Algorithms for the Optimization of Non-orthogonal Orbitals Using Biothogonal Expansions
A Perspective Talk

Jeppe Olsen

Department of Chemistry
qLeap Centre for Theoretical Chemistry
Aarhus University

March 27, 2017

Perspective talks

The purpose is to synthetically and didactically present the state-of-the-art of a specific approach

1. pinpointing questions
2. opening current challenges to solve
3. and providing food for thought for discussion.
4. (Lecturers time: 20-25 minutes)

What is the role of Nonorthogonal methods?

Interpretation and chemical insight

- Realizes that non-orthogonal problems leads to inherent computational problems
- Instead: Gives a better way of understanding bonding
- (From my point of view: far too defensive)

An efficient computational tool

- Methods based on non-othogonal orbitals will not replace standard methods in general
- But, there is a signficant group of molecules containing atoms with partial filled shells and limited overlap, that are best described using non-orthogonal methods

The major remaining challenge in quantum chemistry: Static correlation

Status

- XASSCF, X=C,R,G leads to expansions that will not be able to treat say 30 electrons in 30 orbitals for the next few decades
- A number of newer methods are emerging: DMRG, Tensor-methods, Monte Carlo CI ... - impressive progress

But

- I would like to do the expansions directly in terms of quantities that are chemically meaningfull.
- Thus, whereas most are moving forward (with impressive speed), I will take the opposite approach: Look back (a prerogative of the older generation) and re-examine the valence bond method

When do we perform standard XASSCF(X=C,R,G) calculations?

Molecules with partly occupied shells with limited overlap

- Includes molecules of main-group elements at stretched geometries and transition, lanthanide, and actinice compounds at equilibrium.
- Small overlap $\rightarrow$ limited splitting between bonding and anti-bonding orbitals.
- A RHF calculation produces a wave function with many ionic terms not present in the full wave function.
- The major correction of the MCSCF calculation is to introduce left-right correlation to clean up the mess of RHF.
- (MCSCF is also needed when there are atomic configurations with nearly identical energies)

The non-orthogonal/valence bond approach

The standard procedure

- Use atomic rather than delocalized orbitals
- The number of electrons on each atom/ set of atomic orbitals is specified - several combinations may be needed
- When the partly filled atomic orbitals on different atoms have a small overlap, the covalent configuration is strongly dominating
- Molecules, where the standard MO-RHF-MCSCF approach is most requiring, have (often) simple non-orthogonal expansions.
- Ionic terms are introduced by orbital-optimization allowing limited delocalization of orbitals.

Second Quantization

Describes electronic wave functions as products of creation operators $a_P^\dagger$

- Example, the ground state of H_2: $a_{1\sigma\alpha}^\dagger a_{1\sigma\beta}^\dagger |\text{Vac}\rangle$

- The conjugate operators, a_P removes (if possible an electron in spin-orbital P.

- The simplicity of the standard orthogonal theories arise from the simple anticommutator $[a_P^\dagger, a_Q]_+ = \delta_{QP}$

Non-orthogonal orbitals

- For non-orthogonal orbitals, all the problems arises from anti-commutation relation $[a_P^\dagger, a_Q]_+ = S_{QP}$

- One may introduce another set of orbitals, the bi-orthonormal orbitals, $\tilde{a}_Q = \sum_{Q'} S_{Q'Q}^{-1} a_{Q'}$

- Their anticommutator with the original basis is $[a_P^\dagger, \tilde{a}_Q]_+ = \delta_{QP}$

Algorithms for orbital optimization

Two types two be discussed

- A single set of non-orthogonal orbitals
- Several sets of orbitals
  1. Each set is orthogonal and uses all SDs of the N-electron space, say FCI
  2. Each set is non-orthogonal and uses part of the N-electron space

Algorithms for optimizing a single set of orbitals

Target form of wave function

- A set of orthonormal double occupied inactive orbitals
- A set of variably occupied non-orthogonal orbitals
- A set of orthonormal unoccupied orbitals
- The various spaces are kept orthonormal to each other

Design goals

- Complete and non-redundant parameterization
- Must at all points be faster and simpler that the corresponding CASSCF expansion - (requires that we bypass the three- and four-particle density matrices of standard non-orthogonal orbital optimization).
- Should be easy to extent to response theory for calculating non-linar and excited state propeties including magnetic properties ....

Algorithms for optimizing a single set of orbitals

Parameterization of orbital rotations

Use exponential operators

- $|\tilde{0}\rangle = \exp(-\hat{\kappa}_A)\exp(-\hat{\kappa}_S)|0\rangle / \sqrt{\langle 0|\exp(-2\hat{\kappa}_S)|0\rangle}$
    1. $\kappa_A = \sum_{p>q} \kappa_{pq}^A (E_{pq} - E_{qp})$
    2. $\kappa_S = \sum_{x>y} \kappa_{xy}^S (E_{xy} + E_{yx})$ (sum only over active orbitals)
- $E_{pq} = \sum_{\sigma=\alpha,\beta} a_{p\sigma}^\dagger a_{q\sigma}$

Note

- Operators written in terms of the original operators - not bi-orthogonality here
- The operator $\exp(-\hat{\kappa}_A)$ conserves the metric. The operator $\exp(-\hat{\kappa}_S)$ changes the metric, but only between the active orbitals.
- Internal rotations in the inactive and secondary orbitals are redundant and therefore not in use.

The energy

- $E(\kappa) = \frac{\langle \tilde{0}|\hat{H}|\tilde{0}\rangle}{\langle \tilde{0}|\tilde{0}\rangle}$

- $|\tilde{0}\rangle = \exp(-\hat{\kappa}_A)\exp(-\hat{\kappa}_S)|0\rangle / \sqrt{\langle 0|\exp(-2\hat{\kappa}_S)|0\rangle}$

- $E = \frac{\langle 0|\exp(-\hat{\kappa}_S)\exp(\hat{\kappa}_A)\hat{H}\exp(-\hat{\kappa}_A)\exp(-\hat{\kappa}_S)|0\rangle}{\langle 0|\exp(-2\hat{\kappa}_S)|0\rangle}$

- The κ_A terms may be expanded as usual using the BCH expansion

- No BCH expansion for the κ_S terms

- Bioorthogonal reexpressions are used, either the right or left hand side is expressed in the bioorthogonl expansion

(The energy: $E = \frac{\langle 0| \exp(-\hat{\kappa}_S) \exp(\hat{\kappa}_A) \hat{H} \exp(-\hat{\kappa}_A) \exp(-\hat{\kappa}_S)|0\rangle}{\langle 0| \exp(-2\hat{\kappa}_S)|0\rangle}$)

The gradient (assuming $\langle 0|0\rangle = 1$

1. $E_{A,pq}^{[1]} = \frac{\partial E}{\partial \kappa_{pq}^A} = \langle 0| [E_{pq} - E_{qp}, H] |0\rangle$

2. $E_{S,xy}^{[1]} = \frac{\partial E}{\partial \kappa_{xy}^S} = -2\langle 0|(E_{xy} + E_{yx})(H - E)|0\rangle$

The gradient wrt symmetric operators

- Is the term that in standard algorithms requires the 3-electron density
- Note that the excitation is in the active space
- Proceed therefore instead as
  - Calculate sigma-vector, $|\tilde{S}\rangle = (H - E)|0\rangle$ in biorthogonal basis
  - Calculate one-electron transition density $\langle 0|(E_{xy} + E_{yx})|\tilde{S}\rangle$

# A novel approach to orbital optimization of non-orthogonal expansions

Example of convergence: Cr_2 in ano-4 4321, 3 Ångstöm

Iteration	$E - E_{conv}$	Step-size
1	0.0431467207	0.339×10^{00}
2	0.0018950153	0.162×10^{00}
3	0.0000270094	0.508×10^{-1}
4	0.0000001751	0.414×10^{-2}
5	0.0000000003	0.206×10^{-3}

Comments

- Distance is where 3d-3d bonding occurs
- The accuracy of initial guess (atomic orbitals) is noteworthy
- Optimization takes about 50 seconds on my (now retired) Macbook for one geometry

Convergence towards CASSCF of the hierarchy for the chromium trimer
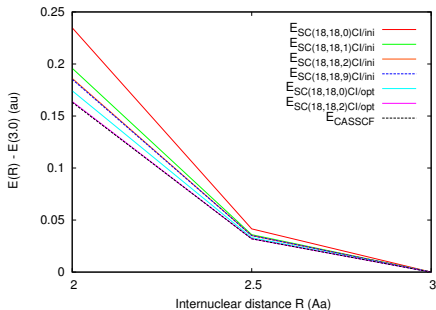
- Figure includes the results using the initial and optimised orbitals
- Energies are relative to those at 3.0 Å.
- Orbital optimisation is perhaps not essential?
- Very accurate results are obtained by including up to double excitations
- Dimensions: N(18,18,0) = 4 862, N(18,18,2) = 1 986 842, N(18,18,9) = 112 318 492.

Current status and the future plans/dreams

- up to 18-20 unpaired electrons can be used. Selected of the crucial parts of the code has been extended to treat 30 unpaired electrons.
- Standard CI is still going in several parts of the code, but these will be removed in the next year.
- Full second order (including orbital-CI coupling is also under development, as are response theory in the frequency domain.
- Non-orthogonal expansions containing hundred of millions of SDs have been realized (yes, in the Mac-book)

Algorithms for optimizing wave functions containing several sets of orbitals

Why several sets of orbitals, consider F_2

- The correlation energy increases (in absolute magnitude) when the two F-atoms approach each other
- Explanations:
  1. The standard answer: When the two F-atoms approaches each other, the electron density increases, and therefore also the correlation energy
  2. The VB person: Yes, but there is an effect more: When the two atoms appoaches each other, there is also a deviation from RHF by the tendency for an electron to jump from on F-atom to the other, forming the $F+$ F- ionic term.
- The 2p-orbitals for F^- are pretty different from the orbitals for the neutral F and therefore to describe the above effect in an efficient manner, one needs several sets of orbitals.
- The simultaneous optimization of several sets of orbitals is also needed to desribe for example mixed valence states.

Algorithms for optimizing wave functions containing several sets of orbitals

Two sets of codes

- One where, for example numerous HF or CAS states, each with different orthonormal orbitals are added and optimized
- One where, say a CAspace is divided into several parts, and each part is given a set of non-orthonormal orbitals
- Not simple coding

Central aspects

- The gradient can again be described only in terms of one- and two-electron densities using a number of tricks and rewrites
- The explicit Hessian requires now a three-body term
- A direct approach, where the Hessian times a vector is calculated, may again be formulated in terms of two-electron terms.