Correlated calculations in local basis sets formalism and iterative techniques

Ville Weijo, Pekka Manninen, Poul Jørgensen, Ove Christiansen, Jeppe Olsen

The Lundbeck Foundation Center for Theoretical Chemistry
Department of Chemistry
University of Aarhus

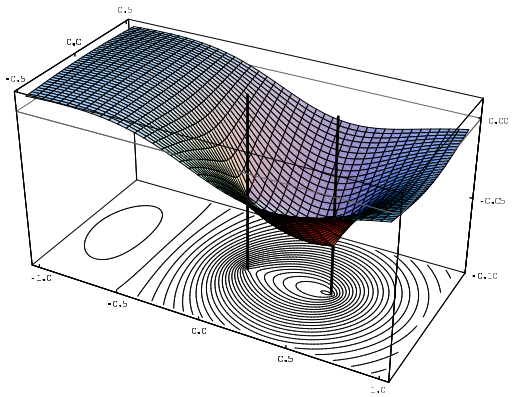
LCC2007

Contents

- Localized basis functions, advantages and disadvantages
- Two groups of methods using local basis functions
- General projected basis sets
- Second quantization in singular nonorthogonal basis
- Solving The MP1 equations in the local basis,
- Some results
- Conclusion

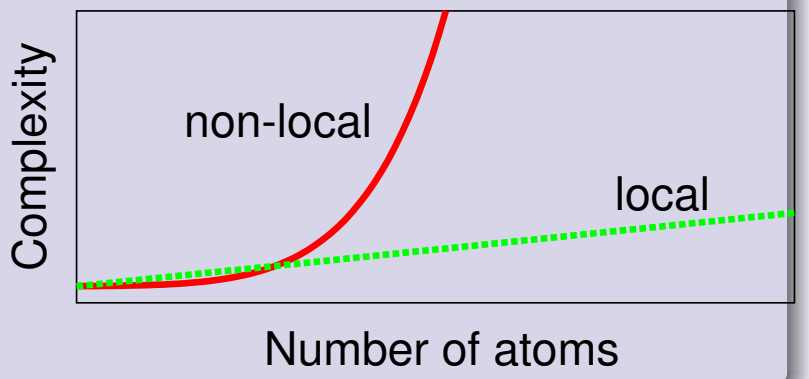
Why use local orbitals in correlated calculations

The Coulomb hole



- Caused by the repulsion of electrons
- Local: only non vanishing when two electrons are close to each other

The Coulomb hole may be described in different ways



- **Non-local** functions $\longrightarrow$ steep rise in complexity
- **Local functions** $\longrightarrow$ linear rise in complexity

Pros and cons of local orbitals

Advantages

- Local functions are expanded in a local basis
- Allows the development of linear scaling or low-scaling methods
- Avoids intermediate constructs like molecular orbitals $\longrightarrow$ Formulation directly in computable terms

Disadvantages

- Our normal mental picture of molecular orbitals is lost
- No simple division into occupied and virtual orbitals
- Loss of diagonal dominance of matrices $\longrightarrow$ iterative methods may be very slowly convergent

Localized orbitals are usefull when it is more important to distinguish between near and far than between occupied and unoccupied

Two main groups of local correlation methods

The domain approach

- Divide molecule into domains
- Localize occupied orbitals, redundant basis for virtual orbitals
- Use domains to discriminate between strong and weak interactions
- Computationally very efficient, does not go to standard MP2 limit
- Some arbitrariness, assumes that occupied orbitals may be localized
- Pioneered by P. Pulay, important work by H.J. Werner, M. Schütz

The domain-free approach

- Develop local redundant basis for occupied and virtual orbitals
- Use sparse arithmetic and storing to exploit locality
- More expensive, dimension of basis for occupied orbitals = dimension of basis set, does go towards MP2 limit
- Pioneered by M. Head-Gordon and G. Scuseria

General local basis expansions of occupied and virtual orbitals

Goals: Basis functions that are both local and provide diagonal dominant matrices

Procedure

1. Standard atomic orbital basis $|\chi\rangle$, dimension N, metric $S^{\chi}_{ij} = \langle \chi_i | \chi_j \rangle$
2. Transform to another local basis $|\phi\rangle = |\chi\rangle \mathbf{Y}$
3. Obtain redundant basis sets of dim N for occupied and for virtual orbitals by $|\psi\rangle = (|\psi_{\mathbf{k}}\rangle, |\psi_{\mathbf{a}}\rangle) = |\chi\rangle \mathbf{Z}$
4. Obtain pseudo-biorthonormal basis to simplify equations

General local basis functions

cont'd

Step 2: $|\phi\rangle = |\chi\rangle \mathbf{Y}$: local basis of dimension $2N$ allowing convergence of equations

- Orthonormal local basis sets provide diagonal dominant matrices!
  1. Cholesky factorization: $\mathbf{S}^{\chi} = \mathbf{U}^T\mathbf{U}$, $\mathbf{Y} = \mathbf{U}^{-1}$ (Used here !)
  2. Symmetric orthogonalization: $\mathbf{Y} = (\mathbf{S}^{\chi})^{-\frac{1}{2}}$
- Both may be realized in linear scaling fashion
- $\mathbf{Y} = \mathbf{1}$ leads to standard biorthonormal approach

Step 3: $|\psi\rangle = |\chi\rangle \mathbf{Z}$: separate basis sets for occ. and virt.

- $\mathbf{Z} = (\mathbf{P}^{\chi}\mathbf{S}^{\chi}\mathbf{Y}, \mathbf{Q}^{\chi}\mathbf{S}^{\chi}\mathbf{Y})$
- $\mathbf{P}^{\chi}$ is density matrix in atomic orbital basis, $\mathbf{Q}^{\chi} = (\mathbf{S}^{\chi})^{-1} - \mathbf{P}^{\chi}$
- Leads to a nonorthogonal basis of $2N$ functions- singular

Less local than projected AO's $(|\chi\rangle \mathbf{P}^{\chi}\mathbf{S}^{\chi}, |\chi\rangle \mathbf{Q}^{\chi}\mathbf{S}^{\chi})$

The Pseudo-biorthonormal basis

Standard biorthonormal approach: example, biorthonormal atomic orbital basis

- Introduce a basis $\langle \bar{\chi}|$ so $\langle \bar{\chi}_i | \chi_j \rangle = \delta_{ij}$
- $\langle \bar{\chi}| = (\mathbf{S}^{\chi})^{-1} \langle \chi|$
- Assumes a nonsingular metric

The Pseudo-biorthonormal basis

Extension to singular basis

- $\mathbf{S}^{\psi} = \langle \boldsymbol{\psi} | \boldsymbol{\psi} \rangle$ is singular $2N \times 2N$ singular metric
- Introduce $\mathbf{S}^{\psi\ominus} = \begin{pmatrix} \mathbf{P}^{\phi} & 0 \\ 0 & \mathbf{Q}^{\phi} \end{pmatrix}$
- Has some of the properties of standard inverses, for example $\mathbf{S}^{\psi}\mathbf{S}^{\psi\ominus}\mathbf{S}^{\psi} = \mathbf{S}^{\psi}$
- Is the standard general (Penrose) inverse if $|\boldsymbol{\phi}\rangle$ is an orthonormal basis, else $(\mathbf{S}^{\psi}\mathbf{S}^{\psi\ominus})^{\dagger} \neq \mathbf{S}^{\psi}\mathbf{S}^{\psi\ominus}$
- Introduce $\langle \bar{\boldsymbol{\psi}}| = \mathbf{S}^{\psi\ominus}\langle \boldsymbol{\psi}|$
- Overlap $\mathcal{P} = \langle \bar{\boldsymbol{\psi}} | \boldsymbol{\psi} \rangle = \begin{pmatrix} \mathbf{P}^{\phi}\mathbf{S}^{\phi} & 0 \\ 0 & \mathbf{Q}^{\phi}\mathbf{S}^{\phi} \end{pmatrix}$
- $\mathcal{P}$ is a projection matrix, for example $\mathcal{P}^2 = \mathcal{P}$
- As close as possible to standard biorthonormal basis

Second quantization in a singular nonorthogonal basis

ψ, dim $2N$

Purpose

Allow the development of methods using SQ directly in the ψ basis without backtransforming from MO basis

	Standard nonorth.	singular nonorth.
Biorthonormal operator	$\bar{a}_P = \sum_R (\mathbf{S})^{-1}_{PR} a_R$	$\bar{a}_P = \sum_R (\mathbf{S})^{\psi\ominus}_{PR} a_R$
Anticommutation	$[a^\dagger_P, \bar{a}_Q]_+ = \delta_{PQ}$	$[a^\dagger_P, \bar{a}_Q]_+ = \mathcal{P}_{QP}$
1-elec operator	$\hat{h} = (\mathbf{S}^{-1}\mathbf{h})_{PQ} a^\dagger_P \bar{a}_Q$	$\hat{h} = (\mathbf{S}^{\psi\ominus}\mathbf{h}^{\psi})_{PQ} a^\dagger_P \bar{a}_Q$

The MP1 equations in the ψ basis

Inserting the expansion

$$\hat{T}_2 = \frac{1}{2}\sum_{abij} T_{ij}^{ab}\bar{E}_{ai}\bar{E}_{bj}, \qquad E_{ai} = \sum_{\sigma} a_{a\sigma}^{\dagger}\bar{a}_{i\sigma} \;\; a,b\text{: virtual, } i,j\text{: occupied}$$

In the N^4 MP1-equations

$$\langle {}_{ij}^{ab}|\hat{H} + [\hat{F}, \hat{T}_2]|\mathrm{HF}\rangle = 0 \tag{1}$$

gives the N^4 equations(χ-indeces neglected)

$$\sum_{cdkl}(\mathbf{QS})_{ac}(\mathbf{QS})_{bd}(\mathbf{PS})_{ki}(\mathbf{PS})_{lj}\left(\bar{g}_{ckdl} + (\mathbf{WT})_{kl}^{cd}\right) = 0$$

$$\mathbf{W} = \bar{\mathbf{F}}\otimes\mathbf{1}\otimes\mathbf{1}\otimes\mathbf{1} + \mathbf{1}\otimes\bar{\mathbf{F}}\otimes\mathbf{1}\otimes\mathbf{1} + \mathbf{1}\otimes\mathbf{1}\otimes\bar{\mathbf{F}}\otimes\mathbf{1} +$$

The **QSQSPSPS**-term come from the general anticommutation; previously neglected by us and others

The essential equations

$\sum_{cdkl}(\mathbf{QS})_{ac}(\mathbf{QS})_{bd}(\mathbf{PS})_{ki}(\mathbf{PS})_{lj}\left(\bar{g}_{ckdl}+(\mathbf{WT})^{cd}_{kl}\right)=0$

Standard(Previous approach)

- Solve $\bar{g}_{ckdl}+(\mathbf{WT})^{cd}_{kl}=0$
- Solves in addition to the essential equations also a set of nonessential equations
- The nonessential equations are typically not automatically fulfilled $\longrightarrow$ slower convergence

Current approach

- Include projection operators to solve only the essential conditions
- Each iteration becomes more complicated, but (hopefully) faster convergence

Improve convergence by modifying the essential equations

$\sum_{cdkl}(\mathbf{QS})_{ac}(\mathbf{QS})_{bd}(\mathbf{PS})_{ki}(\mathbf{PS})_{lj}\left(\bar{g}_{ckdl}+(\mathbf{WT})^{cd}_{kl}\right)=0$

Change W

- Any matrix **M** may be added to **W** without changing the solution if **QSQSPSPSM = 0**
- Does not change solution, but may improve convergence by giving equations that are more easy to precondition
- Exploits that the a singular set of linear equations may be modified without modifying the solution

Possible choice of **M**: **W** with interchange of blocks in $\bar{\mathbf{F}}$

$\bar{\mathbf{F}}$, standard

$$\mathbf{F}=\begin{pmatrix} {}^{\mathbf{P}}\bar{\mathbf{F}} & \mathbf{0} \\ \mathbf{0} & {}^{\mathbf{Q}}\bar{\mathbf{F}} \end{pmatrix}$$

$\bar{\mathbf{F}}$, blocks interchanged

$$\mathbf{A}=\begin{pmatrix} {}^{\mathbf{Q}}\bar{\mathbf{F}} & \mathbf{0} \\ \mathbf{0} & {}^{\mathbf{P}}\bar{\mathbf{F}} \end{pmatrix}$$

Example of convergence of iterative method

C_6H_2 in various basis sets, DIIS with diagonal preconditioner

Improvements compared to standard iterative approach

- Use of general local basis $\longrightarrow$ convergence
- Focus on the essential equations $\longrightarrow$ improved convergence
- $\mathbf{W}$ replaced by $\mathbf{W} + \mathbf{M}$ $\longrightarrow$ improved convergence

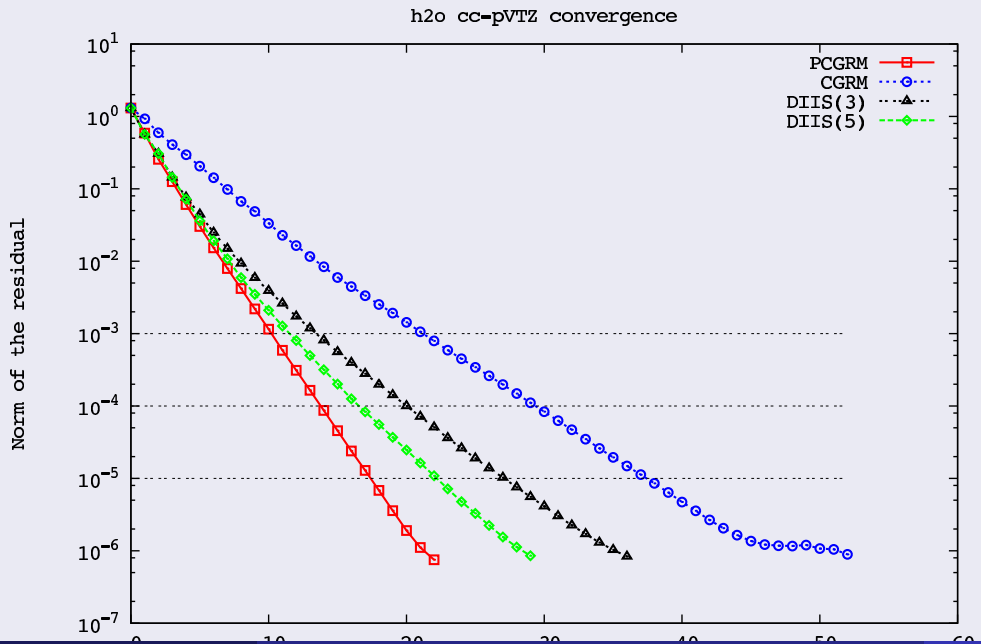
Number of iterations as function of convergence threshold

Basis	Threshold	All eqs. $\mathbf{W}$	Ess. eqs. $\mathbf{W}$	All eqs. $\mathbf{W}+\mathbf{M}$	Ess. eqs. $\mathbf{W}+\mathbf{M}$
STO-3G	10^{-3}	9	8	11	7
STO-3G	10^{-5}	17	13	24	11
cc-pVTZ	10^{-3}	34	24	24	19
cc-pVTZ	10^{-5}	94	53	62	39

The number of iterations may be reduced by upto a factor of 3

Solution of the linear equations using DIIS

- To solve the linear equations we employ DIIS, rather than PCG
- The traditional formulation of DIIS depends on the dimension of the subspace, also for linear equations with symmetric matrices
- Reformulation ensures that three vectors in subspace is mathematically identical to using all vectors in subspace
- Ingredients: save optimal vectors and use norm containing preconditioner



Preliminary scalings

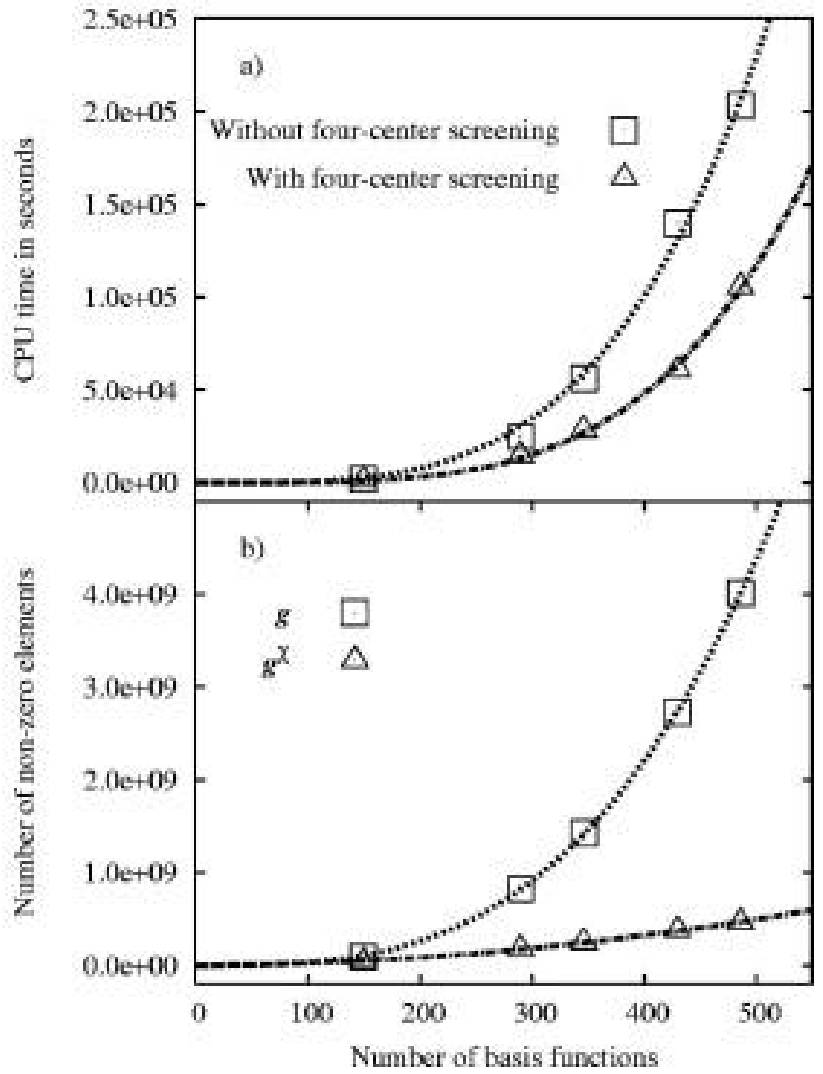
C_nH_2 for $n = 10, 20, 24, 30, 34$ in the cc-pVDZ basis

Upper figure: CPU time

Far from linear scaling (fourth-order scaling!)

Lower figure: # of integrals $> 10^{-8}$

- Two-electron integrals g in the projected basis have significant more non-vanishing elements than in the atomic basis g^{χ}
- Number of projected integrals $\approx$ cubic scaling
- Number of atomic integrals $\approx$ quadratic scaling



Conclusion

- A theoretical framework for working with singular basis sets has been developed
- Transformation to an orthonormal basis before projecting improves convergence significantly
- By solving only the essential equations and modyfying the equations, the iterations may be reduced with upto a factor of 3
- Orbitals obtained by symmetric orthogonalization converge somewhat faster and regular than orbitals obtained by Cholesky-factorization
- Major formal and computational restrictions have been eliminated
- A long way to go...

Acknowledgements

Financial support

- Danish Research Council for a research center
- Finnish Cultural Foundation and the Finnish Academy
- The NANOQUANT EU-network
- CSC in Espoo and DCSC for computational resources
- The Lundbeck Foundation

The papers

- O. Christiansen *et al*, JCP, **124**, 084103 (2006): our initial discussion of approach
- V. Weijo *et al*, JCP, **127**, 074106 (2007): New SQ + numerical examples
- B. Jansik *et al*, JCP, **126**, 124104 (2007): $\mathbf{S}^{-\frac{1}{2}}$