the AP1roG reference determinant $\langle\Phi_0|$. In this case, we can straightforwardly evaluate the overlap $\langle\Phi_0|\text{AP1roG}\rangle$, which equals 1 due to intermediate normalization of the AP1roG wave function. The expression for the second-order energy correction $E^{(2)}$ given in eq. (35) thus simplifies to

$$E^{(2)} = \langle\Phi_0|\hat{V}'_N|\Psi^{(1)}\rangle. \tag{43}$$

Specifically, the energy correction for $\hat{T} = \hat{T}'_2$ is given as

$$E_{\text{d}}^{(2)} = \sum_{iajb} t_{ij}^{ab} (\langle ij||ab\rangle + \langle ij|ab\rangle), \tag{44}$$

where the sum runs over spatial orbitals only, while for single and double excitations we have (again using spatial orbitals)

$$E_{\text{sd}}^{(2)} = 2 \sum_{ia} f_{ia} t_i^a + \sum_{iajb} t_{ij}^{ab} (\langle ij||ab\rangle + \langle ij|ab\rangle). \tag{45}$$

We should emphasize that this PT model is equivalent to the PTa model presented in ref. 119. Note, however, that pair excitations are not excluded in ref. 119 and that the full Hamiltonian $\hat{H}$ is taken as the perturbation Hamiltonian so that $E^{(2)} = \langle\Phi_0|\hat{H}|\Psi^{(1)}\rangle$. Since the PTa amplitude equations for the pair excitations

Katarzyna Bystrek

vanish (as they equal the AP1roG amplitude equations), the PTa energy corrections is nonetheless determined from eq. (43). For reasons of consistency, we will abbreviate these PT models using a single determinant (SD) as dual and a diagonal (d) zero-order Hamiltonian as PT2SDd. The choice of the excitation operator will be indicated in parentheses, that is, PT2SDd(d) for double excitations and PT2SDd(sd) for both single and double excitations.
Our second choice for the dual state includes $\langle$AP1roG|. In this case, we have to evaluate terms as $\langle$AP1roG|AP1roG$\rangle$ in the energy expression, which becomes computationally intractable for large systems. In order to arrive at a computationally feasible model, we will follow ref. 119 to, at least partially, eliminate the overlap $\langle$AP1roG|AP1roG$\rangle$ in the PT equations and energy expression. For that purpose, we redefine the zero-order Hamiltonian of eq. (41) by introducing the inverse of the overlap $\langle$AP1roG|AP1roG$\rangle$ as a scaling factor,

$$(\hat{H}_0)_N = \bar{F}_N^{\mathrm{d}} = \sum_p \frac{f_{pp}}{\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle} \{a_p^\dagger a_p\} = \sum_p \bar{f}_{pp} \{a_p^\dagger a_p\}. \tag{46}$$

By changing the zero-order Hamiltonian of eq. (41), we also have to adjust the corresponding perturbation part given in eq. (42),

$$\hat{V}'_N = \hat{F}_N - \bar{F}_N^{\mathrm{d}} + \hat{W}'_N = \sum_{p,q} (f_{pq} - \bar{f}_{pp}\delta_{pq}) \{a_p^\dagger a_q\} + \frac{1}{2} {\sum_{pqrs}}' \langle pq|rs\rangle \{a_p^\dagger a_q^\dagger a_s a_r\}. \tag{47}$$

To fully avoid the evaluation of the overlap $\langle$AP1roG|AP1roG$\rangle$ in the PT amplitude equations eq. (39), the inverse of the AP1roG wave function overlap will be absorbed in the PT amplitudes. Thus, the first-order wave function contains scaled amplitudes,

$$|\overline{\Psi^{(1)}}\rangle = \sum_p \frac{t_p}{\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle} |\Phi_p\rangle = \sum_p \bar{t}_p |\Phi_p\rangle. \tag{48}$$

By substituting eq. (46) into eq. (39) and introducing the scaled PT amplitudes from eq. (48), we can eliminate the wave function overlap from the PT working equations. Furthermore, scaling the PT amplitudes by the inverse of the overlap $\langle$AP1roG|AP1roG$\rangle$ restores the zero-order Hamiltonian of eq. (41) and we get $(\hat{H}_0)_N = \hat{F}_N^{\mathrm{d}}$. Note that the wave function overlap is still present in the perturbation part $\hat{V}'_N$. Due to the structure of the zero-order wave function and the choice of the projection manifold, the diagonal part of the modified Fock operator in eq. (47) does not contribute to the PT amplitude equations and the resulting perturbation reduces to $\hat{V}'_N = \hat{F}_N^{\mathrm{o}} + \hat{W}'_N$. Since we have chosen $\langle$AP1roG| as dual, the second-order energy correction is determined from $E^{(2)} = \langle \mathrm{AP1roG}|\hat{V}'_N|\overline{\Psi^{(1)}}\rangle$ (cf. eq. (35)), where the first-order correction to the wave function is calculated from the scaled PT amplitudes (eq. (48)),

$$E^{(2)} = \sum_p \bar{t}_p \langle \mathrm{AP1roG}|\hat{V}'_N|\Phi_p\rangle. \tag{49}$$

The sum in the above equation runs over all determinants in the projection manifold

(doubly excited or singly- and doubly-excited determinants). We should note that although we can exactly evaluate the energy correction and the PT equations, the zero- and first-order energy corrections eq. (34) do not vanish. However, we can neglect the weights of the (AP1roG) amplitudes beyond single pair excitations and assume that $E^{(0)} + E^{(1)} \approx 0$.[119] We will label this PT model as PT2MDd as it uses a multi-determinant (MD) wave function as dual and a diagonal (d) zero-order Hamiltonian. Furthermore, PT2MDd(d) indicates that the excitation operator contains only double excitations (without pairs), while in PT2MDd(sd) both single and double excitations are included in $\hat{T}$.

An Off-diagonal one-body zero-order Hamiltonian. If the off-diagonal terms of the Fock operator are included in the zero-order Hamiltonian of eq. (41), we arrive at

$$(\hat{H}_0)_N = \hat{F}_N^{\mathrm{d}} + \hat{F}_N^{\mathrm{o}} = \sum_{p,q} f_{pq} \{a_p^\dagger a_q\}. \tag{50}$$

Then, the perturbation part $\hat{V} = \hat{H} - \hat{H}_0$ in its (shifted) normal-product form contains only the two-electron part,

$$\hat{V}_N' = \frac{1}{2} \sum_{pqrs}{}' \langle pq|rs\rangle \{a_p^\dagger a_q^\dagger a_s a_r\}. \tag{51}$$

To obtain the working equation for the PT amplitudes, we substitute eqs. (50) and (51) into eq. (39). Note that, in contrast to the PT methods with a diagonal zero-order Hamiltonian, the PT amplitudes are now obtained from a set of coupled equations and have to be solved iteratively.
If the dual state is restricted to the reference determinant of the AP1roG wave function, we benefit from the intermediate normalization when evaluating the overlap $\langle \Phi_0 | \mathrm{AP1roG} \rangle$. Analogous to PT2SDd-type methods, the second order energy can be evaluated from eq. (43). Note, however, that only the doubly-excited determinants directly contribute to the energy correction. Since the perturbation Hamiltonian is a two-electron operator, the second-order energy correction of the single excitations vanishes. Single excitations contribute indirectly through coupling to the double excitation manifold in the PT amplitude equations. For both including and excluding the singles projection manifold, the second-order energy $E^{(2)}$ is thus calculated from eq. (44). We will abbreviate the PT corrections using an off-diagonal (o) zero-order Hamiltonian and a single determinant for its dual state as PT2SDo, while the projection manifold will be indicated in parentheses (d for doubles, sd for singles and doubles, respectively).
Similar to the PT methods with a diagonal $\hat{H}_0$ Hamiltonian, choosing $\langle \mathrm{AP1roG}|$ as dual forces us to evaluate the wave function overlap $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle$, which becomes prohibitive for large systems. In order to (partially) remove the wave function overlap from the working equations, we follow the procedure from above and introduce a scaled zero-order Hamiltonian, where we have to modify both the diagonal

F_N^{d} and off-diagonal F_N^{o} Fock operator,

$$\begin{aligned}(\hat{H}_0)_N &= \bar{F}_N^{\mathrm{d}} + \bar{F}_N^{\mathrm{o}} \\ &= \sum_{p,q} \frac{f_{pq}}{\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle} \{a_p^\dagger a_q\} = \sum_{p,q} \bar{f}_{pq} \{a_p^\dagger a_q\}. \end{aligned} \tag{52}$$

Since we use a modified Fock operator as zero-order Hamiltonian, we have to account for it in the definition of the perturbation part,

$$\begin{aligned}\hat{V}'_N &= \hat{F}_N - \bar{F}_N^{\mathrm{d}} - \bar{F}_N^{\mathrm{o}} + \hat{W}'_N \\ &= \sum_{p,q} (f_{pq} - \bar{f}_{pq}) \{a_p^\dagger a_q\} + \frac{1}{2} \sum_{pqrs}{}' \langle pq|rs\rangle \{a_p^\dagger a_q^\dagger a_s a_r\}. \end{aligned} \tag{53}$$

Analogous to PT2MDd-type methods, the scaling factor in the $(\hat{H}_0)_N$ Hamiltonian is absorbed in the PT amplitudes (cf. eq. (48)) so that the zero-order Hamiltonian in the PT amplitude equations eq. (39) contains only the unscaled Fock operator, $(\hat{H}_0)_N = \hat{F}_N^{\mathrm{d}} + \hat{F}_N^{\mathrm{o}}$, while the PT amplitudes t_p are replaced by the scaled amplitudes $\bar{t}_p$. Although we eliminated the wave function overlap in the second-order energy expression and in the $(\hat{H}_0)_N$ part of the PT amplitude equations, $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle$ still remains in the perturbation part (cf. eq. (53)). In contrast to the PT2MDd-type methods discussed above, the perturbation $\hat{V}'_N$ contains also modified off-diagonal elements in the one-electron Fock operator that do not vanish in the PT equations. Therefore, we have to evaluate the wave function overlap $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle$ before we can determine the PT amplitudes. To obtain a computationally feasible model, we will approximate the overlap $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle \approx 1 + \sum_{ia} |c_i^a|^2$, keeping only the quadratic terms in the AP1roG amplitudes. This is usually a good approximation around the equilibrium as the AP1roG wave function amplitudes are typically much smaller than 1 ($|c_i^a| \ll 1$). In the vicinity of dissociation, however, this approximation might be inappropriate as some of the AP1roG amplitudes are close to 1 in absolute value. In such cases, higher order terms have to be included to approximate the wave function overlap $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle$ appropriately. To keep the PT2 correction computationally inexpensive, we can assume that $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle \to \infty$ and use an approximate perturbation Hamiltonian where we disregard the contribution of $\bar{F}_N$. This will result in the PTb model.[119] However, we approximate the scaled off-diagonal Fock matrix elements by $\bar{f}_{pq} \approx \frac{f_{pq}}{1+\sum_{ia} |c_i^a|^2}$, neglecting all higher-order terms. As for PT2MDd, the second-order energy $E^{(2)}$ can be determined from eq. (49) with $\hat{V}'_N$ defined in eq. (53). Note that in contrast to PT2SDo, the single excitations directly contribute to the energy correction through both the one- and two-electron operators in the perturbation Hamiltonian. The PT models with an off-diagonal (o) $\hat{H}_0$ Hamiltonian and a multi-determinant (MD) wave function ($\langle \mathrm{AP1roG}|$) as dual will be abbreviated as PT2MDo, while the projection manifold will be again indicated in parentheses (d for doubles, sd for singles and doubles, respectively). Note that pair-excitations are excluded in the projection manifold due to the orthogonality constraint.

We should emphasize that the PT2MDo model is similar, but not equivalent, to the

Katarzyna Boguslawska

recently presented PTb approach.[119] In the PTb method, the wave function overlap in the perturbation Hamiltonian is neglected, that is, we assume $\frac{1}{\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle} \to 0$ so that $\bar{f}_{pq} \to 0$. By neglecting the scaled components of the Fock operator, the perturbation Hamiltonian reduces to the full quantum chemical Hamiltonian $\hat{V}'_N = \hat{H}'_N$. The second-order energy correction is then given as

$$E^{(2)} = \sum_p \bar{t}_p \langle \mathrm{AP1roG}|\hat{H}'_N|\Phi_p\rangle. \tag{54}$$

As discussed above, this approximation is appropriate in the vicinity of dissociation, where the AP1roG wave function amplitudes are close to 1. Around equilibrium, however, we typically have $\langle \mathrm{AP1roG}|\mathrm{AP1roG}\rangle \approx 1$ and therefore neglecting $\bar{F}_N$ in the perturbation Hamiltonian represents a crude approximation.
Furthermore, in PTb theory, pair excitations are not excluded in the projection manifold and the first-order correction of the wave function contains all double excitations. These pair excitations do not contribute to the energy corrections eq. (54). Their contribution vanishes as the corresponding terms in eq. (54) equal the AP1roG amplitude equations. However, pair excitations indirectly enter the energy correction by coupling to the remaining PT amplitudes in the PT equations as well as to the pair excitations of the AP1roG model. Furthermore, we have extended the original PTb model by including also single excitations in the projection manifold. We also investigate how the pair excitations in the projection manifold influence the PTb energy correction. For that purpose we have excluded the pair excitations in the projection manifold when optimizing the PTb amplitudes. To emphasize the order of the energy correction in PTb, these PT models are abbreviated as PT2b, while the projection manifold is indicated in parentheses (d for doubles, sd for singles and doubles, d\p for doubles without pairs, sd\p for singles and doubles excluding pair excitations).
The computational scaling of all PT corrections discussed above is determined by the first term of eq. (39). Note that we can introduce suitable intermediates for the second term in eq. (39) so that summations are performed only once in the beginning of the calculation. If $(\hat{H}_0)_N$ is restricted to be a diagonal one-body operator, the computational cost of the corresponding PT models scales as $\mathcal{O}(o^2v^2)$, where o is the number of occupied and v the number of virtual orbitals, respectively. Choosing an off-diagonal one-body zero-order Hamiltonian $(\hat{H}_0)_N$ increases the computational cost to $\mathcal{O}(o^2v^3)$ (see also Table 3). Since we now have to solve a coupled set of linear equations iteratively, we have to consider an additional prefactor. However, this prefactor is typically much smaller than v. Thus, PT2SDd and PT2MDd scale similar to AP1roG (or conventional electronic structure methods like MP2), while in PT2SDo, PT2MDo, and PT2b the computational cost increases by a factor of v. All PT approaches presented in this series of publications are summarized in Table 3 for comparison.

c.5.2 Linearized coupled cluster theory with an AP1roG reference function

A reliable way to account for dynamic correlation effects *a posteriori* is to use a multi-reference Linearized Coupled Cluster (LCC) correction. Recently, Zoboki *et al.*

Katerina Boguslawski

Table 3: Summary of PT models with zero-order Hamiltonian $\hat{H}_0$, perturbation $\hat{V}$, dual $\langle\tilde{\Psi}|$, and excitation operator $\hat{T}$. All operators are defined in the text. The computational scaling is given in the last column.

Model	$\hat{H}_0$	$\hat{V}$	$\langle\tilde{\Psi}|$	$\hat{T}$	scaling
PT2SDd	$\hat{F}_N^{\mathrm{d}}$	$\hat{F}_N^{\mathrm{o}} + \hat{W}'_N$	$\langle\Phi_0|$	$\hat{T}'_2$, $\hat{T}_1 + \hat{T}'_2$	$\mathcal{O}(o^2v^2)$
PT2MDd	$\hat{F}_N^{\mathrm{d}}$	$\hat{F}_N^{\mathrm{o}} + \hat{W}'_N$	$\langle\mathrm{AP1roG}|$	$\hat{T}'_2$, $\hat{T}_1 + \hat{T}'_2$	$\mathcal{O}(o^2v^2)$
PT2SDo	$\hat{F}_N^{\mathrm{d}} + \hat{F}_N^{\mathrm{o}}$	$\hat{W}'_N$	$\langle\Phi_0|$	$\hat{T}'_2$, $\hat{T}_1 + \hat{T}'_2$	$\mathcal{O}(o^2v^3)$
PT2MDo	$\hat{F}_N^{\mathrm{d}} + \hat{F}_N^{\mathrm{o}}$	$\hat{F}_N - \bar{F}_N + \hat{W}'_N$	$\langle\mathrm{AP1roG}|$	$\hat{T}'_2$, $\hat{T}_1 + \hat{T}'_2$	$\mathcal{O}(o^2v^3)$
PT2b	$\hat{F}_N^{\mathrm{d}} + \hat{F}_N^{\mathrm{o}}$	$\hat{H}'_N$	$\langle\mathrm{AP1roG}|$	$\hat{T}_2$, $\hat{T}'_2$, $\hat{T}_1 + \hat{T}_2$, $\hat{T}_1 + \hat{T}'_2$	$\mathcal{O}(o^2v^3)$

presented an LCC correction based on an APSG reference function and demonstrated the good performance of the APSG-MRLCC approach.[55] Their findings encouraged us to develop an LCC correction based on an AP1roG reference state. For an LCC correction, dynamic correlation effects are built in the electronic wave function *a posteriori* using an exponential Coupled Cluster ansatz,

$$|\Psi\rangle = \exp(\hat{T})|\mathrm{AP1roG}\rangle, \tag{55}$$

where $\hat{T} = \sum_\nu t_\nu \hat{\tau}_\nu$ is a general cluster operator. The corresponding time-independent Schrödinger equation reads

$$\hat{H}\exp(\hat{T})|\mathrm{AP1roG}\rangle = E\exp(\hat{T})|\mathrm{AP1roG}\rangle. \tag{56}$$

Multiplying from the left by $\exp(-\hat{T})$ and truncating the Baker–Campbell–Hausdorff expansion after the second term,

$$\exp(-\hat{T})\hat{H}\exp(\hat{T}) \approx \hat{H} + [\hat{H}, \hat{T}], \tag{57}$$

we arrive at the Linearized Coupled Cluster Schrödinger equation

$$(\hat{H} + [\hat{H}, \hat{T}])|\mathrm{AP1roG}\rangle = E|\mathrm{AP1roG}\rangle. \tag{58}$$

To obtain the cluster amplitudes t_ν, we multiply from left by $\langle\nu|$

$$\langle\nu|(\hat{H} + [\hat{H}, \hat{T}])|\mathrm{AP1roG}\rangle = 0, \tag{59}$$

where we assume that the excitation operator $\hat{\tau}_\nu$ creates states orthogonal to $|\mathrm{AP1roG}\rangle$, $\langle\nu|\mathrm{AP1roG}\rangle = 0$. The projection manifold $\{\nu\}$ will depend on the choice of the cluster operator $\hat{T}$ (*vide infra*).
The energy can be calculated by projecting against the reference determinant of $|\mathrm{AP1roG}\rangle$, *i.e.*, multiplying eq. (58) by $\langle\Phi_0|$ and using intermediate normalization,

$$\langle\Phi_0|(\hat{H} + [\hat{H}, \hat{T}])|\mathrm{AP1roG}\rangle = E. \tag{60}$$

The only constraint on the cluster operator we have made so far is that it creates states that are orthogonal to the AP1roG reference function. One suitable choice

Katarzyna Boguslawska

for the cluster operator is to include substitutions between the occupied and virtual orbitals with respect to $|\text{AP1roG}\rangle$. If only double excitations are included, the cluster operator as defined in eq. (37) fulfills the orthogonality condition, $\langle\nu|\text{AP1roG}\rangle = 0$. To arrive at a computationally feasible model, we will further restrict the cluster operator of eq. (37) to include only excitations with respect to the reference determinant, thereby excluding possible redundancies in excitations and amplitudes. The projection manifold then contains all doubly-excited determinants with respect to $|\Phi_0\rangle$. The doubles amplitudes $\{t_{ij}^{ab}\}$ are obtained by solving a linear set of equations

$$B_\mu + \sum_\nu A_{\mu,\nu} t_\nu = 0, \tag{61}$$

where the sum runs over all double excitations (without pair excitations) and $B_\mu = B_{iajb} = \langle\overline{{}_{ij}^{ab}}|\hat{H}|\text{AP1roG}\rangle$, while $A_{\mu,\nu} = A_{iajb,kcld} = \frac{1}{2}\langle\overline{{}_{ij}^{ab}}|[\hat{H}, \hat{E}_{ck}\hat{E}_{dl}]|\text{AP1roG}\rangle$. Note that the bar over the projection manifold indicates that the final working equations will be spin-summed. The energy correction $E_{\text{corr}}^{(D)}$ with respect to the AP1roG reference wave function can be evaluated from eq. (44), where the coupled cluster doubles amplitudes have to be substituted for $\{t_{ij}^{ab}\}$.
Similar to our PT models, the contribution of single excitations can be accounted for by including eq. (38) in the cluster operator. Thus, the singles projection manifold contains all singly-excited determinants with respect to $|\Phi_0\rangle$. The single and double amplitudes are obtained by solving a coupled set of linear equations equivalent to eq. (61) where μ and ν now run over all single and double excitations. The energy correction with respect to the AP1roG reference value can be evaluated using eq. (45). Note that, in contrast to canonical Hartree–Fock orbitals, the Fock matrix is not diagonal when the orbitals are optimized within AP1roG. In the AP1roG-LCC approach, the single excitations thus contribute both directly to the energy correction and indirectly through coupling to the doubles equations. We will abbreviate the LCC correction using $\hat{T} = \hat{T}_2'$ as AP1roG-LCCD, while AP1roG-LCCSD indicates that the cluster operator contains single and double excitations, $\hat{T} = \hat{T}_1 + \hat{T}_2'$.
Finally, we should note that the LCCD and LCCSD corrections as outlined above are similar, but not equivalent to the frozen-pCCD (fpCCD) and fpCCSD approaches, respectively.[87] Specifically, in fpCC, first the equations for the pair amplitudes are solved, which in our case is equivalent to solving the AP1roG amplitude equations. Then, the usual CCD/CCSD equations are solved without allowing the pair amplitudes to change. In AP1roG-LCC, we first solve for the AP1roG amplitudes, which is equivalent to the first step of a fpCC calculation, followed by solving for the remaining cluster amplitudes. In contrast to fpCC, our cluster operator is linearized (*cf.*, eq. (58)) and thus all higher-order terms are eliminated, while the reference wave function is $|\text{AP1roG}\rangle$ (instead of a single Slater determinant as in fpCC). Choosing $|\text{AP1roG}\rangle$ as a reference function results in additional terms in the amplitude equations beyond the standard single-reference LCC approach arising from coupling to pair-excited Slater determinants with respect to $|\Phi_0\rangle$. This coupling to pair-excited Slater determinants leads to additional terms in the LCC amplitude equations that are also included in the fpCC amplitude equations. Therefore, our LCC approach can be considered as a simplification of the fpCC method. Specifically, the connection

between our LCCD corrections and fpCC can also be understood by rewriting the corresponding Linearized Coupled Cluster Schrödinger equation (58) using explicitly the exponential ansatz for AP1roG,

$$\exp(-\hat{T}_\mathrm{p})(\hat{H} + [\hat{H}, \hat{T}_\mathrm{np}]) \exp(\hat{T}_\mathrm{p})|\Phi_0\rangle = E|\Phi_0\rangle, \tag{62}$$

where $\hat{T}_\mathrm{p}$ is the AP1roG (seniority zero) cluster operator containing only pair excitations and $\hat{T}_\mathrm{np}$ is the seniority non-zero cluster operator of eq. (37). Note that we have used the labels p (pair) and np (non-pair) to emphasize the connection to fpCC methods. The seniority non-zero cluster amplitudes can be obtained by projection against $\langle\nu|$,

$$\langle\nu|\hat{H} + [\hat{H}, \hat{T}_\mathrm{np}] + [[\hat{H}, \hat{T}_\mathrm{np}], \hat{T}_\mathrm{p}]|\Phi_0\rangle = 0, \tag{63}$$

which further illustrates the coupling to pair-excited Slater determinants generated by $\hat{T}_\mathrm{p}$ (compared to single-reference Linearized Coupled Cluster). The above equation can be compared to the amplitude equations of (fp-)CC approaches in, for instance, Refs. 83,124. Since the most expensive contributions in the amplitude equations are similar in both AP1roG-LCCD/LCCSD and fp-CCD/CCSD, the computational scaling of our LCC correction is as $\mathcal{O}(o^2v^4)$, where o and v are the number of occupied and virtual orbitals, respectively.

c.5.3 Performance of dynamic energy corrections

We have benchmarked the above mentioned models against spectroscopic constants for multiply bonded diatomics (F_2, C_2, N_2, BN, CN^+, and BN) **[H1,H4]** and thermochemical data of 15 reactions containing main-group elements extrapolated to the basis set limit **[H1]**. Most importantly, combining AP1roG with the investigated corrections allows us to reach chemical accuracy in most of the studied systems.

For all investigated diatomics (see Table 4 for C_2 and CO), PT2SDd and PT2SDo provide equilibrium bond lengths and vibrational frequencies that agree well with the corresponding reference values. Furthermore, addition of single excitations in the excitation operator $\hat{T}$ does not increase the accuracy of PT2SDd(d) or PT2SDo(d). In contrast to PT2SDd and PT2SDo, the performance of PT models using a multi-determinant dual state (PT2MDd, PT2MDo, and PT2b) is less satisfying. Specifically, these methods provide equilibrium bond distances and vibrational frequencies that typically deviate most from MRCI-SD/MRCI-SD+Q reference data. Similar to PT2SDd and PT2SDo, addition of single excitations generally worsens spectroscopic constants and hence the excitation operator should be restricted to double excitations only. Furthermore, including pair excitations in $\hat{T}$ (PT2b-type methods) does not significantly affect spectroscopic constants and both models (with and without pair excitations) yield similar values for r_e and ω_e.

In contrast to equilibrium bond distances and vibrational frequencies, the accurate prediction of dissociation energies D_e is more challenging. In general, none of the proposed PT models can reliably predict potential energy well depths, which differ up to 60 kcal/mol from MRCI-SD/MRCI-SD+Q reference data. Note that most PT corrections (PT2MD and PT2b) diverge in the vicinity of dissociation and hence only an estimated dissociation energy is given in Table 4 (indicated by the $*$ in the

Katarzyna Bąkowicz

Table). Furthermore, the choice of the approximate AP1roG wave function overlap in PT2MDo/PT2b-type methods does not affect the accuracy in spectroscopic constants. Both PT approaches predict similar equilibrium bond lengths and vibrational frequencies as well as (estimated) dissociation energies D_e. Thus, including higher-order terms when calculating $\langle$AP1roG|AP1roG$\rangle$ might not cure the observed divergencies in the dissociation limit.
The performance of the LCCD/LCCSD correction on top of AP1roG is more robust in

Table 4: Spectroscopic constants for the dissociation of the homonuclear dimer C_2 and the heteronuclear dimer CO for different quantum chemistry methods. While for CO the differences are with respect to MRCI-SD+Q reference data,[125] for C_2, differences are with respect to MRCI-SD reference data.[126] E_e: ground state energy at r_e.

	Method	E_e [E_h]	r_e [Å]	D_e [$\frac{\text{kcal}}{\text{mol}}$]	ω_e [cm^{-1}]
C_2	AP1roG	−75.58569	1.227(−0.025)	132.9(−8.2)	1780(−56)
	AP1roG-PT2SDd(d)	−75.80222	1.251(+0.001)	160.4(+19.3)	1889(+53)
	AP1roG-PT2SDd(sd)	−75.81041	1.249(−0.003)	154.6(+13.5)	1915(+79)
	AP1roG-PT2MDd(d)	−75.77832	1.239(−0.013)	121.6(−19.5)*	1940(+104)
	AP1roG-PT2MDd(sd)	−75.78630	1.238(−0.014)	115.6(−25.5)*	1963(+127)
	AP1roG-PT2SDo(d)	−75.81778	1.242(−0.010)	156.5(+15.4)	1919(+83)
	AP1roG-PT2SDo(sd)	−75.81783	1.242(−0.010)	156.4(+15.3)	1919(+83)
	AP1roG-PT2MDo(d)	−75.79032	1.231(−0.021)	125.9(−15.2)*	2010(+174)
	AP1roG-PT2MDo(sd)	−75.79108	1.230(−0.022)	119.9(−21.2)*	2019(+183)
	AP1roG-PT2b(d)	−75.78350	1.235(−0.017)	127.2(−13.9)*	1938(+102)
	AP1roG-PT2b(sd)	−75.79400	1.228(−0.024)	113.0(−28.1)*	2049(+213)
	AP1roG-PT2b(d\p)	−75.78381	1.231(−0.021)	123.9(−17.2)*	2016(+180)
	AP1roG-PT2b(sd\p)	−75.79370	1.228(−0.024)	112.9(−28.2)*	2048(+212)
	AP1roG-LCCD	−75.81125	1.240(−0.012)	139.3(−1.8)	1916(+80)
	AP1roG-LCCSD	−75.81257	1.240(−0.012)	143.0(+1.9)	1926(+90)
	NEVPT2	−75.78829	1.244(−0.008)	148.0(+6.9)	1886(+50)
	LCCD	−75.82665	1.196(−0.056)	257.1(+116.0)*	2302(+466)
	LCCSD	−75.89995	1.194(−0.058)	293.0(+151.9)*	2710(+874)
	CCSD	−75.77969	1.325(+0.073)	106.0(−35.1)*	1183(−653)
	CR-CCSD(T)	−75.80395	1.242(−0.010)	152.1(+9.0)	1989(+153)
	CR-CC(2,3)	−75.81278	1.253(+0.001)	131.2(−9.9)	1842(+6)
	MRCI-SD[126]	−75.78079	1.252	141.1	1836
CO	AP1roG	−112.91110	1.116(−0.020)	240.2(−11.6)	2197(+49)
	AP1roG-PT2SDd(d)	−113.16024	1.129(−0.007)	261.2(+9.4)	2028(−120)
	AP1roG-PT2SDd(sd)	−113.16167	1.133(−0.003)	227.8(−24.0)	2128(−20)
	AP1roG-PT2MDd(d)	−113.15041	1.126(−0.010)	236.6(−15.2)	2186(+38)
	AP1roG-PT2MDd(sd)	−113.15196	1.126(−0.010)	190.7(−61.1)*	2204(+56)
	AP1roG-PT2SDo(d)	−113.17603	1.135(+0.001)	253.6(+1.8)	2023(−125)
	AP1roG-PT2SDo(sd)	−113.17606	1.132(−0.004)	—	2070(−78)
	AP1roG-PT2MDo(d)	−113.16343	1.130(−0.006)	242.5(−9.3)	2180(+32)
	AP1roG-PT2MDo(sd)	−113.16428	1.130(−0.006)	196.2(−55.6)*	2196(+48)
	AP1roG-PT2b(d)	−113.15512	1.129(−0.007)	240.6(−11.2)	2210(+62)
	AP1roG-PT2b(sd)	−113.15718	1.130(−0.006)	184.7(−67.1)*	2188(+40)
	AP1roG-PT2b(d\p)	−113.15425	1.129(−0.007)	240.1(−11.7)	2196(+48)
	AP1roG-PT2b(sd\p)	−113.15638	1.130(−0.006)	184.3(−67.5)*	2188(+40)
	AP1roG-LCCD	−113.17264	1.128(−0.008)	245.4(−6.4)	2265(+117)
	AP1roG-LCCSD	−113.17490	1.130(−0.006)	252.5(+0.7)	2182(+36)
	CASSCF[127]	−112.91158	1.136(+0.000)	251.1(−0.7)	2166(+18)
	LCCD	−113.16889	1.130(−0.006)	—	1954(−194)
	LCCSD	−113.17802	1.133(−0.003)	—	2141(−7)
	CCSD	−113.16293	1.129(−0.007)	330.8(+79.1)	2219(+71)
	CR-CC(2,3)	−113.18114	1.137(+0.001)	332.7(+80.9)	2130(−18)
	MRCI-SD+Q[127]	−113.15580	1.136	251.8	2148

* Estimated dissociation energy.

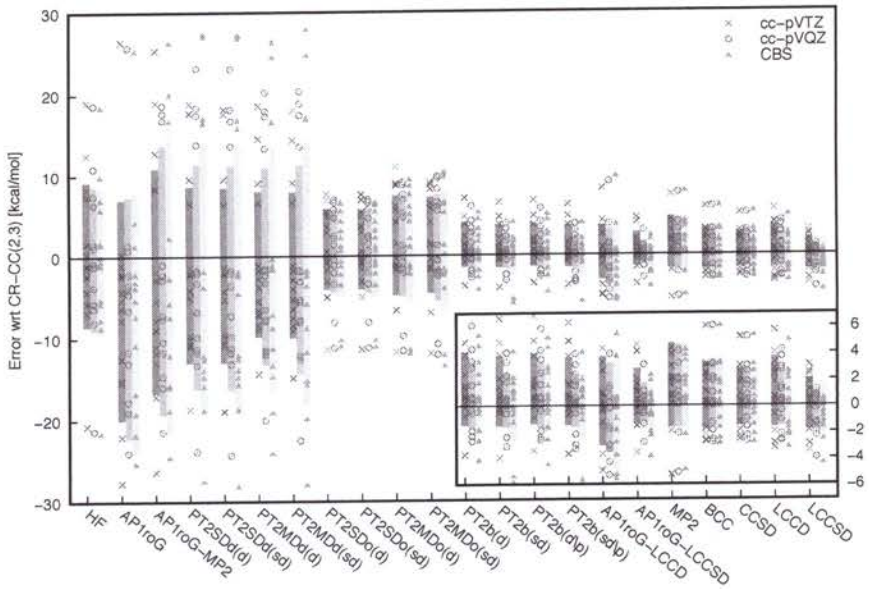
Katayun Boguslawski

predicting spectroscopic constants for all investigated diatomics. Although AP1roG-LCCSD yields r_e and ω_e that deviate more from MRCI-SD/MRCI-SD+Q reference values compared to PT2SDd and PT2SDo results, an LCC correction allows us to reliably model dissociation energies for most of the multiply bonded diatomics (C_2, N_2, CN^+, and CO). Finally, we should emphasize that AP1roG-LCCSD outperforms all investigated CC approaches (CCSD, LCCD, LCCSD, CR-CC(2,3)) in predicting dissociation energies for multiply bonded systems compared to MRCI-SD/MRCI-SD+Q reference values. Although being a simplified version of CCSD, AP1roG-LCCSD significantly reduces the errors of CCSD in dissociation energies (from 35 to 2 kcal/mol for C_2, from 29 to 6 kcal/mol for F_2, from 35 to 6 kcal/mol for N_2, from 79 to 1 kcal/mol for CO, from 88 to 24 kcal/mol for BN; we encountered computational difficulties for CCSD for the CN^+ molecule, while AP1roG-LCCSD results in an error of 2 kcal/mol). Thus, AP1roG-LCCSD provides improved dissociation energies compared to conventional CC approaches, while the computational cost remains similar to CCSD (neglecting the orbital-optimization step).

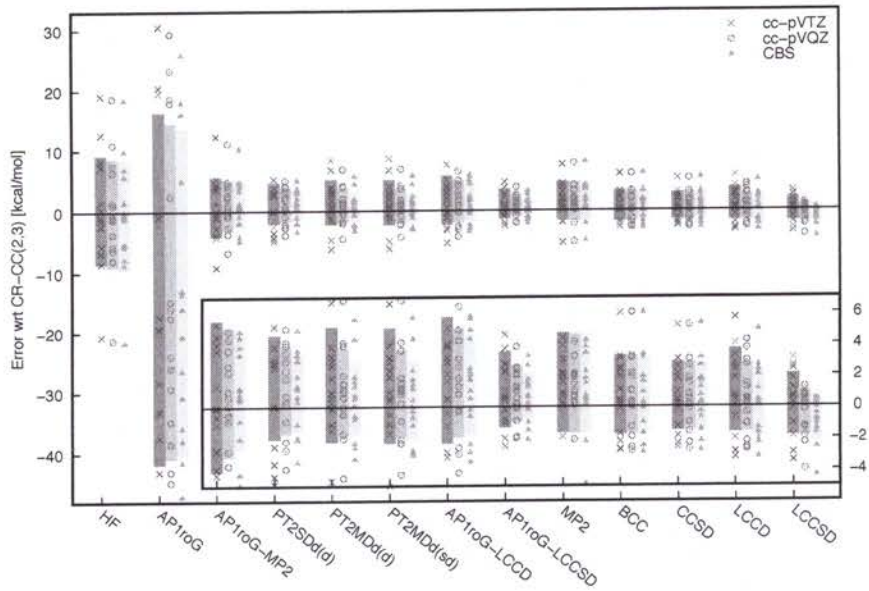
The standard errors in reaction energies with respect to CR-CC(2,3) reference values is displayed in Figure 7 for (a) the optimized AP1roG orbital basis and (b) canonical Hartree–Fock orbitals. In general, the performance of all PT methods can be divided in three different groups: (i) those with a diagonal zero-order Hamiltonian (PT2SDd/PT2MDd), (ii) those with an off-diagonal zero-order Hamiltonian (PT2SDo/PT2MDo), and (iii) those with an off-diagonal zero-order Hamiltonian and the full quantum-chemical Hamiltonian as perturbation operator (PT2b-type methods). The accuracy of the PT corrections with respect to CR-CC(2,3) reference data increases when going from PT methods (i) to (iii) reducing the RMSE from approximately 14 kcal/mol in PT2SDd/PT2MDd, to about 5 kcal/mol in PT2SDo/PT2MDo, to 2 kcal/mol in PT2b-type methods. The choice of the dual state and the inclusion of single excitations in the excitation operator do not significantly affect the accuracy of the PT methods (mean error, root mean square error, mean absolute error, maximum absolute error). Furthermore, excluding pair-excited determinants from the projection manifold in PT2b-type methods improves the accuracy of PT2b only marginally. Since pair-excitations are already described in the AP1roG reference function, it might, however, be advantageous to exclude pair excitations from the excitation operator $\hat{T}$ and hence eliminate the coupling to pair excitations modeled in the AP1roG reference function and pair excitations of the PT method, which both couple to the remaining PT amplitudes in the PT amplitude equations.

If the optimized natural AP1roG orbitals are replaced by canonical Hartree–Fock orbitals, only two distinct PT models persist, namely, PT2SDd (with double excitations) and PT2MDd (with double as well as single and double excitations). In contrast to the natural AP1roG orbitals basis, all PT methods yield similar error measures in the canonical Hartree–Fock basis with a standard error of about 3 kcal/mol. Therefore, the optimization of the molecular orbital basis and the AP1roG reference determinant might be unnecessary if the molecular system is dominated by dynamic correlation and molecular properties around the equilibrium geometry are considered, provided dynamic correlation effects are accounted for in the AP1roG model. If the optimal natural AP1roG orbitals are used in calculations, PT2b-type methods result in the smallest error measures (around 2 kcal/mol) and thus outperform all other

Katayun Boguslawski



(**a**) optimized AP1roG natural orbitals



(**b**) canonical Hartree–Fock orbitals

Figure 7: Errors in the optimized AP1roG orbital basis (a) and in the canonical Hartree–Fock basis (b). The grey bars indicate the standard errors in each basis set.

PT models. If the orbital optimization step is omitted, PT2SDd/PT2SDo provide the smallest errors that are similar to PT2b-type methods in the optimized AP1roG basis (RMSE around 3 kcal/mol).
Finally, a Linearized Coupled Cluster correction with an AP1roG reference function introduced in **[H4]** predicts reaction energies that deviate least from CR-CC(2,3) reference data reducing the RMSE to 1.4 kcal/mol. To minimize the error in AP1roG-LCC, single excitations are indispensable and have to be included in the cluster operator, both using optimized AP1roG orbitals and canonical Hartree–Fock orbitals. Most importantly, the AP1roG orbital basis does not need to be optimized if chemical accuracy (approximately 1 kcal/mol) is desired for predicting equilibrium properties of weakly-correlated systems. To conclude, AP1roG-LCCSD provides the most accurate reaction energies with respect to CR-CC(2,3) reference data, outperforming all investigated PT models as well as conventional electronic structure methods like MP2, BCC, CCSD, CCSD(T), LCCD, and LCCSD. Furthermore, PT2SDd/PT2SDo in the canonical Hartree–Fock basis provides the smallest errors among all investigated PT corrections (slightly better than MP2) and allows us to cheaply model the thermochemistry of main group elements ($\mathcal{O}(o^2v^2)$). For strongly-correlated systems, however, the molecular orbital basis needs to be optimized before a PT2SDd or PT2SDo correction is applied. Rotating the orbital basis increases the computational cost due to the four-index transformation and some additional prefactor of the orbital optimization.

c.6 Targeting excited states

So far, we have discussed AP1roG-based methods that allow us to accurately model ground-state electronic structures. In this section, we will elaborate on how the AP1roG ansatz can be extended to target excited states. Our first excited state models are discussed in **[H2]**. Most importantly, we have presented the first excited state extensions for an AP1roG reference function. In the following, we will briefly summarize our methodology to target excited states within AP1roG.
In order to model excited states in AP1roG, we can exploit the fact that the AP1roG wave function is equivalent to the pCCD ansatz and employ the equation of motion (EOM) formalism.[3,128,129] In EOM-CC, excited states are parametrized by a linear CI-type ansatz,[83]

$$\hat{R} = \sum_{\mu} c_\mu \hat{\tau}_\mu, \tag{64}$$

where the summation is over all excitation operators present in the cluster operator $\hat{T}$ as well as the identity operator $\hat{\tau}_0$. The operator $\hat{R}$ is then used to generate the target state from the initial CC state,

$$|\Psi\rangle = \hat{R}\exp(\hat{T})|\Phi_0\rangle = \sum_{\mu} c_\mu \hat{\tau}_\mu \exp(\hat{T})|\Phi_0\rangle, \tag{65}$$

with $|\Phi_0\rangle$ being the CC reference determinant.
To arrive at the EOM-CC working equations, it is convenient to use the normal-product form of the Hamiltonian, $\hat{H}_N = \hat{H} - \langle\Phi_0|\hat{H}|\Phi_0\rangle$. Furthermore, we will

Katarzyna Boguslawska

disregard any excitation properties, like dipole moments, and focus on excitation energies instead. In that case, we have to solve for the $\hat{R}$ amplitudes only. Our target-state Schrödinger equation then reads

$$\hat{H}_N \hat{R} \exp(\hat{T})|\Phi_0\rangle = \Delta E \hat{R} \exp(\hat{T})|\Phi_0\rangle, \tag{66}$$

where ΔE is the energy difference with respect to the Fermi vacuum expectation value $|\Phi_0\rangle$. Introducing the similarity transformed Hamiltonian in normal-product form $\hat{\mathcal{H}}_N = \exp(-\hat{T})\hat{H}_N \exp(\hat{T})$ and subtracting the equation for the CC ground state, we obtain the EOM-CC equations for the $\hat{R}$ amplitudes,

$$[\hat{\mathcal{H}}_N, \hat{R}]|\Phi_0\rangle = \omega \hat{R}|\Phi_0\rangle, \tag{67}$$

where ω are the excitation energies with respect to the CC ground state, $\exp(\hat{T})|\Phi_0\rangle$. The excitation energies are thus the eigenvalues of a non-Hermitian matrix,

$$\begin{bmatrix} 0 & \langle\Phi_0|\hat{\mathcal{H}}_N|\mu\rangle \\ 0 & \langle\nu|[\hat{\mathcal{H}}_N, \tau_\mu]|\Phi_0\rangle \end{bmatrix}, \tag{68}$$

where the first row is associated with the CC reference state and the subsequent rows correspond to the excited configurations $\nu > 0$. The EOM-CC working equations may be solved using, for instance, non-Hermitian extensions of the Davidson algorithm to determine the lowest-lying excited electronic states.

c.6.1 Electron-pair excitations

Here, we are considering a pCCD reference function |pCCD⟩ (or |AP1roG⟩) as a special CC state confined to electron-pair states. As indicated in eq. (1), the pCCD cluster operator contains only electron-pair excitations, $\hat{T} = \hat{T}_\mathrm{p} = \sum_{ia} t_i^a a_a^\dagger a_{\bar{a}}^\dagger a_{\bar{i}} a_i$. In the corresponding EOM model, the $\hat{R}$ operator is thus restricted to the identity operator $\hat{\tau}_0$ as well as all pair excitations present in the cluster operator $\hat{T}_\mathrm{p}$,

$$\hat{R}_\mathrm{p} = c_0 \hat{\tau}_0 + \sum_{ia} c_{i\bar{i}}^{a\bar{a}} \hat{\tau}_{a\bar{a}i\bar{i}}, \tag{69}$$

where $\hat{\tau}_{a\bar{a}i\bar{i}} = a_a^\dagger a_{\bar{a}}^\dagger a_{\bar{i}} a_i$ creates an electron pair in the virtual orbital a. To obtain the target-state Schrödinger equation of EOM-pCCD, we have to substitute the general cluster operator $\hat{T}$ by the pair-excitation operator $\hat{T}_\mathrm{p}$ in eq. (66),

$$\hat{H}_N \hat{R}_\mathrm{p} \exp(\hat{T}_\mathrm{p})|\Phi_0\rangle = \Delta E \hat{R}_\mathrm{p} \exp(\hat{T}_\mathrm{p})|\Phi_0\rangle. \tag{70}$$

The $\hat{R}_\mathrm{p}$ amplitudes are determined from the EOM-pCCD equations (restricting $\hat{R}$ to $\hat{R}_\mathrm{p}$ and $\hat{T}$ to $\hat{T}_\mathrm{p}$ in eq. (67)),

$$[\hat{\mathcal{H}}_N^{(\mathrm{p})}, \hat{R}_\mathrm{p}]|\Phi_0\rangle = \omega_\mathrm{p} \hat{R}_\mathrm{p}|\Phi_0\rangle, \tag{71}$$

where $\hat{\mathcal{H}}_N^{(\mathrm{p})}$ indicates the similarity transformed Hamiltonian of pCCD, $\hat{\mathcal{H}}_N^{(\mathrm{p})} = \exp(-\hat{T}_\mathrm{p})\hat{H}_N \exp(\hat{T}_\mathrm{p})$, and ω_p are the electron-pair excitation energies. Thus, EOM-

Katarzyna Bąkowicz

pCCD allows us to model electron-pair excited states only. In order to target singly excited or general doubly excited states, we have to extend the pCCD cluster operator to include excitations beyond electron pairs. This can be done either by changing to a frozen-pair CCSD[130] formalism or to a Linearized CCSD correction with an pCCD reference function. However, we will consider a different, cost-effective approach to account for single excitations in the pCCD model that does not scale as $\mathcal{O}(n^6)$ as conventional EOM-CCSD methods. Note that EOM-pCCD scales as $\mathcal{O}(o^2v^2)$, where o is the number of occupied orbitals (equivalent to the number of electron pairs) and v is the number of virtual orbitals.

c.6.2 Accounting for single excitations

As proposed by Forseman *et al.*,[131] configuration interaction with only single substitutions (CIS) represents an accurate model to investigate (singly) excited electronic states, even for large systems. In the CIS method, the reference is a single Slater determinant obtained from an SCF procedure, while the CIS wave function is expanded as

$$|\mathrm{CIS}\rangle = c_0|\Phi_0\rangle + \sum_{ia} c_i^a |_i^a\rangle, \tag{72}$$

with $|_i^a\rangle$ being a singly excited determinant where the (occupied) orbital i of $|\Phi_0\rangle$ has been substituted by the (virtual) orbital a. Similar to CIS, we will include single excitations in EOM-pCCD by extending the $\hat{R}_\mathrm{p}$ operator of eq. (69). In addition to the identity operator $\hat{\tau}_0$ and all pair excitations $\hat{T}_\mathrm{p}$, $\hat{R}_\mathrm{p}$ also contains a summation over all single excitations,

$$\hat{R}_\mathrm{ps} = c_0\hat{\tau}_0 + \sum_{ia} c_i^a \hat{\tau}_{ai} + \sum_{ia} c_{i\bar{i}}^{a\bar{a}} \hat{\tau}_{a\bar{a}i\bar{i}}, \tag{73}$$

where $\hat{\tau}_{ai}$ is a singlet excitation operator $\hat{\tau}_{ai} = a_a^\dagger a_i + a_{\bar{a}}^\dagger a_{\bar{i}}$ that creates a singly excited electronic state with respect to $|\Phi_0\rangle$, $|_i^a\rangle = \hat{\tau}_{ia}|\Phi_0\rangle$. The $\hat{R}_\mathrm{ps}$ amplitudes are determined from solving

$$[\hat{\mathcal{H}}_N^{(\mathrm{p})}, \hat{R}_\mathrm{ps}]|\Phi_0\rangle = \omega_\mathrm{ps}\hat{R}_\mathrm{ps}|\Phi_0\rangle, \tag{74}$$

where we still have the similarity transformed Hamiltonian of pCCD, $\hat{\mathcal{H}}_N^{(\mathrm{p})}$, while ω_ps are the excitation energies of both singly excited and pair excited states. We will label this simplified model as EOM-pCCD+S to indicate that single excitations are included *a posteriori* in the $\hat{R}_\mathrm{p}$ operator. In contrast to CIS that uses a single Slater determinant as reference, EOM-pCCD+S employs the pCCD wave function as reference state. Furthermore, electron correlation effects are included through the $\hat{T}_\mathrm{p}$ operator in the similarity transformed Hamiltonian.
Similar to EOM-pCCD, the excitation energies are obtained by diagonalizing a non-

Katarzyna Boguslawski

Hermitian matrix of the form

$$\begin{bmatrix} 0 & \langle\Phi_0|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_j^b\rangle & \langle\Phi_0|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_{j\bar{j}}^{b\bar{b}}\rangle \\ \langle_i^a|\hat{\mathcal{H}}_N^{(\mathrm{p})}|\Phi_0\rangle & \langle_i^a|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_j^b\rangle & \langle_i^a|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_{j\bar{j}}^{b\bar{b}}\rangle \\ 0 & \langle_{i\bar{i}}^{a\bar{a}}|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_j^b\rangle & \langle_{i\bar{i}}^{a\bar{a}}|\hat{\mathcal{H}}_N^{(\mathrm{p})}|_{j\bar{j}}^{b\bar{b}}\rangle \end{bmatrix}. \tag{75}$$

Note that the final working equations have been spin-summed and the calcuated excitation energies are thus spin-free. Furthermore, in contrast to conventional EOM-CC methods, the first column does not equal zero because single excitations are not included in the cluster operator of pCCD. Thus, terms like $\langle_i^a|\hat{\mathcal{H}}_N^{(\mathrm{p})}|\Phi_0\rangle$ do not vanish as they are not incorporated in the ground-state CC amplitude equations. Although we can account for single excitations in a rather straightforward way, we loose size-intensivity in the EOM model. For the molecular systems investigated so far, the error introduced by extending only the $\hat{R}_\mathrm{p}$ operator is approximately three orders of magnitude smaller than the actual excitation energies, while the computational cost increases insignificantly compared to EOM-pCCD (the Hamiltonian still contains terms that scale as $\mathcal{O}(o^2v^2)$, but with a larger pre-factor). Thus, EOM-pCCD+S represents a cost-effective starting point to study singly excited electronic states in the pCCD model.

c.6.3 Targeting doubly-excited states in all-trans polyenes

All-trans polyenes are model systems for carotenoids and polyene chromophores that play an important role in photoprocesses. The proper description of the two lowest-lying excited states poses a challenge to both experiment and quantum chemistry approaches,[132–145] especially because doubly excited configuration are required to accurately model ground and excited states of longer polyenes. These molecules are thus ideal test systems to assess the performance of our proposed excited state models based on electron-pair states. In **[H2]**, we have studied all-trans polyenes ranging from 2 to 7 double bonds, *i.e.*, from C_4H_6 to $C_{14}H_{16}$.

The excitation energies of the two lowest-lying excited states for DFT-optimized structures are presented in Table 5. EOM-pCCD+S yields excitation energies that are lower than the corresponding CIS(D) values. Nonetheless, EOM-pCCD+S predicts the wrong order of the first bright and dark state in all-trans polyenes using the DFT-optimized structures. For the DMRG-optimized structures, the excitation energies of the first dark states decrease by approximately 0.6 to 0.8 eV compared to the DFT-optimized structures so that the first (adiabatically excited) dark state lies above the first (adiabatically excited) bright state by approximately 0.2-0.4 eV.

Our calculations demonstrate that the EOM-pCCD+S excitation energies (and the character of the excited states) strongly depend on the molecular structures used in calculations (compare Tables 5 and 6). Specifically, the character of the $2^1A_g^-$ state can only be properly predicted if molecular structures are allowed to relax, resulting in a dominant doubly-excited HOMO^2 $\rightarrow$LUMO^2 configuration and two singly-excited configurations (HOMO$\rightarrow$LUMO+1 and HOMO$-1\rightarrow$LUMO). In order to accurately model the first dark state in all-trans polyenes, double excitations beyond electron-pair excitations as well as higher excitations (triples, etc.) might be important to

Katarzyna Boguslawska

Table 5: Vertical excitation energies of the two lowest-lying excited states in all-trans polyenes C_4H_6 to $C_{14}H_{16}$ calculated for EOM-pCCD, EOM-pCCD+S, and different quantum chemistry methods. The molecular structures were optimized using DFT. Note that the 6-31G basis set was used in CIS(D), while the cc-pVDZ basis set was utilized in MRMP. CASSCF was performed in a double-zeta basis set (see corresponding references). The excitation energies of EOM-pCCD and EOM-pCCD+S are determined for the cc-pVDZ basis set, while the corresponding results for the 6-31G basis set are given in parenthesis.

C=C	EOM-pCCD	EOM-pCCD+S	CIS(D) [144]	MRMP [146]	CASSCF [147]
			$2^1A_g^-$		
2	10.56 (10.49)	7.45 (7.37)	9.01	6.31	6.67
3	9.11 (9.04)	6.79 (6.75)	7.81	5.10	5.64
4	8.11 (8.02)	6.15 (6.09)	6.78	4.26	5.16
5	7.42 (7.32)	5.69 (5.63)	6.12	3.68	4.32
6	6.93 (6.82)	5.37 (5.29)	5.55	3.19	–
7	6.58 (6.46)	5.13 (5.05)	5.14	2.80	–
			$1^1B_u^+$		
2	–	7.20 (7.44)	8.09	6.21	7.73
3	–	5.98 (6.16)	6.78	5.25	7.06
4	–	5.19 (5.34)	5.95	4.57	6.62
5	–	4.62 (4.75)	5.43	4.17	6.37
6	–	4.20 (4.31)	5.00	3.87	–
7	–	3.87 (3.97)	4.70	3.60	–

Table 6: Vertical and adiabatic excitation energies of the two lowest-lying excited states in all-trans polyenes $C_{10}H_{12}$ to $C_{14}H_{16}$ calculated with EOM-pCCD, EOM-pCCD+S, and DMRG for different DMRG-optimized geometries. Note that different active spaces are used in DMRG calculations (see computational details of **[H2]** and ref. 148), while all orbitals are active in EOM-pCCD and EOM-pCCD+S. The DMRG reference data is taken from ref. 148. Experimental data is taken from ref. 149. pCCD+S indicates EOM-pCCD+S.

C=C	pCCD+S	DMRG	pCCD+S	DMRG	pCCD+S	DMRG	pCCD+S	DMRG	Exp.
	v		a		v_π		a_π		
					$2^1A_g^-$				
5	6.28	5.43	5.18	4.01	6.07	4.51	4.85	3.36	3.03
6	5.99	4.76	4.61	3.41	5.84	4.15	4.60	2.99	2.69
7	5.76	4.64	4.44	3.22	5.63	3.91	4.45	2.73	2.44
					$1^1B_u^+$				
5	4.91	5.35	5.20	4.98	4.79	5.77	4.63	5.49	3.57
6	4.50	4.98	4.30	4.60	4.41	5.41	4.26	5.13	3.31
7	4.25	4.66	4.04	4.29	4.14	5.16	3.99	4.87	3.12

capture the missing correlation effects in the targeted excited states that cannot be described within pCCD. Nonetheless, our numerical results suggest that EOM-pCCD+S represents a good and cost-effective starting point to investigate singly-excited states in the pCCD model.

Katarzyna Boguslawski

c.7 Modeling heavy-element chemistry with geminals

Heavy elements are immensely difficult to describe theoretically because correlation effects and relativistic effects have to be described on equal footing, which poses a challenge for present-day quantum chemistry, especially when the molecular system under investigation contains more than one heavy element. The treatment of relativistic effects via one-component approaches remains computationally advantageous.[150] As scalar relativistic effects are nowadays relatively straightforward to describe, and there are sufficiently accurate approaches to treat spin-orbit interactions perturbatively, the primary difficulty is the large number of competing highly-correlated electronic states resulting from distributing electrons among the energetically close-lying valence atomic orbitals, which requires a multi-reference treatment. Since AP1roG does not require us to restrict the active space as in conventional multi-configurational method, it represents computationally an ideal ansatz to model heavy-element chemistry. In the following, we will assess the accuracy of our proposed AP1roG-based methods in describing ground and excited states of small actinide complexes. All numerical results are presented in **[H2,H4]**.

The uranyl cation (UO_2^{2+}) is a small building block of a large variety of uranium-containing complexes and clusters.[151] This molecule has a linear structure and a singlet ground-state electronic configuration. Its characteristic symmetric and asymmetric U–O vibrational frequencies are used to identify the presence of UO_2^{2+} in larger molecular assemblies.[151,152] While the electronic structure of the uranyl cation is well-understood around the equilibrium structure,[6,10,17,20,24,28,108,153–155] the complicated nature of the U–O bond hampers a theoretical description at larger U–O distances using standard quantum chemistry approaches.[108] One of the limiting factors that impede theoretical studies is the large number of strongly-correlated electrons distributed among the $5f$-, $6d$-, and $7s$-orbitals. In addition, the $6s$- and $6p$-core-valence orbitals are easily polarizable and have a non-negligible contribution to the correlation energy. However, around the equilibrium structure, the uranyl cation is well described by single-reference CC theory if all important electrons are correlated. This allows us to assess the performance of the above presented AP1roG-based models in describing both static and dynamic correlation effects originating from the $5f$-, $6d$-, and $7s$- as well as the core-valence electrons.

The equilibrium bond lengths and vibrational frequencies of the uranyl cation obtained by different quantum chemistry methods are shown in Table 8 (scalar relativistic effects have been included using relativistic effective core potentials). As expected, AP1roG considerably underestimates the equilibrium bond length, while ω_e is in good agreement with CCSD(T). Adding dynamic correlation effects on top of AP1roG shifts r_e closer to the CCSD(T) reference data. The shape of the potential, however, strongly depends on the AP1roG dynamic correlation model. Specifically, PT2b(d) (originally introduced as PTb) results in a much steeper potential energy surface overestimating vibrational frequencies by more than 330 cm^{-1} compared to CCSD(T), while LCCD and LCCSD preserve the shape of the potential energy surface and yield a vibrational frequency that agree well with AP1roG and CCSD(T) data (differences amount to approximately 20 cm^{-1}). The overall accuracy of AP1roG-LCC lies between CCSD and CCSD(T), being closer to the latter. We should empha-

Katarzyna Boguslawska

size that PT2SDd(d) (equivalent to PTa) completely fails for the UO_2^{2+} molecule and produces a discontinuous potential energy surface around the equilibrium (see also Figure 8). Furthermore, the CASSCF equilibrium distance strongly depends on the size of the active space chosen in CASSCF calculations. Specifically, increasing the active space from CAS(10,10) to CAS(12,12), *i.e.*, including the σ- and σ^*-orbitals, results in spectroscopic constants that are in good agreement with AP1roG-LCCD and CCSD(T) data.

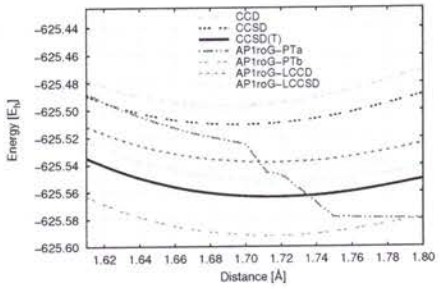


Figure 8: Potential energy surfaces for the symmetric stretching of the UO_2^{2+} molecule around the equilibrium geometry. Note that the CASSCF potential energy surfaces are much higher in energy and are thus not shown.

Method	r_e [Å]	ω_e [cm^{-1}]
AP1roG	1.669(−0.047)	1062(+53)
AP1roG-PT2b(d)	1.715(−0.001)	1340(+331)
AP1roG-LCCD	1.712(−0.004)	997(−12)
AP1roG-LCCSD	1.724(+0.008)	1027(+18)
CAS(10,10)SCF	1.694(−0.022)	1079(+70)
CAS(12,12)SCF	1.707(−0.009)	1034(+25)
CCD	1.690(−0.026)	1125(+116)
CCSD	1.697(−0.019)	1068(+59)
CCSD(T)	1.716	1009

Table 7: Spectroscopic constants for the symmetric dissociation of the UO_2^{2+} molecule for different quantum chemistry. The differences are with respect to CCSD(T) reference data. The CASSCF and CC data were taken from Ref. 108.

Figure 8 shows the fitted potential energy surfaces around the equilibrium for selected quantum chemistry methods. AP1roG-LCC yields total electronic energies that are between CCSD and CCSD(T), while the potential energy surface predicted by AP1roG-PTb is considerably lower than the CCSD(T) reference curve. Note that the potential energy surfaces optimized by CASSCF lie much higher in energy and are thus not shown in Figure 8.

The spectrum of the uranyl cation is well understood[17,20,28,151,156,157] and its lowest-lying excited states are purely singly-excited states. It thus represents an ideal test system to assess the accuracy of EOM-pCCD+S. The vertical and adiabatic excitation energies for the four lowest-lying excited states in EOM-pCCD+S and CIS are summarized in Table 8. Note that for the $UO_2{}^{2+}$ molecule the two lowest-lying excited states can be accurately described within the EOM-CCSD model, which yields excitations energies that are similar to completely renormalized EOM-CCSD(T) reference values.[155] For the higher-lying excited states, however, dynamic correlation effects become important and a triples correction has to be included to accurately model those states. The EOM-CCSD results in Table 8 can thus be considered as upper bounds of the excitation energies for the $\pi_u \rightarrow \delta_u$ state.

In general, EOM-pCCD+S overestimates vertical excitations energies of the two lowest-lying excited states ($\sigma_u \rightarrow \phi_u$ and $\sigma_u \rightarrow \delta_u$) by approximately 0.5 eV, while the corresponding adiabatic excitation energies are overestimated by about 0.7 eV. Note that for the uranyl cation, CIS outperforms EOM-pCCD+S and deviates from EOM-CCSD reference data by approximately 0.3 (vertical excitations) to 0.6 eV (adiabatic excitations). As expected, neither EOM-pCCD+S nor CIS are able to accu-

Katarzyna Boguslawska

Table 8: Vertical and adiabatic excitation energies and equilibrium U–O bond lengths of the four lowest-lying excited states in the UO_2^{2+} molecule calculated for EOM-pCCD+S and different quantum chemistry methods.

Excitation		EOM-pCCD+S	CIS	EOM-CCSD
		ω [eV]		
vertical	$\sigma_u \rightarrow \phi_u$	4.55	4.33	4.02
	$\sigma_u \rightarrow \delta_u$	4.97	4.79	4.36
	$\pi_u \rightarrow \delta_u$	7.99	7.84	5.28
	$\pi_u \rightarrow \delta_u$	8.11	8.00	5.36
adiabatic	$\sigma_u \rightarrow \phi_u$	4.42	4.18	3.70
	$\sigma_u \rightarrow \delta_u$	4.87	4.66	4.08
	$\pi_u \rightarrow \delta_u$	7.62	7.47	4.63
	$\pi_u \rightarrow \delta_u$	7.79	7.62	4.72
Excitation		r_e [Å]		
adiabatic	$\sigma_u \rightarrow \phi_u$	1.712	1.694	1.772
	$\sigma_u \rightarrow \delta_u$	1.709	1.690	1.768
	$\pi_u \rightarrow \delta_u$	1.746	1.719	1.805
	$\pi_u \rightarrow \delta_u$	1.745	1.720	1.805

rately predict the excitation energies for the $\pi_u \rightarrow \delta_u$ states as these excited states are dominated by dynamic correlation effects which are not included in CIS and only marginally accounted for in pCCD. We should emphasize that EOM-pCCD+S provides equilibrium U–O bond lengths of excited states that deviate less from the EOM-CCSD reference values ($\Delta r_e \approx 0.06$ Å compared to $\Delta r_e \approx 0.08$ Å in CIS).
To conclude, the EOM-pCCD+S model predicts the correct order of the lowest-lying excited states in the uranyl cation and provides excitation energies of decent accuracy with errors of about 0.5 eV with respect to EOM-CCSD reference values. However, errors in excitation energies are slightly worse than in the simple CIS model.

c.8 Summary and impact of research

The accurate and inexpensive description of electron correlation effects represents a challenging problem in quantum chemistry. Due to their unfavorable computational scaling with system size, conventional multi-reference methods are usually limitted to small molecular systems or small basis sets and require the definition of active orbital spaces. In the presented series of publications, we scrutinized alternative electronic structure models that allow us to describe both static/nondynamic and dynamic electron correlation effects without introducing active orbital spaces. Our methods are based on wave functions restricted to electron-pair states (geminals) to capture static/nondynamic electron correlation. The missing dynamic electron correlation is added *a posteriori* by so-called broken-pair states. Specifically, we focused on the AP1roG (or pCCD) wave function ansatz presented in ref. 52. Our numerical results demonstrate that the AP1roG model and its extensions allow us to reliably describe

Katarzyna Boguslawski

strongly and weakly correlated systems, where the accuracy of ground- and excited state properties is similar to or better than results predicted by conventional quantum chemistry methods.

c.8.1 Recovering size-consistency

Although being size-extensive by construction, the AP1roG wave function is not size-consistent and hence does not provide reliable potential energy surfaces. Size-consistency can be recovered by optimizing the one-particle basis functions used to construct the geminals. We have presented the first variational orbital optimization protocol using a Lagrange formulation as well as different approximate non-variational orbital optimization techniques that are based on projecting out the seniority-two sector. All orbital-optimization procedure result in size-consistent potential energy surfaces. Nonetheless, numerical results indicate that the variational orbital optimization algorithm is the most robust and stable one and smoothly converges even for difficult molecular systems. Each orbital optimization step, however, requires a four-index transformation and thus deteriorates the scaling of the AP1roG wave function from $\mathcal{O}(o^2v^2)$ to $\mathcal{O}(N^5)$.

c.8.2 Approximating static and nondynamic correlation with electron-pair states

The accuracy of (voo-)AP1roG in describing static and nondynamic correlation has been analyzed using concepts of quantum information theory. Specifically, we used the single-orbital entropy and orbital-pair mutual information to dissect correlation effects into different contributions. Our numerical study focuses on one-dimensional systems where quantum fluctuations have a more pronounced role. Although AP1roG captures the strongest orbital-pair correlations in the weak correlation limit as well as for intermediate interaction strengths, it overestimates orbital-pair correlations in the strong correlation limit. This overcorrelation is introduced by the (variational) orbital optimization procedure. It remains, however, ambiguous if AP1roG represents a good zero-order wave function for *a posteriori* corrections in the strong correlation regime.

c.8.3 Capturing dynamic correlation

To reach chemical accuracy in predicting, for instance, spectroscopic constants and reaction energies, we have to include dynamic electron correlation effects in the AP1roG wave function. Our proposed dynamic energy corrections include dynamic electron correlation *a posteriori* on top of the AP1roG wave function. Specifically, we have developed different flavours of Perturbation Theory models of second order and a Linearized Coupled Cluster correction. In general, our PT2 approaches allow us to accurately predict equilibrium properties, like equilibrium distances or vibrational frequencies. However, all studied PT2 models fail in modeling potential energy surfaces and usually diverge in the dissociation limit. Our LCCD and LCCSD corrections are more robust and result in improved total and relative (reaction) energies as well as energy profiles. Although AP1roG-LCCSD yields equilibrium distances and vibrational frequencies that deviate more from reference values than most PT2 corrections, it considerably improves results obtained by conventional Coupled Cluster

Katayan Bajdich

models (CCSD, BCC, LCCSD). Thus, AP1roG-LCCSD (combined with a variational orbital optimization protocol of AP1roG) allows us to approach chemical accuracy in many molecular systems dominated by both static and dynamic electron correlation.

c.8.4 Targeting excited states

Since AP1roG can be re-written as a Coupled Cluster wave function ansatz, we can exploit the Equation-of-Motion formalism to target excited states. In this series of publication, we have presented two excited state models with an AP1roG reference function. Our first EOM model allows us to target doubly-excited states only as the AP1roG reference functions is limited to electron-pair excitations. In the second model, we have also included single-excitations in the EOM ansatz, restricting the Coupled Cluster reference function to electron-pair excitations. Although being a simple and inexpensive excited state model ($\mathcal{O}(o^2v^2)$), the EOM-pCCD+S (or EOM-AP1roG+S) approach breaks size-intensivity. Numerical studies, however, indicate that the introduced errors are orders of magnitude smaller than the excitation energies and that EOM-pCCD+S represents a good starting point to predict singly-excited states.

c.8.5 Modeling heavy-element chemistry

Finally, we have applied our proposed wave function models to describe the ground- and excited states of heavy-element containing compounds. Specifically, we focused on the uranyl cation that is a small building block of a large variety of uranium-containing complexes and clusters. This particular test systems allows us to assess our AP1roG-based models in modeling the correlation effects originating from the $5f$-, $6d$-, and $7s$- as well as the core-valence electrons. Most importantly, AP1roG-LCCSD provides accurate ground-state potential energy surfaces, while EOM-pCCD+S yields excited states properties that are in good agreement with conventional electronic structure methods like CIS and EOM-CCSD. Thus, we believe that our proposed models are promising alternatives to describe heavy-element-containing complexes and clusters.

c.8.6 Implementation

All working equations have been implemented partly in the HORTON2.0 program suit, an open-source quantum chemistry software package written in Python and C++, and in the PIERNIK program package, an open-source quantum chemistry software package based on the HORTON2.0 program suit. Specifically, we have implemented the following modules in PIERNIK as well as HORTON2.0

- the geminal module (restricted AP1roG with and without orbital optimization)
- the dynamic correlation module (Møller-Plesset Perturbation Theory of second order, PTa and PTb corrections on top of AP1roG)
- the orbital entanglement module (single-orbital entropy and mutual information for seniority-zero wave functions)
- the orbital localization module (Pipek-Mezey localization)

- extensions to the linear algebra module (two- and four-index transformation, general tensor contractions and other tensor manipulations)

PIERNIK is being maintained by an interdisciplinary software development team at NCU in Torun. Our contributions to PIERNIK include

- the extension of the dynamic correlation module (Linearized Coupled Cluster correction based on a Hartree–Fock and AP1roG reference function as well as PT2SD- and PT2MD-type methods)
- the Equation-of-Motion module, which includes EOM-pCCD and EOM-pCCD+S
- the Davidson diagonlization module
- the perturbation-based quasi-Newton solver module (including DIIS)

Currently, the PIERNIK developer team is working on the first release (PIERNIK1.0). As soon as completed, PIERNIK1.0 will be available to the general public.

c.8.7 Outlook

Our numerical studies presented in this series of publications highlight the good performance of AP1roG and its dynamic energy corrections in predicting electronic energies and energy-derived quantities, like spectroscopic constants. Nevertheless, we have also pointed out that variational orbital optimization within AP1roG overestimates orbital-pair correlations. It remains ambiguous whether an LCCSD correction with an orbital-optimized AP1roG reference function provides accurate electronic wave functions and whether it can cure the overcorrelation introduced by voo-AP1roG. The orbital-pair correlations within AP1roG-LCCSD can be determined from the response one-, two-, three-, and four-particle reduced density matrices as outlined in **[H5]**. Most importantly, an orbital-pair correlation analysis will allow us to decide whether an LCCSD correction yields accurate and reliable electronic wave functions. Furthermore, our simple excited state models do not include dynamic electron correlation effects and hence might yield inaccurate spectra for systems where dynamic correlation is important. A dynamic energy correction that can be extended to target excited states using the EOM formalism is LCCSD, which results in the AP1roG-LCCSD model. The performance of EOM-pCCD-LCCSD (or EOM-AP1roG-LCCSD) is currently under investigation and will be the subject of future work.

5 Discussion of other scientific achievements, after PhD studies

In additional to the development of electronic structure approaches based on electron-pair states, our research covers the theoretical modeling of heavy-element-containing compounds and their properties using conventional and unconventional electron correlation methods as well as application of concepts of quantum information theory to

Katarzyna Boguslawska

obtain a qualitative understanding of electronic structures. In the following, we will briefly highlight the most important findings.

5.1 Modeling of heavy-element chemistry using conventional and unconventional electronic structure methods

Our work focuses on CASSCF/CASPT2 and DMRG calculations in transition metal and actinide chemistry. We further use these multi-reference methods to benchmark DFT results.

5.1.1 Assessment of DFT magnetization densities and magnetic properties

We presented a systematic theoretical study of electronic structures, magnetization densities, and magnetic properties of iridium PNP pincer-type complexes containing various non-innocent ligands, like nitrido, azide, and nitrosyl ligands **[P3]**. The quality and accuracy of various Density Functional Approximations in predicting magnetization densities is assessed by comparing them to CASSCF reference distributions. Our analysis points to qualitative differences in DFT magnetization densities at the iridium metal center and the pincer ligand backbone compared to CASSCF reference data when the non-innocent ligands are changed from nitrido, to azide, to nitrosyl. These observations are reflected in large differences in hyperfine couplings calculated for the iridium metal center. In summary, we find that none of the tested exchange–correlation functionals is able to provide a satisfactory description of the magnetization densities and magnetic properties in the investigated iridium complexes. Similar problems have been already observed for iron complexes containing non-innocent ligands. In contrast to previous findings, however, conventional electron correlation methods, like CASSCF, are sufficient to accurately describe the electronic structure of the considered iridium compounds. This study emphasizes the importance of analyzing density functionals to understand their failures and weaknesses and to improve current approximations to the exchange–correlation functional to be applicable to challenging problems in transition-metal chemistry.

5.1.2 Cation-cation-interactions in actinyl dications

We investigated cation-cation interactions (CCIs) between different actinyl ions, such as uranyl **[P2]** and neptunyl dications (under revision in *Inorg. Chem.*). In **[P2]**, we present a state-of-the-art DFT computational study of the uranyl(VI) and uranyl(V) CCIs in aqueous solution. Most importantly, we provide reliable electronic structures of two interacting uranyl(VI) and uranyl(V) subunits as well as those of the uranyl(VI) and uranyl(V) clusters for the first time. Our theoretical study elucidates the impact of CCIs on changes in the molecular structure as well as changes in vibrational and UV-Vis spectra of the bare uranyl(VI) and uranyl(V) moieties for different total spin-states and total charges of the dications. Specifically, the formation of both diamond- and T-shaped CCIs introduces structural asymmetries to at least one uranyl subunit resulting in one longer interior U–O and one shorter terminal U–O bond. These structural changes affect the vibrational spectra of all CCI

Katarzyna Bączek

clusters and prevent any discrimination between and identification of the uranyl(VI) and uranyl(V) species in the supramolecular compounds. In contrast to vibrational spectra, the analysis of the electronic spectra of CCI clusters is more useful and provides direct information about the oxidation state of the uranium atom. Specifically, the $[(UO_2)_2]^{4+}$ compounds have similar spectral characteristics as the bare $[UO_2]^{2+}$ molecule, while the electronic transitions calculated for $[(UO_2)_2]^{3+}$ and $[(UO_2)_2]^{2+}$ agree well with the characteristic f–f transitions observed in the $[UO_2]^{+}$ unit. Our study suggests that the differentiation between the uranyl(VI) and uranyl(V) subunits in CCI clusters based on a pure vibrational frequency analysis remains ambiguous, primarily because of the overlay or shift of the spectral characteristics of both uranyl cations. A frequency analysis can, however, be used to confirm if the CCI cluster underwent major structural changes compared to the bare uranyl subunits. Finally, electronic spectroscopy allows us to identify the oxidation state of the uranium atom in CCI clusters and thus to reliably dissect the CCI supra-molecule in its uranyl building blocks.
In a follow-up publication, we investigate the T-shaped and diamond-shaped neptunyl(V) and neptunyl(VI) dimers (under revision in *Inorg. Chem.*). In this work, we scrutinize their molecular structures, solvation effects, the interplay of static and dynamic correlation, and the influence of spin orbit coupling on the ground-state and lowest-lying excited states for different total spin-states and total charges of the neptunyl dications. Most importantly, our study highlights the complex interplay of correlation effects and relativistic corrections in the description of the ground and lowest-lying excited states of neptunyl dications. The ground-state of the investigated clusters is strongly affected by both electron correlation effects and spin–orbit coupling. Specifically, accounting for dynamic correlation using a CASPT2 corrections changes the ground-state from a quintet to a triplet state. Inclusion of spin-orbit coupling entails mixing between triplet and quintet states for the T-shaped CCI, while the diamond-shaped CCI contains only quintet states.

5.1.3 The mysterious interaction between CUO and a noble gas matrix. A DMRG perspective

In collaboration with the group of Prof. Legeza, we have performed electronic structure calculations on actinide molecules using the DMRG algorithm. In **[P10]**, we present the first DMRG study on actinide chemistry, where we scrutinize the anticipated singlet-triplet spin crossover of the CUO molecule diluted in a noble gas matrix (within a scalar relativistic treatment) and elucidated the mysterious interaction of the CUO unit with the noble gas environment. Specifically, the interaction of the CUO molecule with the surrounding noble gas matrix was investigated in terms of complexation energies and dissected using orbital correlation patterns. The complexation of the CUO molecule by noble gases lowers the first excited $^3\Phi$ state with respect to the $^1\Sigma^+$ state compared to the bare CUO complex, whose ground state is a $^1\Sigma^+$ state. In general, the largest coordination energy is found for the $^3\Phi$ state for both the neon and argon noble gas matrix. With addition of spin–orbit coupling, the energy difference between the CUO moiety embedded in neon and argon atoms is brought down to 0.02 eV, and therefore the anticipated ground-state spin crossover

Katarzyna Boguslawski

might occur. Furthermore, the strongest uranium–noble gas interaction is found for the $CUOAr_4$ complex in its triplet state. Most importantly, **[P10]** was the first theoretical study confirming the experimentally anticipated singlet-triplet ground state change of the CUO molecule.

5.1.4 Singlet ground-state actinide chemistry with geminals

Our numerical results presented in **[P8]** demonstrate that voo-AP1roG provides an accurate, cheap, and robust alternative to standard multi-reference quantum chemistry methods in studying single- and multiple-bond breaking processes in closed-shell systems. These findings motivated us to investigate the performance of (voo-)AP1roG in modeling ground-state actinide chemistry. **[P7]** is the first application of the voo-AP1roG method to singlet-state actinide chemistry. In this work, we assess the accuracy and reliability of the AP1roG ansatz in modelling the ground-state electronic structure of small actinide compounds by comparing it to standard quantum chemistry approaches. Our study of the ground state spectroscopic constants (bond lengths and vibrational frequencies) and potential energy curves of actinide oxides (UO_2^{2+} and ThO_2) as well as the energetic stability of ThC_2 isomers reveals that voo-AP1roG accurately describes the electronic structure of heavy-element compounds. Specifically, AP1roG provides qualitatively correct potential energy surfaces for the dissociation of UO_2^{2+} and ThO_2, while conventional methods fail, and predicts spectroscopic constants that are in good agreement with CCSD(T), the gold standard of quantum chemistry. Our study further highlights the failure of conventional multi-reference methods like CASSCF and suggests that 12-orbital active spaces are unbalanced active spaces when the U–O bonds are stretched. To sum up, our first work on small actinide complexes illustrates the good performance of AP1roG for describing the electronic structure of (closed-shell) actinide compounds and motivates computational studies on larger, more realistic actinide-containing materials. The advantage of AP1roG over conventional multi-reference methods is its cheap computational cost that allows AP1roG to be easily applied to larger molecular systems.

5.2 Qualitative interpretation of electronic wave functions using concepts of quantum information theory

The interaction of orbitals is a useful concept in chemistry. It is frequently used to understand chemical processes and reaction mechanisms. Unfortunately, the interaction of orbitals is commonly understood using qualitative arguments, like molecular-orbital diagrams, Frontier-orbital theory, and ligand field theory. We are developing quantitative means to measure the interaction of orbitals using concepts of quantum information theory. Specifically, we have shown that orbital entanglement and correlation are particularly useful and intuitive measures to quantify the interaction of orbitals and to elucidate electronic structures and changes in electronic structure that accompany chemical processes. For instance, orbital entanglement and correlation can be used to predict bond orders, identify transition states, and dissect electron correlation effects.

Katerina Boguslawski

5.2.1 Extracting bond order from entropy measures

The chemical bond is an important local concept to understand chemical compounds and processes. Unfortunately, like most local concepts, the chemical bond and the bond order do not correspond to any physical observable and thus cannot be determined as an expectation value of a quantum chemical operator. In **[P11]**, we elaborate on how concepts of quantum information theory can be applied to study chemical processes and to redefine bond orders. We demonstrate that the one- and two-orbital entropy measures can be utilized to monitor bond-breaking and—equivalently—bond-forming processes. Upon dissociation of a chemical bond, the bonding and antibonding molecular orbitals associated with the bond of interest become strongly entangled. Hence, the corresponding single-orbital entropies gradually increase if two atoms are pulled apart. Moreover, the entanglement analysis resolves the bond breaking of different bond types (σ, π, etc.) individually in multi-bonded centers. An entropy-based bond order can thus be deduced from the single-orbital entropy diagram. As molecules with prototypical bonds, we have investigated the dissociation process of the diatomic molecules N_2 (triple bond), F_2, and CsH (both single bond), which represent characteristic examples of single- or multi-reference problems.
In a follow-up publication **[P9]**, we demonstrate that the orbital entanglement and correlation analysis can be extended to polyatomic molecules to understand chemical bonding. Specifically, we apply our entropy-based bond order analysis to carbon–carbon, silicon–silicon, and carbon–phosphorus centers. Our approach correctly reproduces bond multiplicities in simple polyatomic molecules like ethane, ethene, and acetylene and confirms the triple bonding between the carbon–phosphorus centers in the $[CP]^-$ and HCP molecules. The behavior of phosphorus in these prototypical phosphoalkynes closely resembles the bonding situation in the N_2 and C_2H_2 molecules. Furthermore, our analysis confirms that the nature of the chemical bond in the C_2 molecule is far more complicated than for their higher substituted analogs like C_2H_6, C_2H_4, and C_2H_2. Stretching the C–C bond in C_2 does not yield a gradually increasing entanglement and correlation pattern, which is an indication of the complexity of chemical bonding in the carbon dimer.

5.2.2 Monitoring bond-formation processes along the reaction coordinate

We extended our entropy-based bond-order analysis presented in **[P9,P11]** to dissect bond-formation processes in metal-driven catalysis along the reaction coordinate. Similar to the procedure outlined above, we exploit the entanglement and correlation among molecular orbitals to analyze changes in electronic structure that accompany chemical processes. Specifically, we demonstrate how the orbital-pair mutual information can be used to monitor bond-formation processes **[P6]** and to identify points along the reaction coordinate when chemical bonds are formed and broken. As a proof of principle example, we investigated the nickel–ethene complexation reaction, where the metal–olefin bond is made possible through metal-to-ligand back-donation. This example allows us to monitor the interplay of back-bonding and π-donation along the reaction coordinate. Our orbital entanglement and correlation analysis shows that metal-ligand bonding is initialized by back-donation which establishes around the transition state. This back-bonding then entails π-donation from the ethene ligand

Katayun Boguslawski

to the metal center. To conclude, our study supports the crucial role of metal-to-ligand back-donation in the bond-forming process of nickel–ethene.
In a follow-up work **[P4]**, we studied the bonding mechanism of ethene to a nickel or palladium center using different quantum chemistry methods (DFT, CASSCF, and DMRG). Specifically, we focus on the interaction between the metal atom and bis-ethene ligands in perpendicular and parallel orientations. The bonding situation in these structural isomers is further scrutinized using an energy decomposition analysis and our orbital entanglement and correlation analysis. In particular, the orbital-pair mutual information and single-orbital entropy highlight the fact that when two ethene ligands are oriented perpendicular to each other, the complex is stabilized by the metal-to-ligand double-back-bonding mechanism. Finally, we demonstrate that nickel–ethene compounds feature a stronger and more covalent interaction between the ligands and the metal center than palladium–ethene complexes with similar coordination spheres.

5.2.3 Automatic selection of active orbital spaces: correlation and unrestricted natural orbital criteria

Conventional multi-reference methods can treat active spaces that are often at the upper limit of what is required for a proper treatment of species with complex electronic structures due to the large number of degenerate or quasi-degenerate electronic states, leaving no room for verifying their suitability. In **[P1]**, we address the issue of properly defining active orbital spaces and introduce a protocol to determine optimal active spaces based on the use of the DMRG algorithm and the orbital-pair mutual information. Specifically, we suggest to construct the active orbital space based on the distribution of the orbital-pair mutual information. Our selection criteria provide optimal active spaces that allow for an accurate description of static/non-dynamic electron correlation. In our approach, the active space should include all orbital pairs that are strongly correlated. Starting from an (unconverged) reference calculation containing a large active space, an optimal active space can then be defined by only selecting the strongly correlated orbitals from this large active space so that the optimal active space calculation reproduces the orbital-pair correlation diagram of the reference calculation. This recipe has been used to define active orbitals spaces in larger actinide-containing compounds, like $PuO_2(OH)_2$ and actinyl dications, where multi-configurational reference calculations were not available in the literature. Most importantly, this selection procedure facilitates black-box active space calculations, where no *a priori* knowledge of the electronic structure of the molecule under study is required. Note that the DMRG reference calculation does not have to be fully converged as already 4–6 sweeps of the DMRG algorithm yield sufficiently accurate correlation diagrams that can be used in the automatic selection of active orbital spaces.
Finally, together with the group of Prof. Pulay, we have assessed the performance of the unrestricted natural orbital (UNO) criterion to construct active orbital spaces in a black-box fashion **[P5]**. The UNO criterion has been rigorously tested against DMRG reference calculations, where active orbital occupancies in UNO-CAS and CASSCF calculations are benchmarked against DMRG reference values in a number

Katayun Boguslaw

of strongly correlated molecules (F_2, ozone, NO_2, polyenes, naphthalene, azulene, anthracene, nitrobenzene, phenoxy, benzyl, o-, m-, and p-benzyne, nickel–acetylene, and Cr_2). Our results suggest that the UNO criterion works well in all these cases.

5.3 Further development of electronic structure codes

Besides the above presented wave function models, we have developed and implemented additional AP1roG-based approaches into the PIERNIK program suite. Specifically, we extended existing modules to include

- a CCS correction on top of AP1roG/pCCD (pCCD-CCS)
- the EOM-CCS/CIS, EOM-pCCD-CCS, EOM-LCCSD, and EOM-pCCD-LCCSD models
- the linear-response module based on EOM-pCCD, EOM-pCCD+S, and EOM-pCCD-CCS
- different 1-electron integrals (electric dipole moment integrals)

5.4 List of publications not included in section 4, after PhD studies

[P1] **K. Boguslawski**, F. Real, P. Tecmer, C. Duperrouzel, A. S. P. Gomes, Ö. Legeza, P. W. Ayers, V. Vallet, *On the multi-reference nature of plutonium oxides: PuO_2^{2+}, PuO_2, PuO_3 and $PuO_2(OH)_2$*, *Phys. Chem. Chem. Phys.* 19 **2017**, 4317-4329 (**corresponding author**).

[P2] P. Tecmer, S. W. Hong, **K. Boguslawski**, *Dissecting the cation-cation interaction between two uranyl units*, *Phys. Chem. Chem. Phys.* 18 **2016**, 18305-18311.

[P3] D. Stuart, P. Tecmer, P. W. Ayers, **K. Boguslawski**, *The Effect of Nitrido, Azide, and Nitrosyl Ligands on Magnetization Densities and Magnetic Properties of Iridium PNP Pincer-Type Complexes*, *RSC Adv.* 5 **2015**, 84311-84320 (**corresponding author**).

[P4] Y. Zhao, **K. Boguslawski**, P. Tecmer, C. Duperrouzel, G. Barcza, O. Legeza, P. W. Ayers, *Dissecting the bond formation process of d^{10}-metal-ethene complexes with multireference approaches*, *Theor. Chem. Acc.* 134 **2015**, 120 (**corresponding author**).

[P5] S. Keller, **K. Boguslawski**, T. Janowski, M. Reiher, P. Pulay, *Selection of active spaces for multiconfigurational wavefunctions*, *J. Chem. Phys.* 142 **2015**, 244104.

[P6] C. Duperrouzel, P. Tecmer, **K. Boguslawski**, G. Barcza, O. Legeza, P. W. Ayers, *A quantum informational approach for dissecting chemical reactions*, *Chem. Phys. Lett.* 621 **2015**, 160-164 (**corresponding author**).

Katarzyna Boguslawski

[P7] P. Tecmer, **K. Boguslawski**, P. W. Ayers, *"Singlet ground-state actinide chemistry with geminals"*, *Phys. Chem. Chem. Phys.* 17 **2015**, 14427-14436.

[P8] P. Tecmer, **K. Boguslawski**, P. A. Johnson, P. A. Limacher, M. Chan, T. Verstraelen, P. W. Ayers, *"Assessing the accuracy of new geminal-based approaches"*, *J. Phys. Chem. A* 118 **2014**, 9058-9068.

[P9] M. Mottet, P. Tecmer, **K. Boguslawski**, O. Legeza, M. Reiher, *Quantum entanglement in carbon–carbon, carbon–phosphorus and silicon–silicon bonds*, *Phys. Chem. Chem. Phys.* 16 **2014**, 8872-8880 (**corresponding author**).

[P10] P. Tecmer, **K. Boguslawski**, O. Legeza, M. Reiher, *Unraveling the quantum-entanglement effect of noble gas coordination on the spin ground state of CUO*, *Phys. Chem. Chem. Phys.* 16 **2014**, 719-727 (**corresponding author**).

[P11] **K. Boguslawski**, P. Tecmer, G. Barcza, O. Legeza, M. Reiher, *Orbital entanglement in bond-formation processes*, *J. Chem. Theory Comput.* 9 **2013**, 2959-2973.

5.5 Biometric data

(As of March 19, 2018)	
h-index:	14
number of citations:	516
(without self-citations)	(390)
number of published papers:	28

Katarzyna Boguslawska

Bibliography

[1] L. Gagliardi and B. O. Roos, *Chem. Soc. Rev.*, 2007, **36**, 893.

[2] D. Wang, W. F. van Gunsteren and Z. Chai, *Chem. Soc. Rev.*, 2012, **41**, 5836–5865.

[3] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.*, 2007, **79**, 291–350.

[4] J. Olsen, P. Jørgensen and J. Simons, *Chem. Phys. Lett.*, 1990, **169**, 463–472.

[5] C. Clavaguéra-Sarrio, V. Vallet, D. Maynau and C. J. Marsden, *J. Phys. Chem.*, 2004, **121**, 5312–5321.

[6] P. Tecmer, A. S. P. Gomes, S. Knecht and L. Visscher, *J. Chem. Phys.*, 2014, **141**, 041107.

[7] B. Roos and P. R. Taylor, *Chem. Phys.*, 1980, **48**, 157–173.

[8] J. Olsen, B. O. Roos, P. Jørgensen and H. J. A. Jensen, *J. Chem. Phys.*, 1988, **89**, 2185–2192.

[9] P.-A. Malmqvist, A. Rendell and B. O. Roos, *J. Phys. Chem.*, 1990, **94**, 5477–5482.

[10] S. Matsika, Z. Zhang, S. R. Brozell, J.-P. Blaudeau, Q. Wang and R. M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 3825–3828.

[11] D. I. Lyakh, M. Musiał, V. F. Lotrich and . J. Bartlett, *Chem. Rev.*, 2012, **112**, 182–243.

[12] K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Woliński, *J. Phys. Chem.*, 1990, **94**, 5483–5488.

Katarzyna Boguslawski

[13] K. Andersson, P.-A. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218–1226.

[14] B. O. Roos, in *Radiation Induced Molecular Phenomena in Nucleic Acids*, ed. M. K. Shukla and J. Leszczynski, Springer, 2008, pp. 125–156.

[15] P.-A. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer and L. Gagliardi, *J. Chem. Phys.*, 2008, **128**, 204109.

[16] L. Gagliardi and B. O. Roos, *Chem. Phys. Lett.*, 2000, **331**, 229.

[17] K. Pierloot and E. van Besien, *J. Phys. Chem.*, 2005, **123**, 204309.

[18] K. Pierloot, E. van Besien, E. van Lenthe and E. J. Baerends, *J. Chem. Phys.*, 2007, **126**, 194311.

[19] F. Ruipérez, C. Danilo, F. Réal, J.-P. Flament, V. Vallet and U. Wahlgren, *J. Phys. Chem. A*, 2009, **113**, 1420–1428.

[20] F. Réal, V. Vallet, C. Marian and U. Wahlgren, *J. Phys. Chem.*, 2007, **127**, 214302.

[21] T. Fleig, H. J. A. Jensen, J. Olsen and L. Visscher, *J. Chem. Phys.*, 2006, **124**, 104106.

[22] G. L. Macchia, I. Infante, J. Raab, J. K. Gibson and L. Gagliardi, *Phys. Chem. Chem. Phys.*, 2008, **48**, 7278.

[23] S. Spencer, L. Gagliardi, N. C. Handy, A. G. Ioannou, C.-K. Skylaris and A. Willetts, *J. Phys. Chem. A*, 1999, **103**, 1831.

[24] F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet and E. Eliav, *J. Phys. Chem. A*, 2009, **113**, 12504–12511.

[25] I. Infante, L. Andrews, X. Wang and L. Gagliardi, *Chem. Eur. J.*, 2010, **43**, 12804–12807.

[26] B. O. Roos, P.-O.Widmark and L. Gagliardi, *Faraday Discuss.*, 2003, **124**, 57–62.

[27] L. Gagliardi and B. Roos, *Nature*, 2005, **433**, 848–851.

[28] P. Tecmer, A. S. P. Gomes, U. Ekström and L. Visscher, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6249–6259.

[29] X. Wang, L. Andrews, P.-A. Malmqvist, B. O. Roos, A. P. Gonçalves, C. C. L. Pereira, J. Marçalo, C. Godart, and B. Villeroy, *J. Am. Chem. Soc.*, 2010, **132**, 8484.

[30] A. Kovacs and R. J. M. Konings, *J. Phys. Chem. A*, 2012, **115**, 6646–6656.

[31] T. Yang, R. Tyagi, Z. Zhang and R. M. Pitzer, *Mol. Phys.*, 2009, **107**, 1193–1195.

Katayun Bagheri

[32] P. Pogány, A. Kovács, Z. Varga, F. M. Bickelhaupt and J. Rudy, *J. Phys. Chem. A*, 2012, **116**, 747.

[33] S. R. White, *Phys. Rev. Lett.*, 1992, **69**, 2863–2866.

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

[35] K. H. Marti and M. Reiher, *Z. Phys. Chem.*, 2010, **224**, 583–599.

[36] O. Legeza, R. Noack, J. Sólyom and L. Tincani, in *Computational Many-Particle Physics*, ed. H. Fehske, R. Schneider and A. Weiße, Springer, Berlin/Heidelberg, 2008, vol. 739, pp. 653–664.

[37] S. R. White and R. L. Martin, *J. Chem. Phys.*, 1999, **110**, 4127–4130.

[38] S. Wouters and D. Van Neck, *Eur. Phys. J. D*, 2014, **68**, 272.

[39] J. Hachmann, W. Cardoen and G. Chan, *J. Chem. Phys.*, 2006, **125**, 141101.

[40] K. H. Marti, I. M. Ondik, G. Moritz and M. Reiher, *J. Chem. Phys.*, 2008, **128**, 014104.

[41] Y. Kurashige and T. Yanai, *J. Chem. Phys.*, 2009, **130**, 234114.

[42] K. Boguslawski, K. H. Marti, O. Legeza and M. Reiher, *J. Chem. Theory Comput.*, 2012, **8**, 1970–1982.

[43] K. Boguslawski, P. Tecmer, O. Legeza and M. Reiher, *J. Phys. Chem. Lett.*, 2012, **3**, 3129–3135.

[44] Y. Kurashige, G. K.-L. Chan and T. Yanai, *Nature Chem.*, 2013, **5**, 660–666.

[45] S. Wouters, P. A. Limacher, D. Van Neck and P. W. Ayers, *J. Chem. Phys.*, 2012, **136**, 134110.

[46] S. Wouters, W. Poelmans, P. W. Ayers and D. V. Neck, *Comput. Phys. Commu.*, 2014, **185**, 1501–1514.

[47] P. Tecmer, K. Boguslawski, O. Legeza and M. Reiher, *Phys. Chem. Chem. Phys*, 2014, **16**, 719–727.

[48] P. R. Surjan, in *Correlation and Localization*, Springer, 1999, pp. 63–88.

[49] W. Kutzelnigg, *Theoret. Chim. Acta*, 1965, **3**, 241–253.

[50] R. McWeeny and B. T. Sutcliffe, *Proc. R. Soc. Lond. A*, 1963, **273**, 103–116.

[51] P. Cassam-Chenaï, *J. Chem. Phys.*, 2006, **124**, 194109.

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

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

Katarzyna Boguslawska

[54] P. R. Surján, A. Szabados, P. Jeszenszki and T. Zoboki, *J. Math. Chem.*, 2012, **50**, 534–551.

[55] T. Zoboki, A. Szabados and P. R. Surjan, *J. Chem. Theory Comput.*, 2013, **9**, 2602–2608.

[56] K. Pernal, *J. Chem. Theory Comput.*, 2014, **10**, 4332–4341.

[57] P. Jeszenszki, P. R. Nagy, T. Zoboki, Á. Szabados and P. R. Surján, *Int. J. Quantum Chem.*, 2014, **114**, 1048–1052.

[58] E. Pastorczak and K. Pernal, *Phys. Chem. Chem. Phys.*, 2015, **17**, 8622–8626.

[59] A. J. Coleman, *J. Math. Phys.*, 1965, **6**, 1425–1431.

[60] A. J. Coleman, *Int. J. Quantum Chem.*, 1997, **63**, 23–30.

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

[62] C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi and G. E. Scuseria, *J. Chem. Phys.*, 2012, **136**, 164109.

[63] S. Bratoz and P. Durand, *J. Chem. Phys.*, 1965, **43**, 2670–2679.

[64] D. M. Silver, *J. Chem. Phys.*, 1969, **50**, 5108–5116.

[65] D. M. Silver, *J. Chem. Phys.*, 1970, **52**, 299–303.

[66] G. Náray-Szabó, *J. Chem. Phys.*, 1973, **58**, 1775–1776.

[67] G. Náray-Szabó, *Int. J. Qunatum Chem.*, 1975, **9**, 9–21.

[68] P. R. Surján, *Phys. Rev. A*, 1984, **30**, 43–50.

[69] P. R. Surján, *Phys. Rev. A*, 1985, **32**, 748–755.

[70] P. R. Surján, *Int. J. Quantum Chem.*, 1994, **52**, 563–574.

[71] P. R. Surján, *Int. J. Quantum Chem.*, 1995, **55**, 109–116.

[72] E. Rosta and P. R. Surján, *Int. J. Quantum Chem.*, 2000, **80**, 96–104.

[73] E. Rosta and P. R. Surján, *J. Chem. Phys.*, 2002, **116**, 878–889.

[74] A. C. Hurley, J. Lennard-Jones and J. A. Pople, *Proc. R. Soc. Lond. A*, 1953, **220**, 446–455.

[75] W. A. Goddard and A. Amos, *Chem. Phys. Lett.*, 1972, **13**, 30–35.

[76] W. A. Goddard, T. H. Dunning Jr., W. J. Hunt and P. J. Hay, *Acc. Chem. Res.*, 1973, **6**, 368–376.

[77] D. W. Small, K. V. Lawler and M. Head-Gordon, *J. Chem. Theory Comput.*, 2014, **10**, 2027–2040.

Katarzyna Bąkowska

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

[79] T. M. Henderson, G. E. Scuseria, J. Dukelsky, A. Signoracci and T. Duguet, *Phys. Rev. C*, 2014, **89**, 054305.

[80] P. R. Surján, A. Szabados, P. Jeszenszki and T. Zoboki, *J. Math. Chem.*, 2012, **50**, 534–551.

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

[82] P. Tecmer, K. Boguslawski, P. A. Limacher, P. A. Johnson, M. Chan, T. Verstraelen and P. W. Ayers, *J. Phys. Chem. A*, 2014, **118**, 9058–9068.

[83] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, Chichester, 2000.

[84] J. Pian and C. S. Sharma, *J. Phys. A: Math. Gen.*, 1981, **14**, 1261–1270.

[85] B. Levy and G. Berthier, *Int. J. Quantum Chem.*, 1968, **11**, 307–319.

[86] T. C. Chang and W. H. E. Schwarz, *Theor. Chim. Acta*, 1977, **44**, 45–59.

[87] L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis and G. E. Scuseria, *J. Chem. Phys.*, 2011, **135**, 044119.

[88] P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck and P. Bultinck, *Mol. Phys.*, 2014, **112**, 853–862.

[89] T. Stein, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2014, **140**, 214113.

[90] J. Hachmann, W. Cardoen and G. K.-L. Chan, *J. Chem. Phys.*, 2006, **125**, 144101.

[91] T. Tsuchimochi and G. E. Scuseria, *J. Chem. Phys.*, 2009, **131**, 121102.

[92] L. Stella, C. Attaccalite, S. Sorella and A. Rubio, *Phys. Rev. B*, 2011, **84**, 245117.

[93] N. Lin, C. A. Marianetti, A. J. Millis and D. R. Reichman, *Phys. Rev. Lett.*, 2011, **106**, 096402.

[94] K. Bhaskaran-Nair, O. Demel and J. Pittner, *J. Chem. Phys.*, 2008, **129**, 184105.

[95] D. W. Whitman and B. K. Carpenter, *J. Am. Chem. Soc.*, 1982, **104**, 6473–6474.

[96] F. A. Evangelista, E. Prochnow, J. Gauss and H. F. Schaefer III, *J. Chem. Phys.*, 2010, **132**, 074107.

Katarzyna Bytautas

[97] E. Prochnow, F. A. Evangelista, H. F. Schaefer III, W. D. Allen and J. Gauss, *J. Chem. Phys.*, 2009, **131**, 064109.

[98] J. Rissler, R. M. Noack and S. R. White, *Chem. Phys.*, 2006, **323**, 519–531.

[99] G. Barcza, O. Legeza, K. H. Marti and M. Reiher, *Phys. Rev. A*, 2011, **83**, 012508.

[100] Generalized Correlation Functions and Entanglement Patterns, 2013.

[101] K. Boguslawski, P. Tecmer, G. Barcza, O. Legeza and M. Reiher, *J. Chem. Theory Comput.*, 2013, **9**, 2959–2973.

[102] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider and Ö. Legeza, *Int. J. Quantum Chem.*, 2015, –.

[103] G. Barcza, R. Noack, J. Sólyom and Ö. Legeza, *Phys. Rev. B*, 2014, **92**, 125140.

[104] M. Mottet, P. Tecmer, K. Boguslawski, O. Legeza and M. Reiher, *Phys. Chem. Chem. Phys.*, 2014, **16**, 8872–8880.

[105] T. Szilvasi, G. Barcza, O. Legeza and J. Sólyom, *arXiv*, 2015, **X**, 1509.04241.

[106] K. Boguslawski and M. Reiher, *The Chemical Bond: Chemical Bonding Across the Periodic Table*, 2014, 219–252.

[107] L. Freitag, S. Knecht, S. F. Keller, M. G. Delcey, F. Aquilante, T. B. Pedersen, R. Lindh, M. Reiher and L. Gonzalez, *Phys. Chem. Chem. Phys.*, 2015, **17**, 13769–13769.

[108] P. Tecmer, K. Boguslawski and P. W. Ayers, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14427–14436.

[109] Y. Zhao, K. Boguslawski, P. Tecmer, C. Duperrouzel, G. Barcza, Ö. Legeza and P. W. Ayers, *Theor. Chem. Acc.*, 2015, **134**, 120.

[110] V. Murg, F. Verstraete, R. Schneider, P. R. Nagy and Ö. Legeza, *J. Chem. Theory Comput.*, 2015, **11**, 1027–1036.

[111] E. Fertitta, B. Paulus, G. Barcza and Ö. Legeza, *Phys. Rev. B*, 2014, **90**, 245129.

[112] C. Duperrouzel, P. Tecmer, K. Boguslawski, G. Barcza, O. Legeza and P. W. Ayers, *Chem. Phys. Lett.*, 2015, **621**, 160–164.

[113] S. Keller, K. Boguslawski, T. Janowski, M. Reiher and P. Pulay, *J. Chem. Phys.*, 2015, **142**, 244104.

[114] C. J. Stein and M. Reiher, *J. Chem. Theory Comput.*, 2016, **12**, 1760–1771.

[115] Ö. Legeza and J. Sólyom, *Phys. Rev. B*, 2003, **68**, 195116.

[116] Ö. Legeza and J. Sólyom, *Phys. Rev. Lett.*, 2006, **96**, 4–7.

Kostya Boguslawski

[117] F. Weinhold and E. B. Wilson Jr., *J. Chem. Phys.*, 1967, **46**, 2752–2758.

[118] E. H. Lieb and F. Y. Wu, *Phys. Rev. Lett.*, 1968, **20**, 1445–1448.

[119] P. Limacher, P. Ayers, P. Johnson, S. De Baerdemacker, D. Van Neck and P. Bultinck, *Phys. Chem. Chem. Phys*, 2014, **16**, 5061–5065.

[120] Z. Tóth, P. R. Nagy, P. Jeszenszki and A. Szabados, *Theor. Chem. Acc.*, 2015, **134**, 100.

[121] A. J. Garza, I. W. Bulik, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2015, **142**, 044109.

[122] A. J. Garza, I. W. Bulik, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2015, **17**, 22412–22422.

[123] P. Pulay, *Int. J. Quantum Chem.*, 2011, **111**, 3273–3279.

[124] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, Cambridge University Press, Cambridge, 2009.

[125] K. A. Peterson, *J. Chem. Phys.*, 1995, **102**, 262–277.

[126] K. A. Peterson, R. A. Kendall and T. H. Dunning, *J. Chem. Phys.*, 1993, **99**, 9790–9805.

[127] K. A. Peterson and T. H. Dunning, *J. Mol. Struct. THEOCHEM*, 1997, **400**, 93–117.

[128] J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.*, 1993, **98**, 7029–7039.

[129] J. Jankowski, K. Kowalski and P. Jankowski, *Chem. Phys. Lett.*, 1994, **222**, 608–614.

[130] T. M. Henderson, I. W. Bulik, T. Stein and G. E. Scuseria, *J. Chem. Phys.*, 2014, **141**, 244104.

[131] J. B. Foresman, M. Head-Gordon, J. A. Pople and M. J. Frisch, *J. Phys. Chem.*, 1992, **96**, 135–149.

[132] B. Hudson and B. Kohler, *Chem. Phys. Lett.*, 1972, **14**, 299–304.

[133] K. Schulten, I. Ohmine and M. Karplus, *J. Chem. Phys.*, 1976, **64**, 4422–4441.

[134] R. J. Cave and E. R. Davidson, *J. Chem. Phys.*, 1987, **91**, 4481–4490.

[135] P. Tavan and K. Schulten, *Phys. Rev. B*, 1987, **36**, 4337–4358.

[136] R. J. Cave and E. R. Davidson, *J. Chem. Phys.*, 1988, **92**, 614–620.

[137] W. J. Buma, B. E. Kohler and K. Song, *J. Chem. Phys.*, 1991, **94**, 6367–6376.

[138] G. Orlandi, F. Zerbetto and M. Z. Zgierski, *Chem. Rev.*, 1991, **91**, 867–891.

Katarzyna Bączek

[139] L. Serrano-Andres, M. Merchan, I. Nebot-Gil, R. Lindh and B. O. Roos, *J. Chem. Phys.*, 1993, **98**, 3151–3162.

[140] L. Serrano-Andres, R. Lindh, B. O. Roos and M. Merchan, *J. Chem. Phys.*, 1993, **97**, 9360–9368.

[141] R. P. Krawczyk, K. Malsch, G. Hohlneicher, R. C. Gillen and W. Domcke, *Chem. Phys. Lett.*, 2000, **320**, 535–541.

[142] C.-P. Hsu, S. Hirata and M. Head-Gordon, *J. Phys. Chem. A*, 2001, **105**, 451–458.

[143] K. B. Wiberg, A. E. de Oliveira, and G. Trucks, *J. Phys. Chem. A*, 2002, **106**, 4192–4199.

[144] J. H. Starcke, M. Wormit, J. Schirmer and A. Dreuw, *Chem. Phys.*, 2006, **329**, 39–49.

[145] R. L. Christensen, M. G. I. Galinato, E. F. Chu, J. N. Howard, R. D. Broene and H. A. Frank, *J. Phys. Chem. A*, 2008, **112**, 12629–12636.

[146] Y. Kurashige, H. Nakano, Y. Nakao and K. Hirao, *Chem. Phys. Lett.*, 2004, **400**, 425–429.

[147] K. Nakayama, H. Nakano and K. Hirao, *Int. J. Quantum Chem.*, 1998, **66**, 157–175.

[148] W. Hu and G. K.-L. Chan, *J. Chem. Theory Comput.*, 2015, **11**, 3000–3009.

[149] B. E. Kohler, *J. Chem. Phys.*, 1988, **88**, 2788–2792.

[150] M. Reiher and A. Wolf, *Relativistic Quantum Chemistry. The Fundamental Theory of Molecular Science*, Wiley, 2009.

[151] R. G. Denning, *J. Phys. Chem. A*, 2007, **111**, 4125–4143.

[152] V. Vallet, U. Wahlgren and I. Grenthe, *J. Phys. Chem. A*, 2012, **115**, 12373–12380.

[153] R. G. Denning, *Struct. Bonding*, 1992, **79**, 215–276.

[154] W. A. de Jong, L. Visscher and W. C. Nieuwpoort, *J. Mol. Struct. THEOCHEM*, 1999, **458**, 41.

[155] P. Tecmer, N. Govind, K. Kowalski, W. A. de Jong. and L. Visscher, *J. Chem. Phys*, 2013, **xx**, xx.

[156] Z. Zhang and R. M. Pitzer, *J. Phys. Chem. A*, 1999, **103**, 6880.

[157] P. Tecmer, R. Bast, K. Ruud and L. Visscher, *J. Phys. Chem. A*, 2012, **116**, 7397–7404.

Katarzyna Boguslawski