Efficient Approximations to SCF Energies and Derivatives

Hartree-Fock theory plays a central role in Quantum Chemistry. It serves on the one hand as a basis for electron correlation approaches while on the other it is also essential for density functional theory. The evaluation of the HF exchange term is a rather time consuming operation for larger molecules and as a result for a wide range of chemically or biologically interesting systems. This motivates the development of various efficient approaches to exchange evaluation. The Chain of Spheres exchange (COSX) algorithm, one of the most efficient methods available for this purpose today, has been developed in the institute since 2009. The algorithm has favorable scaling and contraction properties which makes it especially suitable for larger systems. In combination with the resolution of identity for the Coulomb term (RI-J) it enables large scale HF and DFT calculations.

The possibility of robust fitting in the COSX context is of a more theoretical interest which has also been investigated during the last few years. Several other ideas are also being considered including various projection based approaches, linear scaling RI approaches and range separation.

To realize the full potential of any approximation to the SCF energy, it is inevitable to obtain the corresponding gradients and the Hessian. These can then be used in geometry optimizations and to calculate various properties. Gradients are determined in an approximate manner for the COSX algorithm, while for Hessian calculation its main advantage is in the CP-SCF step.

Basis Set Incompleteness

Basis set incompleteness affects quantum chemical methods in various ways. Since practical methods correspond to the expansion of the wave function in a truncated basis, it becomes important to find a good way to approximate the basis set limit. Basis set extrapolation and explicitly correlated (F12) methods are two different ways to address the same question. At the MP2 level, the additional cost of F12 methods is comparatively high. Nevertheless, our investigations show that combined with various approximate SCF approaches the F12 is fair competitor to extrapolation methods with respect to efficiency as well as accuracy.

Incompleteness effects are also connected to numerical errors, in fact the incompleteness of basis sets and the incompleteness of numerical grids are closely related. There is a conceptual similarity between complementary auxiliary basis (CABS) functions of F12 approaches and the so-called dealiasing functions of seminumerical methods, which are used to cancel numerical errors. We argue that the CABS sets, or other basis sets developed based on similar principles can generalize the notion of grid specific dealiasing functions. This procedure is also a generalization of the overlap fitting method in which the dealiasing set is identical to the atomic orbital basis.

Exchange Methods for Correlation

The COSX method can also be used beneficially in correlation methods. In these cases the SCF density is replaced by an effective density in which the amplitudes of various excited determinants are contracted with molecular orbital coeffcients. Thus, the evaluation of the so called singles Fock term in the LPNO-CCSD method, and the external exchange term in the SCS-MP3 method can both be significantly accelerated without significant loss of accuracy. In the former case the contraction involves the singles, in the latter case the doubles amplitudes, but in both cases the approximated terms are form the bottleneck of the corresponding method.
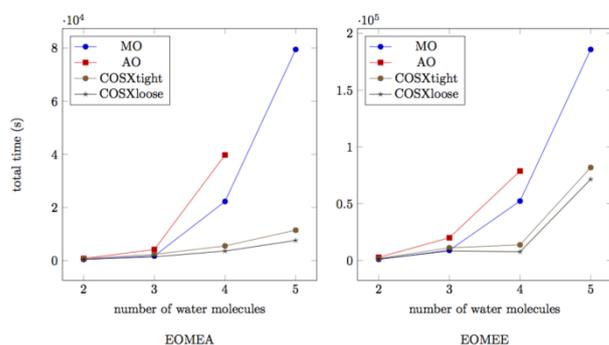
Figure 1

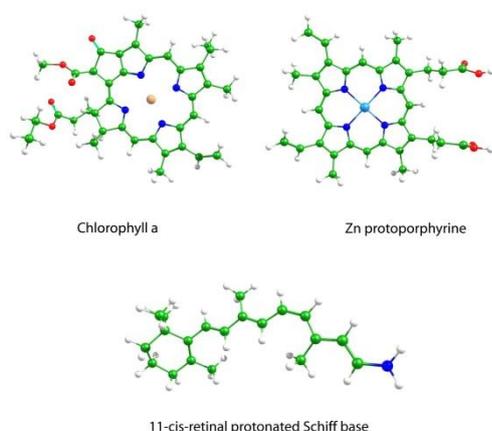Fig. 1: The Performance of COSX-EOM Methods

Chlorophyll a

Zn protoporphyrine

11-cis-retinal protonated Schiff base

Figure 2

Fig. 2: Molecules Accessible to back-transformed PNO-(ST)EOM Methods

Equation of Motion Coupled Cluster Method

Equation of Motion (EOM) Coupled Cluster theory is a standard way of obtaining reliable excitation energies and properties of excited states. The COSX approximation is particularly well suited to accelerate the most expensive term involved in the EOM equations with very good accuracy, see Figure 1. We were able to calculate the first two excited states of the trans retinal molecule (539 basis functions) in less then two days on eight cores only, which is the first such EOM calculation reported in the literature for such a large non-symmetric system.

Ionization operators can also be obtained within the EOM framework. The resulting equations are special cases of the general EOM equations for ionization potentials (IP) and electron attachments (EA). The advantages of the COSX approximation are even more significant in the EA case than for regular EOM; calculations up to a hundred atoms can now be carried out using our code. The accurate calculation of IP and EA values can be a valuable contribution to redox potential calculations, such as those involved in Photosystem II.

In order to speed up EOM calculations to an even larger extent, we currently use a strategy which employs the domain based pair natural orbital (DLPNO) framework for the ground state coupled cluster calculation in combination with the similarity transformed equation of motion (STEOM) method for the excited state part. The DLPNOs are back transformed for the purposes for the excited state calculation, while the excitation manifold is truncated to the singles space, while the effect of the doubles is still taken care of using a similarity transformation. The latter is parametrized in terms

of IP and EA eigenstates, which also benefit from the COSX approximation. This way, we have been able to calculate excitation energies for molecules of 80-90 atoms (e.g., Chlorophyll A, more than a 1000 basis functions), thus providing an alternative for time dependent density functional theory for systems which are out of reach for traditional EOM calculations, see Figure 2.

Spectroscopic Applications

In order to calculate properties like UV-Vis transition moments within the EOM framework, one needs to evaluate the left vectors in an efficient way, which is the subject of an ongoing study. The calculation of core level excitations is another field which we currently investigate especially since such calculations may contribute to the evaluation of X-Ray spectroscopy experiments. Various improvements on the STEOM scheme are also possible, among others a scheme for the automatic selection of active spaces, and a DLPNO variant of the IP/EA equations, which take a significant amount of time in STEOM calculations.

The open shell variants of the STEOM/EOM equations are currently also being implemented, using an automatic code generation scheme. Code generation in conjunction with the DLPNO framework is a promising field of research for the development of novel methods of spectroscopic and other interests. A spin adapted formulation of an open shell ionization potential scheme (ROCI-IP) is also underway, which may be applied to transition metal chemistry and XPS spectroscopy. A further research interest is the development of robust theoretical methods for the calculation of the relevant transition rates in photochemical processes. Also more recently, we also wish to explore efficient ways to calculate NMR/EPR parameters within the DLPNO-CC framework.

These methodologies will provide useful tools for the investigation of various redox processes and spectroscopic problems studied in other groups in the institute.
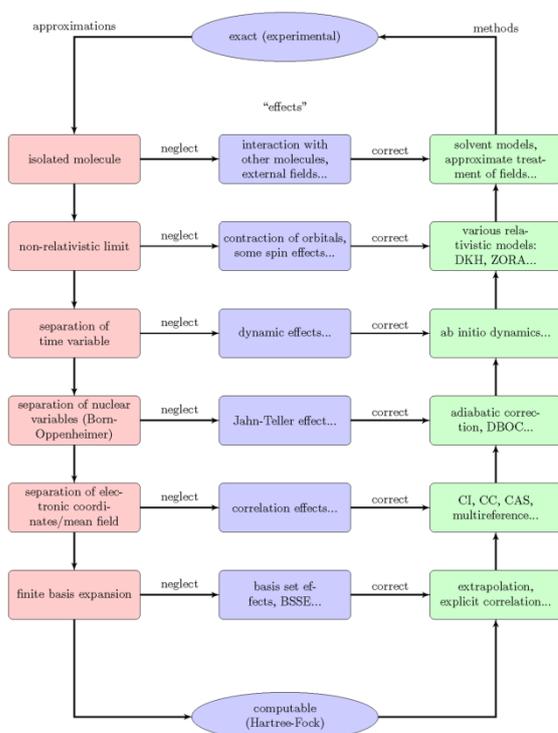

Figure 3

Fig. 3 Method Development: Approximations and Methods

Method Development More Generally

What follows is a more general view on (wave function based) method development illustrated by Figure 3. As with any attempt that tries to give a bird's eye view of a large research area, not everything can be fitted neatly into the scheme presented here. Nevertheless, the purpose of theoretical science is to build theories with quantitative predictive power from only a few principles taken for granted. These predictions can then be checked experimentally, the underlying assumption being that an ideally perfect experiment should match the predictions coming from an ideally "exact" theory. In practice there are many obstacles both on the experimental and on the theoretical side. As there are no experiments without error, there are no theories without some flaws. A realistic goal is to keep these under control.

In quantum chemistry, the "exact" side of the coin basically consist of the principal assumptions of quantum theory and (special) relativity. The resulting equations are often not amenable to practical computations, and, therefore, a system of approximations is designed to arrive at something "computable", something that can be practically evaluated. This role is taken by Hartree-Fock theory in quantum chemistry. In order to arrive at this theory, various approximations were made, but since in principle we know what we neglected, we know what effects this may cause. For example, if we ignore relativity, we cannot expect our method to account for the color of gold, which we then consider as a "relativistic effect". On the other hand in principle we are also able to correct for what we neglected, i.e., we have control over the error we allow. This may not be an easy task, however. While Hartree-Fock theory recovers 98% of the total electronic energy, the recovery of the remaining few percent can be an extremely costly computational assignment, and yet a necessary one since most chemistry is affected by this "correlation energy". For this reason, the purpose of method development is to devise efficient methods which allow us to accurately describe chemically interesting systems.