



---

Budapest University of Technology and Economics  
Faculty of Chemical Technology and Biotechnology  
George Olah Doctoral School

# Efficient Excited-State Quantum Chemical Approaches for Large Molecules

Thesis Statements

**Dávid Mester**

Advisor: Prof. Mihály Kállay

Department of Physical Chemistry and Materials Science  
MTA-BME Lendület Quantum Chemistry Research Group

**2019**

# 1 Introduction

Our efforts on the development of quantum chemical methods focus on the efficient and accurate solution of the Schrödinger equation. Over the past decades, one significant avenue of the developments has been the elaboration and implementation of more accurate methods. Many high-accuracy quantum chemical methods were successfully developed, which allow the theoretical prediction of experimentally measurable physical and chemical properties. The accuracy of the obtained results is comparable to or more and more often even better than that of the experiments.

However, the solution of the corresponding systems of equations containing billions of unknown variables is extremely demanding. Therefore, the upper limit of the applicability of the most accurate methods is only a few atoms. Accordingly, the substantial aim of the method development in recent years has been to reduce the calculation times of our existing, highly reliable methods, or to elaborate new schemes by modifying or combining the theories.

External perturbations, such as electromagnetic radiation of certain strength can result in to electronic excitation. There are several actively researched phenomena related to the excited electronic states of molecular systems. For instance, excited states play an important role for photochromic materials, photo-initialized chemical processes, and energy transfer and storage. Quantum-chemical methods have now become routine tools in the investigation of excited-state properties and processes in the fields of spectroscopy, analytical- and biochemistry, where often extensive molecular systems are studied. Fortunately, the error of the calculated ground- and excited-state energy cancel each other in general, thus it is permissible to use lower scaling methods for fairly accurate excited-state properties. However, the upper limit of the applicability of such methods is still easily accessible, despite their relatively low scaling.

Our goal is to expand the current limit of accurate excited-state methods and to be able to model chemically relevant systems that can not be treated currently due to the extensive computational requirements.

## 2 Overview

Nowadays, time-dependent density functional theory (TDDFT), which is derived from density functional theory (DFT) through the linear-response formalism, is the most common choice to investigate time-dependent properties of molecular systems, such as excitation spectra, polarizabilities, and chiroptical properties. The time-dependent Hartree–Fock (TDHF) method is an analogue of TDDFT where one chooses the Hartree–Fock (HF) solution as the ground-state reference. Another related HF-based excited-state method is the configuration interaction singles (CIS) approach. It is well known that TDDFT is generally more accurate despite the similar computational requirements, at least for valence excitations, due to the consideration of dynamic electron correlation.

The computational expenses of the CIS, TDHF, and TDDFT methods, as long as the DFT functional contains HF exchange contribution, scale with the fourth power of the system size. Accordingly, the upper limit for the applicability of these theories, if no further approximations are introduced, is roughly 150 atoms. Since the size of the chemically relevant systems in many cases exceeds this region, a number of attempts have been made over the last two decades to reduce the computation time and the resources required for the calculations. These approximations have many advantages, but some deficiencies are also present. For instance, some of them cannot be applied in a black box manner, which means that the user has to define the subsystems of the molecule. Particular methods cannot be used with hybrid functionals, whereas other schemes can only be applied to periodic systems. Furthermore, quantitative results cannot be obtained by most of the approaches due to the large error of the approximations, or, compared to genuine TDDFT calculations, the size of the investigated systems cannot be significantly increased. To accelerate TDDFT calculations utilizing hybrid functionals with a tolerable error, the chain of spheres exchange approach

and the auxiliary-density matrix method were developed by Neese<sup>1</sup> and Kjærgaard<sup>2</sup> and co-workers. Both schemes reduce the computation time required to calculate the rate-determining exact exchange contributions. For the calculation of the exchange contribution, an asymptotic linear scaling can be achieved with both methods.

While hybrid TDDFT excitation energies and spectral intensities are quite good for valence excitations, those for Rydberg and charge transfer states, or excitations of extended  $\pi$ -electron systems can still be qualitatively incorrect. Besides the aforementioned problems, in the well-established adiabatic approximation of TDDFT only states dominated by one-electron excitations can be modeled.

The performance of density functional approximations for ground-state calculations can be improved by combining them with wave function methods. Nowadays, the most popular mixed approaches are the double hybrid (DH) functionals. The application of DHs to excited states is a relatively unexplored field. The first step in this direction was made by Grimme and Neese<sup>3</sup>. In their approach, which has been followed in all the excited-state DH schemes, a TDDFT calculation is performed omitting the second-order perturbation theory correction of the functional, and subsequently, the effect of double excitations for the excitation energy is added by calculating the second-order (D) correction of the CIS(D) method using the TDDFT excitation amplitudes. The transition moments are evaluated using the conventional TDDFT expressions, and no second-order correction is added. Recently Schwabe and Goerigk<sup>4</sup> successfully combined time-dependent DH DFT theory with spin scaling.

One of the most accurate electronic structure theories is the coupled-cluster (CC) approach. The hierarchical CC methods based on the exponential parametrization of the wave function enable the consideration of electron correlation with arbitrary accuracy. The CC methods can also be generalized to excited states using the linear-response theory. The simplest CC model, the approximate second-order CC (CC2) method supplies

---

<sup>1</sup>T. Petrenko, S. Kossmann, and F. Neese, *J. Chem. Phys.*, **2011**, 134:054116

<sup>2</sup>C. Kumar *et al.*, *Int. J. Quantum Chem.*, **2018**, 118:e25639

<sup>3</sup>S. Grimme and F. Neese, *J. Chem. Phys.*, **2007**, 127:154116

<sup>4</sup>T. Schwabe and L. Goerigk, *J. Chem. Theory Comput.*, **2017**, 13:4307

excitation energies and transition moments with a moderate error, at least for valence states, with respect to higher-order CC methods. Besides that, among the theories suitable for excited-state property calculations algebraic-diagrammatic construction (ADC) is one of the most promising approaches. It is a Hermitian and size-consistent method, and it is relatively easy to implement. It has been proven that the second-order ADC [ADC(2)] method is practically as accurate as CC2. Both methods scale as the fifth power of the systems size.

Despite the excellent numerical results the applicability of CC and ADC methods is limited due to the steep scaling of their costs with the system size. An alternative solution instead of the TDDFT methods could be the reduction of computational expenses of CC2 and ADC(2) approaches, which provide results consistently better than common TDDFT methods. For this purpose, a widely used technique is the density fitting (DF) approach, where the four-center electron repulsion integrals are approximated. Very efficient methods can be developed by determining the molecular orbitals (MOs) that play an important role in the excitation. One of the most popular schemes is the natural orbital (NO) approximation, with which the MO space where the equations are solved can be effectively reduced. In the NO approach, a one-particle density matrix, which is formed using a lower-level wave function, is diagonalized, and the orbitals with significant importance are selected from the resulting NOs. Further computational savings can be achieved if one takes advantage of the locality of the MOs. In this case, not only the time required for the calculations is decreased, but at the same time the scaling of the methods is also reduced. Among the many attempts worth mentioning is the recent local correlated natural transition orbital framework approach of Baudin and Kristensen<sup>5</sup>, which is an encouraging combination of the NO and the local approaches. A comprehensive study was recently published on the topic of reduced-cost and reduced-scaling approximations by Crawford, Kumar, and co-workers<sup>6</sup>.

---

<sup>5</sup>P. Baudin and K. Kristensen, *J. Chem. Phys.*, **2017**, 146:214114

<sup>6</sup>T. D. Crawford *et al.*, *WIREs Comput. Mol. Sci.*, **2019**, e1406

### 3 Results

In this dissertation first, a reduced-cost implementation of the CC2 and ADC(2) methods is presented using the virtual NO and natural auxiliary function (NAF) approximations. For the construction of the NOs for the ground-state calculations the second-order Møller–Plesset (MP2) density matrix is a plausible choice. For excited states the selection of an appropriate wave function is less trivial. The CIS(D) approach seems to be a good choice since it is an excited-state generalization of the MP2 method and accounts for about the same amount of correlation for the excited state as the MP2 method does for the ground state. In order to solve the ground- and excited-state equations in the same basis, a single, state-specific, “state-averaged” density matrix is assembled for each excited state for the construction of the NOs. For this purpose we simply take the average of the MP2 and CIS(D) densities. The number of the variables could be further reduced utilizing the NAF approach, which is very similar to the NO approximation, however, it does not contain any physical or chemical considerations. We assess several possibilities concerning at what stage the NAF basis should be constructed and what block of the integral list is to be used for the determination of the NAFs. On the basis of our numerical results, the use of the entire integral list turns out to be the most advantageous choice. The various approximations are carefully benchmarked for a representative test set including all important types of excitations, and conservative truncation thresholds are selected which guarantee errors much smaller than the intrinsic error of the methods. All in all, an average speedup of more than an order of magnitude can be achieved at the expense of a mean absolute error of 0.02 eV in the calculated excitation energies for the CC2 and ADC(2) methods compared to the canonical results, while that for oscillator strengths is 0.001 in the case of ADC(2). The rigorous cutoff parameters together with the significantly reduced operation count and storage requirements allow us to obtain accurate excitation energies and transition properties using triple- $\zeta$  basis sets for systems of up to one hundred atoms.

Thereafter a reduced-scaling algorithms for further computational savings are pre-

sented utilizing the locality of the molecular orbitals. First, an approximation is presented which can efficiently decrease the computational expenses of the CIS and TDDFT methods, as well as of the related TDHF and TDA-TDDFT approaches. The approximation is the adaptation of the local density fitting scheme developed for Hartree–Fock calculations for excited states and reduces the quartic scaling of the methods to cubic. A robust algorithm to construct a local fitting domain for each occupied orbitals is discussed in detail. Obviously these sets should contain the auxiliary functions which are located on the atoms on which the corresponding localized MOs are localized. To determine these atoms Löwdin atomic charges are calculated for each MO, and the atoms whose charge is greater than a predefined threshold are selected. Subsequently, we supplement the domain taking into account the information carried by the excitation. For this purpose, we analyze the solution vector transformed to the atomic orbital and localized MO bases. The elements of the normalized vector are sorted into ascending order, and starting from the largest one the squares of the elements are summed up until the sum becomes larger than a predefined threshold. If the element of the vector is included in the sum, we add those auxiliary functions to the domain which reside on the atom of the corresponding atomic orbital. Our benchmark calculations show that, for molecules of 50 to 100 atoms, average speedups of 2 to 4 can be achieved for CIS and TDDFT calculations at the expense of negligible errors in the calculated excitation energies and oscillator strengths. The mean absolute error of the excitation energies and oscillator strengths does not exceed 1 meV and 0.001, respectively, while the maximum error is less than 5 meV and 0.01 in every case. For bigger systems or molecules of localized electronic structure significantly larger speedups can be gained.

Then a framework for the reduced-scaling implementation of excited-state correlation methods is presented. An algorithm is introduced to construct excitation-specific local domains, which include all the important molecular orbitals for the excitation as well as for the electron correlation. The basic assumption of the theory is that only a subset of molecular orbitals contribute dominantly to an excited-state wave func-

tion, and the number of these orbitals does not increase with the size of the system. To determine the important MOs involved in the excitation, we first solve the CIS eigenvalue problem and the solution is transformed to the localized MO and projected atomic orbital bases. The selected most important orbitals are collected to the domain. Subsequently, the domains are supplemented with the environment of the selected orbitals since the CIS wave function could be a poor approximation for the corresponding state. The final domain presumably contains all the important orbitals to describe the excitation. The sizes of the resulting compact domains are further decreased utilizing our reduced-cost techniques based on the NAF and local NO approaches. Additional methodological improvements for the evaluation of density matrices are also discussed. The results of benchmark calculations performed at the ADC(2) level are presented, and it is demonstrated that the speedups achieved are significant even for systems of fewer than 100 atoms, while the errors introduced by our approximation are highly acceptable. Our results show that the new reduced-scaling algorithm allows us to carry out correlated excited-state calculations using triple- $\zeta$  basis sets with diffuse functions for systems of up to 400 atoms or 13000 atomic orbitals.

Finally, a new type of combined double hybrid TDDFT-ADC(2) method is proposed for the calculation of spectral properties of molecular systems. The new ansatz is an extension of the previous double hybrid methods replacing the non-iterative doubles correction of CIS(D) by iteratively evaluated ADC(2)-like second-order terms. Similar to the CIS(D) approach, ADC(2) can also be regarded as a natural excited-state extension of the MP2 method. Thus, ADC(2) is also a plausible candidate for the development of combined DFT-wave function methods for excited states. The major difference with respect to the existing DH methods is that, in the new schemes, the double excitations are treated iteratively. The anticipated advantage of the new models over the previous CIS(D)-based DHs is twofold. First, concerning excitation energies, ADC(2) moderately but consistently outperforms CIS(D), thus, an improvement in the calculated transition energies is expected, especially when the weight of double excitations is relatively large in the excited-state wave function. Second, in contrast to

the CIS(D)-based DHs, the present methods also allow us to evaluate the transition moments at a higher level taking into account the effect of double excitations, which should considerably raise the quality of the computed spectral intensities. Our results show that, compared to the existing double hybrids, the ADC(2)-based methods do not increase the average accuracy of the computed excitation energies for Rydberg states but consistently improve the quality of the transition energies for valence states, especially for those of sizable double excitation character. For oscillator strengths, the performance of the new methods is significantly better. The computational requirement is compared to the existing perturbative methods and it is demonstrated that our ansatz is highly competitive, especially when the NAF approach is used. The proposed fourth-power-scaling method with a mean absolute error of about 0.15 eV for excitation energies, outperforms other double hybrid approaches, ADC(2), its spin-scaled variants, and also the computationally more demanding CC2 model.

## 4 Applications

The developed approximations and implemented methods are available for the scientific community as part of the MRCC program suite. Free, open-source license is available for academic use. The algorithms can be employed for accurate calculation of excitation energies and absorption- or circular dichroism spectra. In addition, it can be used to investigate the rearrangement in the charge distribution upon excitation, or to identify the molecular orbitals involved in the excitation. These results may influence the research in the field of organic-, biomolecular-, and photochemistry. Our calculations demonstrate that the new approaches enable the efficient computation of excited-state properties of molecular systems composed of up to 400 atoms with reasonable basis sets. Accordingly, new materials can be designed using these theoretical tools that have not been possible so far due to their sizes. Finally, the developed local framework is quite general and flexible, hence it can be utilized in numerous related quantum chemical approaches with minor modifications.

## 5 New scientific results

- 1 I have demonstrated the efficiency of the NAF approximation for excited-state CC2, ADC(2), and double hybrid TDDFT calculations. I have proved that the size of the auxiliary basis can be halved regardless of the system without any significant error [1,2,5].
- 2 I have proposed a state-averaged density matrix to obtain suitable virtual NOs for excited-state CC2 and ADC(2) calculations. I have proved that the size of the virtual space can be halved and the approach can be combined with the NAF approximation. With the final algorithm, an average speedup of more than an order of magnitude can be achieved at the expense of moderate error in the calculated excitation energies and oscillator strengths [1,2].
- 3 I have adapted the local density fitting approximation for the CIS, TDHF, and (TDA-)TDDFT methods. I have introduced an algorithm to construct the local fitting domains for excited-state calculations. The developed approximation is practically error-free and scales as the third power of the system size [3].
- 4 I have developed an algorithm to construct excitation-specific local domains for excited-state correlation methods. The scheme is efficiently combined with the NO and NAF approximations and enables us to carry out calculations for systems of up to 400 atoms using triple-zeta basis sets with diffuse functions [4].
- 5 I have elaborated an ADC(2)-based double hybrid TDDFT ansatz, with which the effect of the double excitations is treated iteratively. I have proved that the proposed fourth-power-scaling model outperforms not only the existing CIS(D)-based double hybrid approaches and ADC(2) variants but also the considerably more expensive coupled-cluster methods [5].

## 6 Publications and presentations

### 6.1 Publications related to the thesis

- 1 Reduced-cost linear-response CC2 method based on natural orbitals and natural auxiliary functions, D. Mester, P. R. Nagy, and M. Kállay,  
*Journal of Chemical Physics*, **2017**, 146:194102  
IF: 2.843, non-self citations: 8, contribution: 100%
- 2 Reduced-cost second-order algebraic-diagrammatic construction method for excitation energies and transition moments, D. Mester, P. R. Nagy, and M. Kállay,  
*Journal of Chemical Physics*, **2018**, 148:094111  
IF: 2.997, non-self citations: 4, contribution: 100%
- 3 Reduced-scaling approach for configuration interaction singles and time-dependent density functional theory calculations using hybrid functionals, D. Mester and M. Kállay,  
*Journal Chemical Theory and Computation*, **2019**, 15:1690  
IF: 5.313, non-self citations: 0, contribution: 100%
- 4 Combined density functional and algebraic-diagrammatic construction approach for accurate excitation energies and transition moments, D. Mester and M. Kállay,  
*Journal Chemical Theory and Computation*, **2019**, 15:4440  
IF: 5.313, non-self citations: 0, contribution: 100%
- 5 Reduced-scaling correlation methods for the excited states of large molecules: implementation and benchmarks for the second-order algebraic-diagrammatic construction approach, D. Mester, P. R. Nagy, and M. Kállay,  
*Journal Chemical Theory and Computation*, **2019**, 15:6111  
IF: 5.313, non-self citations: 0, contribution: 100%

## 6.2 Other publications

- Unconventional bond functions for quantum chemical calculations, D. Mester, J. Csontos, and M. Kállay,  
*Theoretical Chemistry Accounts*, **2015**, 134:74  
IF: 1.806, non-self citations: 6
- A study on the optical resolution of 1-isopropyl-3-methyl-3-phospholene 1-oxide and its use in the synthesis of borane and platinum complexes, P. Bagi, K. Juhász, I. Tímári, K. E. Kövér, D. Mester, M. Kállay, M. Kubinyi, T. Szilvási, P. Pongrácz, L. Kollár, K. Karaghiosoff, M. Czugler, L. Drahos, E. Fogassy, and G. Keglevich,  
*Journal of Organometallic Chemistry*, **2015**, 797:140  
IF: 2.336, non-self citations: 2
- Adsorption of an active molecule on the surface of halloysite for controlled release application: interaction, orientation, consequence, J. Hári, P. Polyák, D. Mester, M. Micusik, M. Omastova, M. Kállay, and B. Pukánszky,  
*Applied Clay Science*, **2016**, 132-133:167  
IF: 3.101, non-self citations: 0
- Uracil-linked hydroxyflavone probe for the recognition of ATP, M. Bojtár, P. Z. Janzsó-Berend, D. Mester, D. Hessz, M. Kállay, M. Kubinyi, and I. Bitter,  
*Beilstein Journal of Organic Chemistry*, **2018**, 14:747  
IF: 2.595, non-self citations: 2
- Hydrogen bonding effects on the fluorescence properties of 4'-diethylamino-3-hydroxyflavone in water and water-acetone mixtures, D. Hessz, M. Bojtár, D. Mester, Z. Szakács, I. Bitter, M. Kállay, and M. Kubinyi,  
*Spectrochimica Acta Part A*, **2018**, 203:96  
IF: 2.931, non-self citations: 1

### 6.3 Presentations

- Poster presentation: **A Symposium in Honour of Jeppe Olsen on the Occasion of his 60th Birthday**, 2016, Oslo, Norway  
Reduced-cost LR-CC2 method based on natural orbitals, D. Mester, P. R. Nagy, and M. Kállay
- Oral presentation: **MTA Material and Molecular Structure Working Committee Meeting**, 2017, Mátrafüred, Hungary  
Lokális korrelációs módszerek fejlesztése gerjesztett állapotokra, D. Mester, P. R. Nagy, and M. Kállay
- Poster presentation: **16th International Congress of Quantum Chemistry**, 2018, Menton, France  
Reduced-cost ADC(2) method for excitation energies and transition moments, D. Mester, P. R. Nagy, and M. Kállay
- Poster presentation: **10th Triennial Congress of the International Society for Theoretical Chemical Physics**, 2019, Tromsø, Norway  
Reduced-scaling correlation methods for the excited states of large molecules, D. Mester, P. R. Nagy, and M. Kállay
- Oral presentation: **MTA Reaction Kinetics and Photochemistry Working Committee Meeting**, 2019, Mátrafüred, Hungary  
Hatékony kvantumkémiai módszerek nagy molekulák gerjesztett állapotainak számításához, D. Mester, P. R. Nagy, and M. Kállay