Final report on the OTKA/NKFI project K 115744 entitled "Methodological developments in many-electron theory" conducted in the period 2015.09.01–2019.11.30

1 Brief overview

Investigation of open problems in theoretical chemistry was the broad purpose of the project, with several topics mentioned as challenges to be faced. In accordance with the proposal, geminal based methodologies and correction schemes built upon geminal wavefunctions has been a central research task. As envisaged in the proposal, the element of unplanned novelty has been present during the project, in the form of studies originating in newly surfaced ideas. Many of these mini projects were related to perturbation theory, a long term research interest of the senior investigators (SI) of the project. Some studies performed in the framework of the project do not fit any of the major research lines mentioned in the proposal. This exemplifies the problem driven research attitude of the SIs and explains the diverse topics appearing in a report summarizing four years' work.

Apart from the two SIs, four students have been taking part in the research for shorter or longer periods of time. Two students obtained a BSc degree, one student graduated at the MSc level and the fourth one received her PhD degree during the time of the project. The current PhD students are expected to graduate in 2020 and 2021, respectively.

The final publication record of the project, referencing the OTKA/NKFI grant includes 11 scientific papers, published in international, peer-reviewed journals, 2 summary chapters in the Reference Module in Chemistry, Molecular Sciences and Chemical Engineering maintained by Elsevier and 19 presentations (lecture or poster) at international scientific meetings. Two manuscripts under preparation indicate that only the OTKA/NKFI project is closed, the work carries on. The research performed in the framework of the project can be considered a continuation of a K type OTKA grant of the principal investigator (PI), in the time interval 2010-2014. Studies that have got published in 2015 appeared between the two granted periods. These five additional papers are included in the list of publication for their subject being related to the goals of the proposal. The fact, that these papers lack reference to the OTKA/NKFI grant are duly indicated. The PI has been awarded the Bolyai fellowship of the Hungarian Academy of Sciences in 2015, that is acknowledged in publications beside the OTKA/NKFI grant. Grants termed ÚNKP (Új Nemzeti Kiválóság Program) received by the students are similarly acknowledged.

The four years of the project has naturally brought about some shift of focus as compared to the original workplan. E.g. partial spin projection has been favoured over efficient implementation of the full spin projected then varied unrestricted geminal Ansatz. Or, ring coupled-cluster doubles theory has been pursued instead of the planned diagrammatic formulation of the linearized coupled-cluster doubles methodology, built upon a geminal product reference. We consider these alterations as minor changes characteristic of scientific research and not interfering with the declaration that the reseach plan has been fulfilled.

Summarizing the main achievements of the project in a nutshell, exploration of the spin projected variants of the unrestricted geminal product wavefunction has to be stressed. Interest in geminal based wavefunction lies with their advantageous computational cost together with their potential to serve as reference for incorporating electron correlation often needed for quantitative accuracy. We have pinpointed the defective fragment spin in multiple covalent bond dissociation processes described by the singlet geminals' product wavefunction. This unfortunately rules out the Antisymmetrized Product of Strongly orthogonal Geminals (APSG) wavefunction as a generally applicable reference. Geminal Ansätze built with singlet/triplet mixed geminals are capable of correcting the defect, they however introduce spin contamination. With the aim of examining suitability of such, unrestricted geminal Ansätze as reference we have worked out and performed numerical tests with a perturbation theory (PT) based correlation correction. Our results on singlet-triplet gaps of biradical systems have revealed that significant spin contamination in the reference undermines performance of the subsequent correlation treatment. This points at the need for correcting the spin of the reference function. Investigation of full spin projection followed by variation of the unrestricted geminal Ansatz has been performed with this impetus, revealing substantial size-consistency violation. The latter property being highly adverse, a compromise method has been investigated in the form of half-projection. The magnitude of size-inconsistency has been found to beneficially diminish for the half-projected Ansatz at the expense of a persisting spin-contamination. It remains to be seen whether spin impurity of the half-projected Ansatz is negligible enough to provide a suitable reference for correlation treatment, resulting in a description of the low lying electronic states of biradicals correct at the quantitative level.

Correction schemes built upon an initial, reference function constitute another main chapter of the research perfomed. In this respect our newly devised tools for the resummation of divergent perturbation series have to be mentioned. Generalization of the ring approximation of coupled-cluster theory for the multireference case is another notable achievement, exhibiting a relatively moderate, $\mathcal{O}(N^6)$ scaling applicable not only to APSG but also for the half-projected and full-projected geminal Ansatz.

In the following we give a more detailed account of the work performed, grouped according to research topic.

2 Results at a closer look

2.1 Geminal wavefunctions

One of the simplest specimen of the so-called strongly orthogonal geminal product involves singlet geminals, giving the APSG wavefunction. Though potential energy surfaces (PES) obtained by APSG appear well behaving for covalent bond breaking, the good looking PES conceals a conceptual problem in the case of multiple bond breaking, which lies with incorrect spin of the dissociated fragments. The problem becomes manifest when e.g. developing PT based on APSG, possibly resulting divergence of the correction scheme. We diagnosed the lack of certain elements of the configuration interaction (CI) space at order zero as the source of the effect and included

the necessary components thereby achieving a proper description of the process.[1]

The obvious explanation for the failure of APSG in multiple bond breaking is the lack of triplet geminals coupling to singlet becoming an important part of the wavefunction in the dissociated limit. The improvement achieved in fragment spin by incorporating triplet geminals has been demonstrated in a joint numerical study with Vitaly Rassolov (University of South Carolina, USA).[2] The downside of allowing for triplet geminals is the appearance of spin contamination that can in turn be corrected by spin projection. In a study examining spin projection followed by variation we reported transparent formulae on the relationship between the unrestricted Hartree-Fock (HF) model and strongly orthogonal geminals with singlet/triplet mixed two-electron units. Spin projection of the former leads to the long known spin projected Extended HF (EHF) wavefunction. The other, we termed spin projected Extended APSG (EAPSG). Formal and numerical studies revealed that while EHF violates size consistency in case of already two interacting geminals, EAPSG is size consistent if not more that three interacting geminals are present. Size consistency violation appears also for EAPSG for four or more interacting geminals. At the same time, EAPSG interpolates smoothly between the weakly and strongly correlated geminals' regime, while EHF and APSG fail at the former or the latter, respectively.[3]

A compromise between spin contaminated singlet/triplet mixed geminals' product and spin pure but size inconsistent EAPSG is provided by partial spin projection. Equations for determining the parameters of the half-projected geminal Ansatz based on the variation principle have been derived and implemented. Our first results are according to expectations: considerably (but not completely) purified total spin is accompanied by slight (an order of magnitude smaller than EAPSG) but nonzero size inconsistency.[4]

In response to an invitation of the Elsevier publishing house, a review article was written on the topic of geminal wavefunctions.[5]

2.2 Correction schemes based on geminal wavefunctions

Devising correction schemes to geminal type wavefunctions has been a principal aim of the plan. The breakdown of APSG based linearized coupled cluster (LCC) correction, describen in Section 2.1. was in fact the origin of exploring triplet including geminal models. Knowing of the failure of LCC the lack of ill effect with the extended random phase approximation (ERPA) based correction developed for APSG by Katarzyna Pernal (Łódź University of Technology, Poland) was surprising. We set out to examine the origin of the qualitatively different behaviour of the two methods. A comparative study performed in cooperation with the group of Pernal revealed that the ERPA correction has to be computed with special attention to instabilities in order to avoid discontinuity in the range where the LCC correction breaks down. In fact, exclusion of individual single excitations from the excitation operator lying at the heart of ERPA is needed in order to avoid deficient PES. When computed properly, ERPA provides qualitatively correct PES, it however brings no improvement upon the spin coupling of dissociation fragments (essentially wrong by APSG). At difference with this, LCC has the potential to correct on spin coupling. It however takes effect only in the intermediate regime, not along the entire PES.[6]

The known correspondence between HF based RPA and the ring approximation of the coupled-cluster doubles methodology (rCCD) initiated another line of reasearch. We worked out an extension of rCCD at the multireference (MR) level. Our formulation was tailored in a way that the companion MR RPA equations coincide with Pernal's ERPA amplitude equations. Our investigation at the formal level reveals that the equivalence between RPA and rCCD is not preserved in the MR situation. We have also shown that the alternatives for the ground state energy correction formulated based on CC theory, the so-called plasmon formula of RPA or the idea of the consistent ground state are all different and neither match Pernal's ERPA-APSG expression.[7]

With the aim of numerical comparison we implemented the proposed MR rCCD theory and applied it to the APSG reference function. Numerical tests of the originally published approach revealed that the approximations introduced can not always be justified. We rectified the problems by incorporating cumulants up to rank four, introducing a novel way of redundancy handling and applying a pruning in the excitation manifold. First results of the revised approach are promising. We obtained accuracy comparable to (occasionally better than) the related ERPA based correction with $\mathcal{O}(N^6)$ scaling applying to both. The material is now being prepared for publication. Two manuscipts are being prepared in parallel. One reports on the performance of rCCD, the other presents a theoretical framework for the extension towards ladder diagrams and some further terms of the amplitude equations, remaining in the $\mathcal{O}(N^6)$ regime of computational cost.

We have attempted to devise corrections to a triplet including geminal reference as well. Our study in this line produced a crossbreed between UHF as reference and a geminal based PT suggested by our laboratory in 2002. In the resulting approach natural orbitals (NO) of the UHF wavefunction and its geminal structure (i.e. NOs with occupation numbers adding up to two) are the ingredients defining a zero order approximation for the subsequent PT treatment. The previously formulated PT had to be extended to allow for singlet-triplet mixed geminals instead of pure singlets. Results on singlet triplet gaps of biradical systems show good performance in cases characteristically ill-behaving with UHF based MP, due to serious spin contamination. Good results are typically obtained when geminal coefficient relaxation, performed at order zero of the approximation eradicates spin contamination. Incorrect spin of the geminal coefficient relaxed zero order has harmful influence, undermining accuracy or even yielding essentially wrong results in severely contaminated cases.[8]

2.3 The nature of static correlation

Failure of APSG based PT in case of multiple bond breaking inspired the study where we examined whether the divergence of a PT series can be used to identify the onset of static correlation in a bond dissociation process. Our first, analytical studies were performed on the H_2 molecule, as the simplest test case of single bond rupture. We examined three criteria for identifying the presence of static correlation: (i) appearance of the UHF solution (Coulson-Fischer point); (ii) crossing of the RHF and a minimal open-shell Ansatz (c.f. two-determinantal singlet) energy curves; (iii) divergence of Eptein-Nesbet PT built upon RHF. Interestingly, criteria (ii) and (iii) match exactly both at the CNDO and minimal basis set ab initio level, setting the onset of static correlation at cca.

1.6 Angstrøm in the latter case. This value is larger than 1.2 Angstrøm, where the Coulson-Fischer point appears. Our results on this topic remain unpublished for lack of numerical calculations in non minimal basis and consideration of multiple bond dissociation.

2.4 Lower bounds to energy eigenvalues

We worked out a computational scheme for the approximate evaluation of Löwdin's bracketing function for providing approximate lower bounds to matrix eigenvalues. The approach is based on the Taylor-series expansion of the inverse matrix, partitioned for a zero order and a remainder. Implementation of the method in a direct full CI code is relatively stratightforward, generating essentially just I/O increase. Due to the series approximation, the strict lower bound property is lost. Numerical tests are however encouraging, as the error of the approximate lower bound is often of same order of magnitude, as the usual upper bound, provided by the expectation value.[9]

Our initial approach was based on a partitioning of the Hamiltonian matrix where the first row, first column and the diagonal provides the zero order matrix. We considered an improvement where several rows and the same number of columns are included at order zero apart from the diagonal elements. Initial numerical tests were unsatisfactory, producing larger errors than the simpler zero order. At the same time, lower bound behaviour is exhibited more often by the more involved approach. Publication has to be preceded by a cross check of the implementation and detailed analysis of the results.

2.5 Studies related to PT

A former suggestion of our laboratory on maximizing convergence radius by appropriate choice of the partitioning was tested on molecular systems in 2017. The key element is the observation that increasing orders of PT involve the reduced resolvent (Q) multiplied by the perturbation (W) on increasing powers. Minimizing the matrix norm of QW by the choice of appropriate level shift parameters is therefore expected to increase the convergence radius of the PT series.[10]

Novel ideas on PT evolved around the possibility of applying the technique of analytic continuation for divergent series. It is well known, that divergence of the series at value 1 of the coupling strength parameter implies the presence of a square root branch point at some point on the unit circle of the complex plane. The distance of the branch point from the origin in fact defines the radius of convergence. Our investigations aimed at utilizing the terms of the series in the convergent domain for extrapolating in the divergent interval on the real axis till value 1.[11] We have also considered utilizing a convex domain of the complex plane for extrapolation.[12] Extrapolation on the complex plane was combined with the Cauchy integral formula and error minimization in a study that appeared in the Communications section of the Journal of Chemical Physics in 2019.[13]

Our experience in multireference PT was the basis of a collaborative project aiming at the PT analysis of cumulant approximations implied in density matrix functional theory (DMFT) based approaches. The main idea is that cumulant expressions involving the 1-particle reduced density matrix (1-RDM) and one- and two-electron integrals should be correct up to the leading order of

PT in the limit of vanishing inter electron interaction. We explored the weak correlation regime in the situation where the reference function is a combination of several determinants, deriving an analogue of the exact constraint for the 2-cumulant formulated by Kutzelnigg and Mukherjee for the single reference case. Examination of the existing 1-matrix functional reveals that none of them complies with the newly obtained asymptotic formula. This points to the necessity of novel constructions satisfying the constraint bilinear in the 2-cumulant.[14]

In response to an invitation of the Elsevier publishing house, a review article was written on the topic of perturbation theory.[15]

2.6 Further studies beyond the plan

Generating a set of $N-1$ orthonormal (ON) vectors, treated on an equal footing and orthogonal to a specific vector of an N-dimensional space is a problem that has been solved previously. An interesting aspect is that the solution can be given in closed form, without any need of relying on numerical algorithms. During the project we investigated the extended problem where M ON vectors ($M < N$) are prescribed as elements of the final N-dimensional ON set.[16] The obtained formulae has been put to use in formulating a multi-state version of a multiconfigurational PT framework, previously proposed by our laboratory.

In a continuation of a project aiming at the semi-empirical level modelling of Raman Optical Activity (ROA) spectra we incorporated line group symmetry in our approach. This allowed us to compute the Raman and ROA spectra of chiral carbon nanotubes.[17]

Variation of the energy functional in the case where the Hamiltonian depends on its solution, i.e. the eigenvector has been examined. We found that the energy is not stationary at the eigensolution but its variation is determined by a generalized Hellmann-Feynman theorem. A functional different from the energy can be constructed which is stationary at the eigenvector.[18]

References

[1] P. R. Surján, P. Jeszenszki, and Á. Szabados. Role of triplet states in geminal-based perturbation theory. *Molecular Physics*, 113:2960–2963, 2015.

[2] P. Jeszenszki, V. Rassolov, P. R. Surján, and Á. Szabados. Local spin from strongly orthogonal geminal wavefunctions. *Molecular Physics*, 113:249–251, 2015.

[3] P. Jeszenszki, P. R. Surján, and Á. Szabados. Spin symmetry and size consistency of strongly orthogonal geminals. *Journal of Chemical Theory and Computation*, 11:3096–3103, 2015.

[4] Zs. É. Mihálka, P. R. Surján, and Á. Szabados. Half-projection of the strongly orthogonal unrestricted geminals' product wavefunction. *Journal of Chemical Theory and Computation*, DOI 10.1021/acs.jctc.9b00858, 2019.

[5] P. R. Surján. Geminal approach. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier, 2016. ISBN 978-0-12-409547-2. doi: 10.1016/B978-0-12-409547-2.11468-4.

[6] Á. Margócsy, P. Kowalski, K. Pernal, and Á. Szabados. Multiple bond breaking with APSG-based correlation methods: comparison of two approaches. *Theoretical Chemistry Accounts*, 137:159, 2018.

[7] Á. Szabados and Á. Margócsy. Ring coupled-cluster doubles correction to geminal wavefunctions. *Molecular Physics*, 115:2731–2742, 2017.

[8] D. Földvári, Zs. Tóth, P. R. Surján, and Á. Szabados. Geminal perturbation theory based on the unrestricted Hartree–Fock wavefunction. *The Journal of Chemical Physics*, 150:034103, 2019.

[9] Zs. Tóth and Á. Szabados. Energy error bars in direct configuration interaction iteration sequence. *The Journal of Chemical Physics*, 143:084112, 2015.

[10] Zs. É. Mihálka, Á. Szabados, and P. R. Surján. Effect of partitioning on the convergence properties of the Rayleigh–Schrödinger perturbation series. *The Journal of Chemical Physics*, 146:124121, 2017.

[11] Zs. É. Mihálka and P. R. Surján. Analytic-continuation approach to the resummation of divergent series in Rayleigh–Schrödinger perturbation theory. *Phys. Rev. A*, 96:062106, 2017.

[12] P. R. Surján, Zs. É. Mihálka, and Á. Szabados. The inverse boundary value problem: application in many-body perturbation theory. *Theoretical Chemistry Accounts*, 137:149, 2018.

[13] Zs. É. Mihálka, Á. Szabados, and P. R. Surján. Application of the Cauchy integral formula as a tool of analytic continuation for the resummation of divergent perturbation series. *The Journal of Chemical Physics*, 150:031101, 2019.

[14] J. Cioslowski, Zs. É. Mihálka, and Á. Szabados. Bilinear constraints upon the correlation contribution to the electron–electron repulsion energy as a functional of the one-electron reduced density matrix. *Journal of Chemical Theory and Computation*, 15:4862–4872, 2019.

[15] Á. Szabados. Perturbation theory: Time-independent aspects of the theory applied in molecular electronic structure description. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier, 2017. ISBN 978-0-12-409547-2. doi: 10.1016/B978-0-12-409547-2.11467-2.

[16] Zs. Tóth, P. R. Nagy, P. Jeszenszki, and Á. Szabados. Novel orthogonalization and biorthogonalization algorithms. *Theoretical Chemistry Accounts*, 134:1–6, 2015.

[17] P. R. Nagy, J. Koltai, P. R. Surján, J. Kürti, and Á. Szabados. Resonance Raman optical activity of single walled chiral carbon nanotubes. *The Journal of Physical Chemistry A*, 120: 5527–5538, 2016.

[18] Zs. É. Mihálka, Á. Margócsy, Á. Szabados, and P. R. Surján. On the variational principle for the non-linear Schrödinger equation. *Journal of Mathematical Chemistry*, 2019. doi: 10.1007/s10910-019-01082-5.