

New techniques of self-consistent field theory

Trygve Helgaker

Centre for Theoretical and Computational Chemistry, University of Oslo, Norway

Filip Pawłowski, Simen Reine, Erik Tellgren

Stinne Høst, Branislav Jansík, Poul Jørgensen, Jeppe Olsen, University of Aarhus, Denmark

Sonia Coriani, University of Trieste, Italy

Paweł Sałek, Royal Institute of Technology, Sweden

Norwegian Theoretical Chemistry

From Molecules To Nanostructures

Britannia Hotel, Trondheim, Norway

November 8–11, 2007

The one-electron density matrix

- Traditional SCF theory rests on the concept of molecular orbitals (MOs):

$$\mathbf{FC} = \mathbf{SC}\epsilon \leftarrow \text{Roothaan-Hall equations}$$

- cost of diagonalization is cubic—expensive for large systems
- ill suited to parallelization

- Fortunately, MOs are not needed: all information is contained in the density matrix

$$\mathbf{D} = \mathbf{C}_{\text{occ}}^{\text{T}} \mathbf{C}_{\text{occ}} \leftarrow \text{density matrix}$$

- it is sufficient to work in terms of the one-electron density matrix
- this is particularly advantageous if it can be compactly (sparsely) represented

- This approach is taken in our work, but then several questions arise:

- Q1: In what basis do we represent the density matrix?
- Q2: How do we optimize the density matrix, without recourse to orbitals?
- Q3: How do we calculate its response to perturbations?

- All tasks should be accomplished in terms of elementary matrix manipulations

- additions, multiplications, trace operations—no diagonalizations
- with sparsity, linear scaling is then achievable for large systems
- diagonalization-free methods are simpler to parallelize

Overview

- the generation of a sparse, orthonormalized atomic-orbital (OAO) basis
 - diagonalization-free evaluation of $\mathbf{S}^{\pm 1/2}$
 - multiplication-rich globally convergent iterative method
 - Jansik *et al.*, *J. Chem. Phys.* **126**, 124104 (2007)
- energy optimization for large systems
 - direct optimization of density matrix \mathbf{D} in OAO basis
 - multiplication-rich diagonalization-free self-consistent field (SCF) theory
 - Salek *et al.*, *J. Chem. Phys.* **126**, 114110 (2007)
 - multiplication-rich quasi-Newton method, superior to RH/DIIS
- linear response theory for large systems
 - direct perturbation of density matrix \mathbf{D} in OAO basis
 - calculation of polarizabilities and excitation energies
 - Coriani *et al.* *J. Chem. Phys.* **126**, 154108 (2007)

Part I: Orthonormalization of the AO basis

- For large systems, the **atomic-orbital (AO) basis** is ideal in the sense that it is sparse
 - unfortunately, many computational tasks are ill-conditioned in this basis
- This may be avoided by transforming to an **orthonormalized AO (OAO) basis**

$$\mathbf{Z}^T \mathbf{S} \mathbf{Z} = \mathbf{I}, \quad \mathbf{Z} = \begin{cases} \mathbf{S}^{-1/2} & \text{Löwdin} \\ \mathbf{U}^{-1} \text{ (upper triangular)} & \text{Cholesky} \end{cases}$$

- Millam and Scuseria, JCP **106**, 5569 (1997); Challacombe, JCP **110**, 2332 (1999)
- The **Löwdin basis** resembles the AO basis most closely (Carlson and Keller, 1957)
 - it should therefore retain sparsity to the greatest possible extent
 - however, its construction traditionally involves diagonalization
- The **Cholesky basis** has therefore been preferred instead
 - many uses in quantum chemistry (rank reduction techniques)
- However, since the Löwdin basis is (presumably) the sparsest of all OAO bases, it would be surprising if it could not be generated in linear time!
- We have recently presented a robust linear-scaling algorithm for Löwdin factorization
 - Jansik *et al.*, J. Chem. Phys. **126**, 124104 (2007)

A globally convergent Newton–Schulz method for the inverse square root

- Application of Newton’s elementary method for root finding gives

$$Z^{-2} - S = 0 \quad \Rightarrow \quad Z_{n+1} = \frac{3}{2}Z_n - \frac{1}{2}Z_n^3S \quad \leftarrow \text{Newton–Schulz iteration}$$

- Z_n converges quadratically but not globally to $S^{-1/2}$; no inversion!

- In 2004, Niklasson proposed the corresponding matrix iteration

$$\mathbf{Z}_{n+1} = \frac{1}{2}\mathbf{Z}_n \left(3\mathbf{I} - \mathbf{Z}_n^T \mathbf{S} \mathbf{Z}_n \right) \quad \leftarrow \text{stabilized Newton–Schulz matrix iteration}$$

- higher-order formulas may also be derived
- for $\mathbf{Z}_0 = \mathbf{I}$, there are severe convergence restrictions $\|\mathbf{S} - \mathbf{I}\|_2 < 1$ for such methods
- intended for refinement of approximate inverse square roots

- However, convergence is guaranteed with the scaled $\lambda_{\min} \mathbf{S}$ that minimizes $\|\lambda \mathbf{S} - \mathbf{I}\|_2$:

$$\lambda_{\min} = \frac{2}{\epsilon_{\min} + \epsilon_{\max}} \quad \leftarrow \text{expensive}$$

- to avoid eigenvalues, we replace the 2-norm by a ratio of Frobenius norms

$$f(\lambda) = \sqrt{\frac{\text{Tr}(\lambda \mathbf{S} - \mathbf{I})^4}{\text{Tr}(\lambda \mathbf{S} - \mathbf{I})^2}} \quad \leftarrow \text{cheap lower bound to the 2-norm}$$

- the evaluation and minimization of $f(\lambda)$ costs only **one** matrix multiplication

Number of iterations for convergence to 10^{-10}

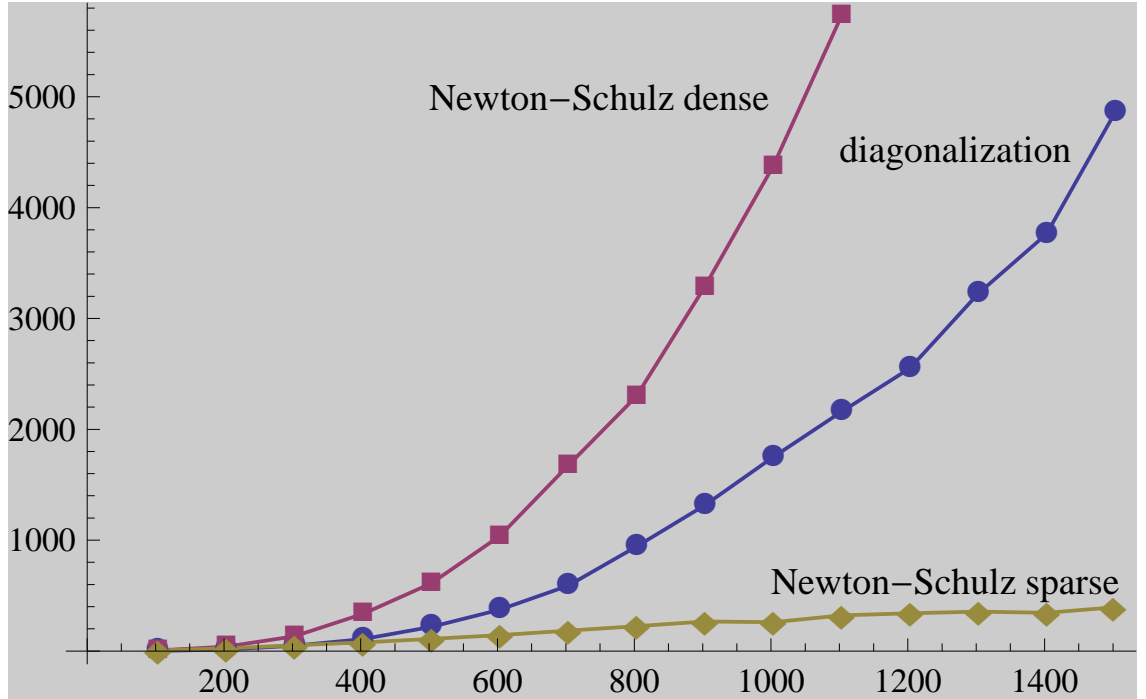
	$\ \mathbf{S} - \mathbf{I}\ _2$	$N^{(2)}$	$N^{(3)}$
Fullerene C₆₀	6.0	17	11
Polysaccharide (438)	5.2	14	10
Polysaccharide (875)	5.2	14	10
DNA fragment (583)	4.9	15	11
PRC model (1082)	5.7	16	11
Red phosphorus (506)	1.9	8	6
Red phosphorus (810)	4.7	14	10
Crystobalite (984)	4.4	13	9
Tridymite (1314)	5.2	15	11
Random molecule (64)	16.0	23	27

Performance of the scaled Newton–Schulz method

- About 15 scaled Newton–Schulz iterations needed for convergence to 10^{-10}
 - less than 50 matrix multiplies
 - example: polyalanine peptide residues (6-31G) (cost vs. number of atoms)

- In dense-matrix algebra:
 - of cubic complexity
 - two–three times more expensive than diagonalization
 - easy to parallelize
 - beats diagonalization on 4–6 processors

- In sparse-matrix algebra:
 - of linear complexity
 - beats diagonalization for a few hundred atoms



- Transformation to orthonormal basis is an essential preparatory step to optimization

Part II: Energy optimization

- Consider the direct optimization of the density matrix:

$$E(\mathbf{D}) = \text{Tr } \mathbf{D}\mathbf{h} + \text{2-el. part}$$

- there are constraints on the density matrix:

$$\underbrace{\mathbf{D} = \mathbf{D}^T}_{\text{symmetry}}, \quad \underbrace{\text{Tr } \mathbf{D} = N}_{\text{trace}}, \quad \underbrace{\mathbf{D}^2 = \mathbf{D}}_{\text{idempotency}} \quad (\text{orthonormal basis})$$

- any optimization must obey these constraints

- Many strategies are based on purification of the density matrix

$$\tilde{\mathbf{D}} = 3\mathbf{D}^2 - 2\mathbf{D}^3 \quad (\text{McWeeny purification, 1960})$$

- Li, Nunes and Vanderbilt (1993)

$$\tilde{E} = \text{Tr } \tilde{\mathbf{D}}\mathbf{h} + \mu(\text{Tr } \mathbf{D} - N) + \text{2-el. part}$$

- Millam and Scuseria (1997), Challacombe (1999)
- Palser and Manolopoulos (1998), Niklasson (2002)

- We shall pursue a different approach, based on an explicitly parameterization of \mathbf{D}

Exponential parameterization of the density matrix

- In a real, nonorthogonal AO basis, with $\mathbf{S} \neq \mathbf{I}$, let \mathbf{D} be a valid HF/KS matrix:

$$\underbrace{\mathbf{D} = \mathbf{D}^T}_{\text{symmetry}}, \quad \underbrace{\text{Tr } \mathbf{D} \mathbf{S} = N}_{\text{trace}}, \quad \underbrace{\mathbf{D} \mathbf{S} \mathbf{D} = \mathbf{D}}_{\text{idempotency}}$$

- Any other valid density matrix $\mathbf{D}(\mathbf{X})$ can then be generated from this matrix:

$$\underbrace{\mathbf{D}(\mathbf{X}) = \exp(-\mathbf{X} \mathbf{S}) \mathbf{D} \exp(\mathbf{S} \mathbf{X})}_{\text{exponential parameterization}}, \quad \underbrace{\mathbf{X}^T = -\mathbf{X}}_{\text{antisymmetric}}$$

- Helgaker, Jørgensen and Olsen: *Molecular Electronic-Structure Theory* (Wiley, 2000)
- Head-Gordon and coworkers, MolPhys **101**, 37 (2003), JCP **118**, 6144 (2003)

- We can obtain any valid density matrix, in the AO basis, without recourse to MOs!
 - in particular, we may optimize the energy by freely varying $X_{\mu\nu}$ with $\mu > \nu$:

$$E_{\min}(\mathbf{X}) = \min_{\mathbf{X}} [\text{Tr } \mathbf{D}(\mathbf{X}) \mathbf{h} + \text{2-el. part}]$$

- Is the use of $\mathbf{D}(\mathbf{X})$ a practical proposition?
 - we shall in this talk demonstrate that it is indeed so

Two questions about $\mathbf{D}(\mathbf{X}) = \exp(-\mathbf{X})\mathbf{D}\exp(\mathbf{X})$

- Can it be evaluated efficiently?

- we use a Baker–Campbell–Hausdorff (BCH) expansion:

$$\mathbf{D}(\mathbf{X}) = \exp(-\mathbf{X})\mathbf{D}\exp(\mathbf{X}) = \mathbf{D} + [\mathbf{D}, \mathbf{X}] + \frac{1}{2} [[\mathbf{D}, \mathbf{X}], \mathbf{X}] + \dots$$

- converges rapidly (purification may be necessary), in about 10 matrix multiplications

- Are redundancies a problem?

- the orbital space consists of two parts: the occupied space and the virtual space

$$\mathbf{P}_o = \mathbf{D} \text{ (onto occupied space)}, \quad \mathbf{P}_v = \mathbf{I} - \mathbf{D} \text{ (onto virtual space)}$$

- only rotations between the occupied and virtual spaces are nonredundant;

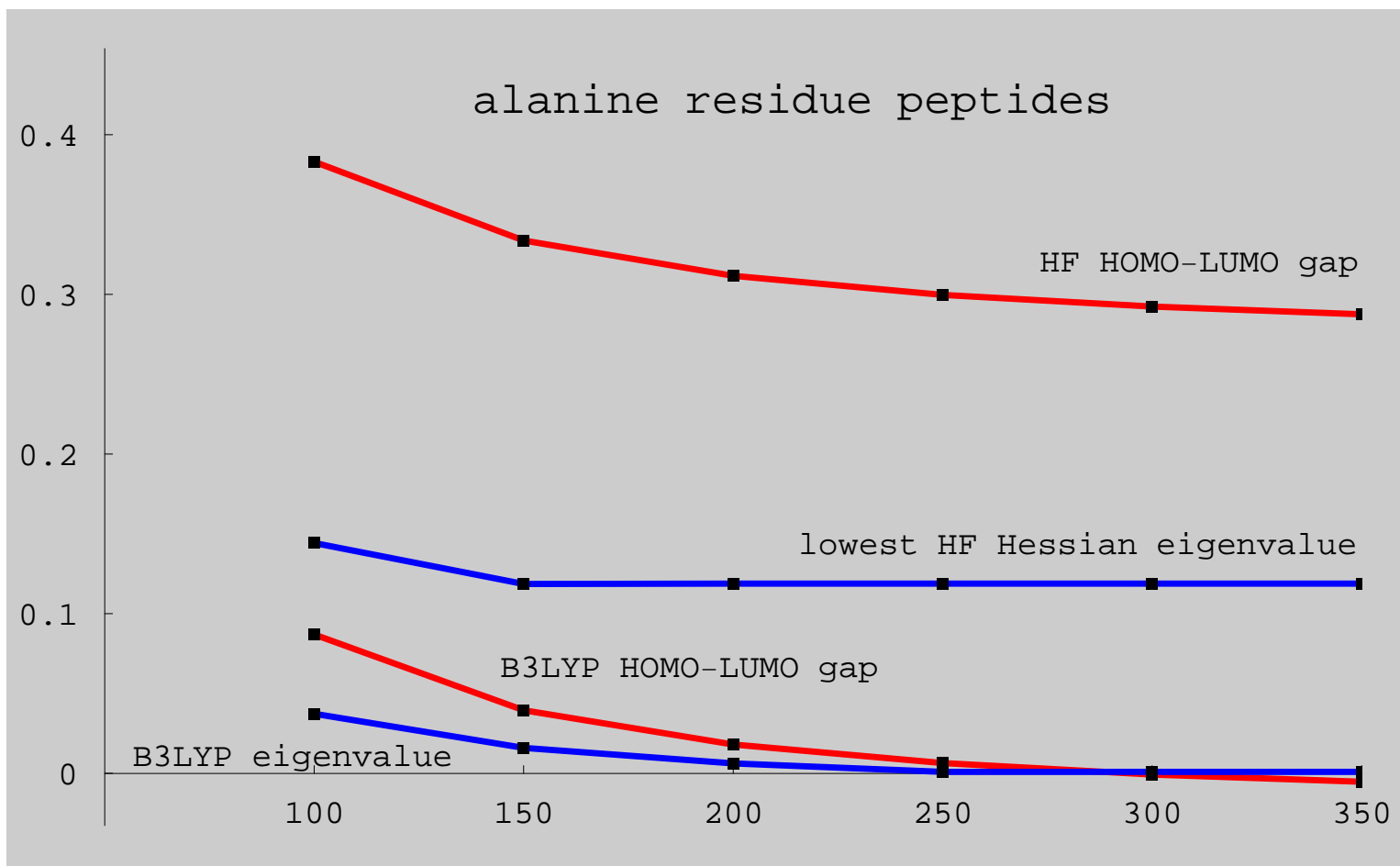
$$\mathbf{X} = \underbrace{\mathbf{P}_o\mathbf{X}\mathbf{P}_o + \mathbf{P}_v\mathbf{X}\mathbf{P}_v}_{\mathbf{X}_{\text{red}}} + \underbrace{\mathbf{P}_o\mathbf{X}\mathbf{P}_v + \mathbf{P}_v\mathbf{X}\mathbf{P}_o}_{\mathbf{X}_{\text{nr}}}$$

- to avoid problems with redundancies, we use the projected parameterization

$$\mathbf{D}(\mathbf{X}) = \exp(-\mathbf{X}_{\text{nr}})\mathbf{D}\exp(\mathbf{X}_{\text{nr}}), \quad \mathbf{X}^T = -\mathbf{X}$$

SCF optimizations in large molecules

- SCF convergence is typically more difficult in larger systems



- small (or negative) HOMO–LUMO gaps and small Hessian eigenvalues in DFT
- 6-31G HOMO–LUMO gap and lowest Hessian eigenvalue in alanine residue peptides

Newton's method applied to the SCF energy

- Let us now consider the iterative optimization of the SCF energy
- At the n th iteration, we parameterize the electronic energy in the manner

$$\mathcal{E}(\mathbf{X}) = E[\mathbf{D}_n(\mathbf{X})] = E[\exp(-\mathbf{X})\mathbf{D}_n \exp(\mathbf{X})]$$

- We expand the energy in orders of X_μ :

$$\mathcal{E}(\mathbf{X}) = \mathcal{E}_0 + \sum_{\mu} \mathcal{F}_{\mu} X_{\mu} + \frac{1}{2} \sum_{\mu\nu} \mathcal{G}_{\mu\nu} X_{\mu} X_{\nu} + \dots$$

- Truncating at second order and setting the gradient to zero, we obtain the Newton step

$$\text{Hessian} \times \text{step} \Rightarrow \sum_{\nu} \mathcal{G}_{\mu\nu} X_{\nu} = -\mathcal{F}_{\mu} \quad \Leftarrow \text{negative gradient}$$

- the iterative application of the Newton step yields fast, second-order convergence
- For guaranteed ground-state convergence, a **trust-region scheme** is employed
 - at each Newton iterations, a best step is obtained subject to constraint $\|\mathbf{X}\| \leq h$
 - constraint introduces level shifting
- We shall now study the Newton method and some approximations to it

The Newton equations

- A basic quantity in SCF optimizations is the Fock (Kohn–Sham) matrix

$$\mathbf{F}_n = \mathbf{F}(\mathbf{D}_n) = \mathbf{h} + \mathbf{G}(\mathbf{D}_n)$$

- it may be decomposed into four different projections:

$$\mathbf{F}_n = \mathbf{F}_n^{\text{oo}} + \mathbf{F}_n^{\text{ov}} + \mathbf{F}_n^{\text{vo}} + \mathbf{F}_n^{\text{vv}}, \quad (\mathbf{P}_o = \mathbf{D}, \mathbf{P}_v = \mathbf{I} - \mathbf{D})$$

- Expressed in terms of the Fock matrix, the Newton equations become

$$(\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) + \frac{d(\mathbf{F}_n^{\text{ov}} - \mathbf{F}_n^{\text{vo}})}{d\mathbf{D}_n} [\mathbf{D}_n, \mathbf{X}] = \mathbf{F}_n^{\text{vo}} - \mathbf{F}_n^{\text{ov}}$$

- the left-hand side represents a linear Hessian transformation of the matrix \mathbf{X}
- all contributions that involve the undifferentiated \mathbf{F}_n are easily constructed
- the contributions that involve $d\mathbf{F}_n/d\mathbf{D}_n$ are expensive
- These Newton equations are solved iteratively (about 10 micro iterations)
 - each micro iteration requires integral and XC reevaluations because of $d\mathbf{F}_n/d\mathbf{D}_n$
- We shall now explore simplifications to this basic SO scheme
 - the Roothaan–Hall (RH) scheme
 - the augmented Roothaan–Hall (ARH) scheme

The Roothaan–Hall Newton method

- Let us first simply **ignore** the time-consuming part of the Newton equations

$$(\mathbf{F}^{\text{vv}} - \mathbf{F}^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}^{\text{vv}} - \mathbf{F}^{\text{oo}}) = \mathbf{F}^{\text{vo}} - \mathbf{F}^{\text{ov}}$$

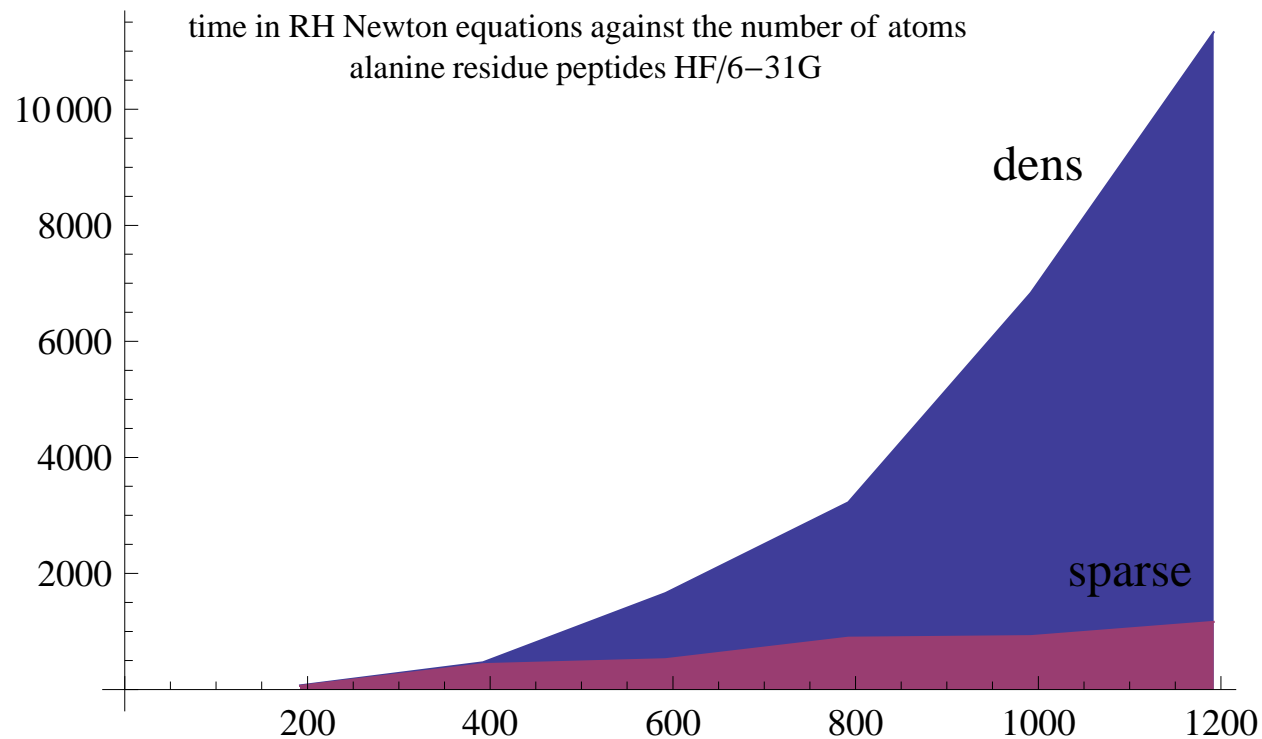
- The solution of these equations is simple and fast
 - **the preconditioned conjugate-gradient method** (typically 10 iterations)
 - elementary (sparse) matrix manipulations (typically less than 100 multiplications)
 - use of an orthonormal basis essential for rapid convergence
- Since we do not calculate the Newton step of the total energy $\mathcal{E}(\mathbf{X})$, convergence suffers
- On the other hand, we are now instead carrying out a Newton minimization of the energy

$$\mathcal{E}_{\text{RH}}(\mathbf{X}) = \text{Tr} \mathbf{D}(\mathbf{X})\mathbf{F} = \sum_i \varepsilon_i \quad \leftarrow \quad \text{sum of orbital energies}$$

- **we are in fact carrying out a Roothaan–Hall diagonalization step!**
- Diagonalization-free SCF optimization is simple—ignore the difficult Hessian terms!
 - we call this approach the **Roothaan–Hall Newton method**
- Larsen *et al.*, J. Chem. Phys. **113** 8908 (2000)
Sąlek *et al.*, *ibid.* **126** 114110 (2007)

Performance of the Roothaan–Hall Newton method

- We have successfully avoided Roothaan–Hall diagonalization
 - minimization rather than the solution of a generalized eigenvalue problem
 - rapidly convergent: 50–100 sparse matrix multiplications needed
 - well suited to parallelization
- Linear scaling is obtained by employing sparse-matrix algebra
 - compressed sparse-row (CSR) representation of few-atom blocks



The augmented Roothaan–Hall method

- The full Newton equations of SCF theory are given by

$$(\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) + d(\mathbf{F}_n^{\text{ov}} - \mathbf{F}_n^{\text{vo}})/d\mathbf{D}_n [\mathbf{D}_n, \mathbf{X}] = \mathbf{F}_n^{\text{vo}} - \mathbf{F}_n^{\text{ov}}$$

- For efficiency, it is not necessary to neglect the derivative (red) contribution completely
 - during the optimization, we have built up a set of density matrices $\mathbf{D}_{in} = \mathbf{D}_i - \mathbf{D}_n$
 - derivatives of the Fock/KS matrix (approximately) are known in these directions

$$\frac{d\mathbf{F}_n(\mathbf{D}_n)}{d\mathbf{D}_n} \mathbf{D}_{in} \approx \mathbf{F}(\mathbf{D}_i) - \mathbf{F}(\mathbf{D}_n) = \mathbf{F}_i - \mathbf{F}_n = \mathbf{F}_{in} \leftarrow \text{available}$$

- Let us therefore project $[\mathbf{D}_n, \mathbf{X}]$ onto the set of \mathbf{D}_{in}

$$\frac{d(\mathbf{F}_n^{\text{ov}} - \mathbf{F}_n^{\text{vo}})}{d\mathbf{D}_n} [\mathbf{D}_n, \mathbf{X}] \approx \frac{d(\mathbf{F}_n^{\text{ov}} - \mathbf{F}_n^{\text{vo}})}{d\mathbf{D}_n} \mathbf{P}_n(\mathbf{X}) [\mathbf{D}_n, \mathbf{X}] = \sum_i (\mathbf{F}_{in}^{\text{ov}} - \mathbf{F}_{in}^{\text{vo}}) P_i(\mathbf{X})$$

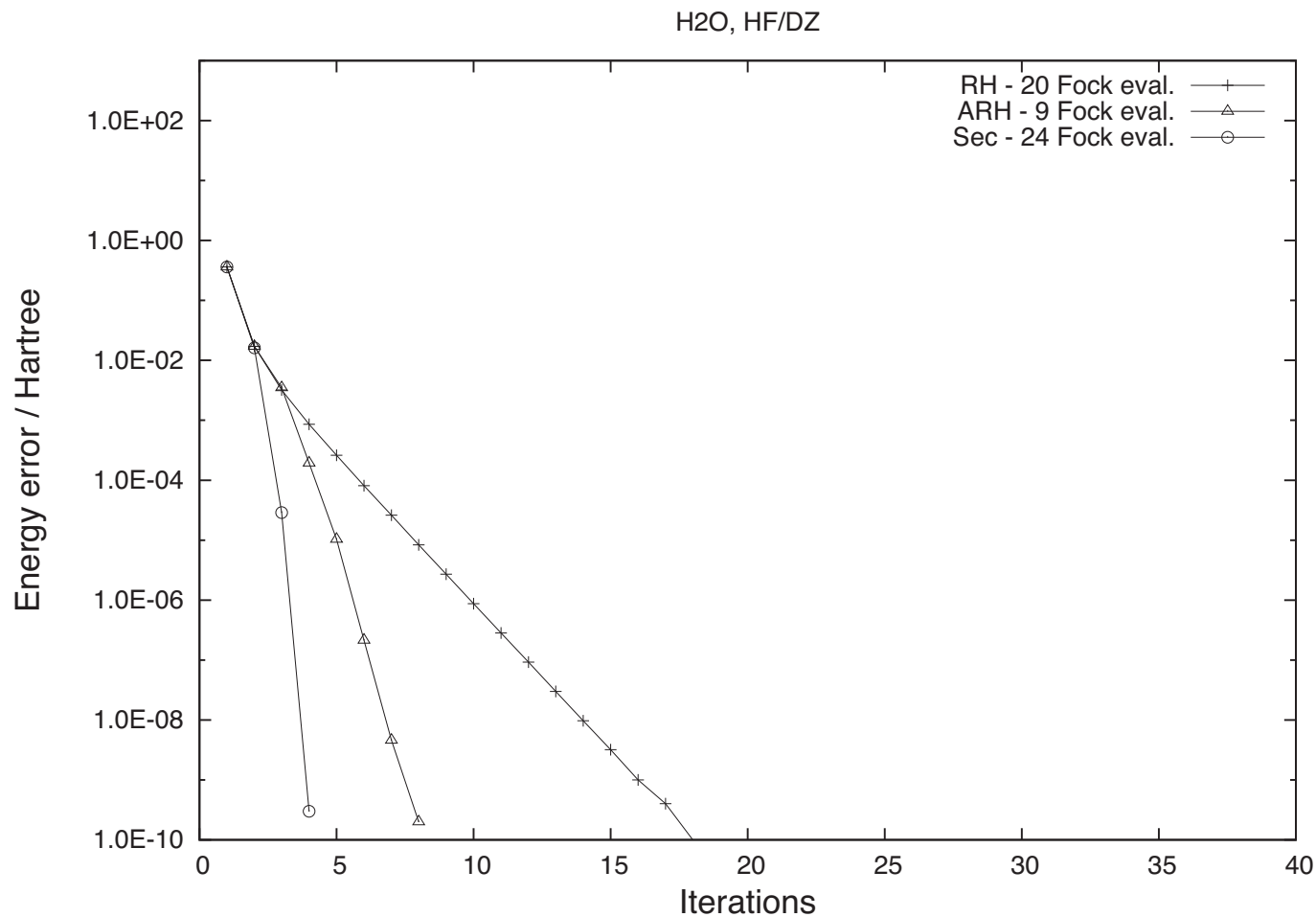
- this evaluation is inexpensive, requiring **no new Fock/KS matrix evaluations**
- we can thus evaluate the Hessian cheaply and accurately in the \mathbf{D}_{in} directions
- The resulting method is called **the augmented Roothaan–Hall method (ARH)**

$$(\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) + \sum_i (\mathbf{F}_{in}^{\text{ov}} - \mathbf{F}_{in}^{\text{vo}}) P_i(\mathbf{X}) = \mathbf{F}_n^{\text{vo}} - \mathbf{F}_n^{\text{ov}}$$

- the ARH method is vastly superior to the standard RH scheme (with DIIS?)

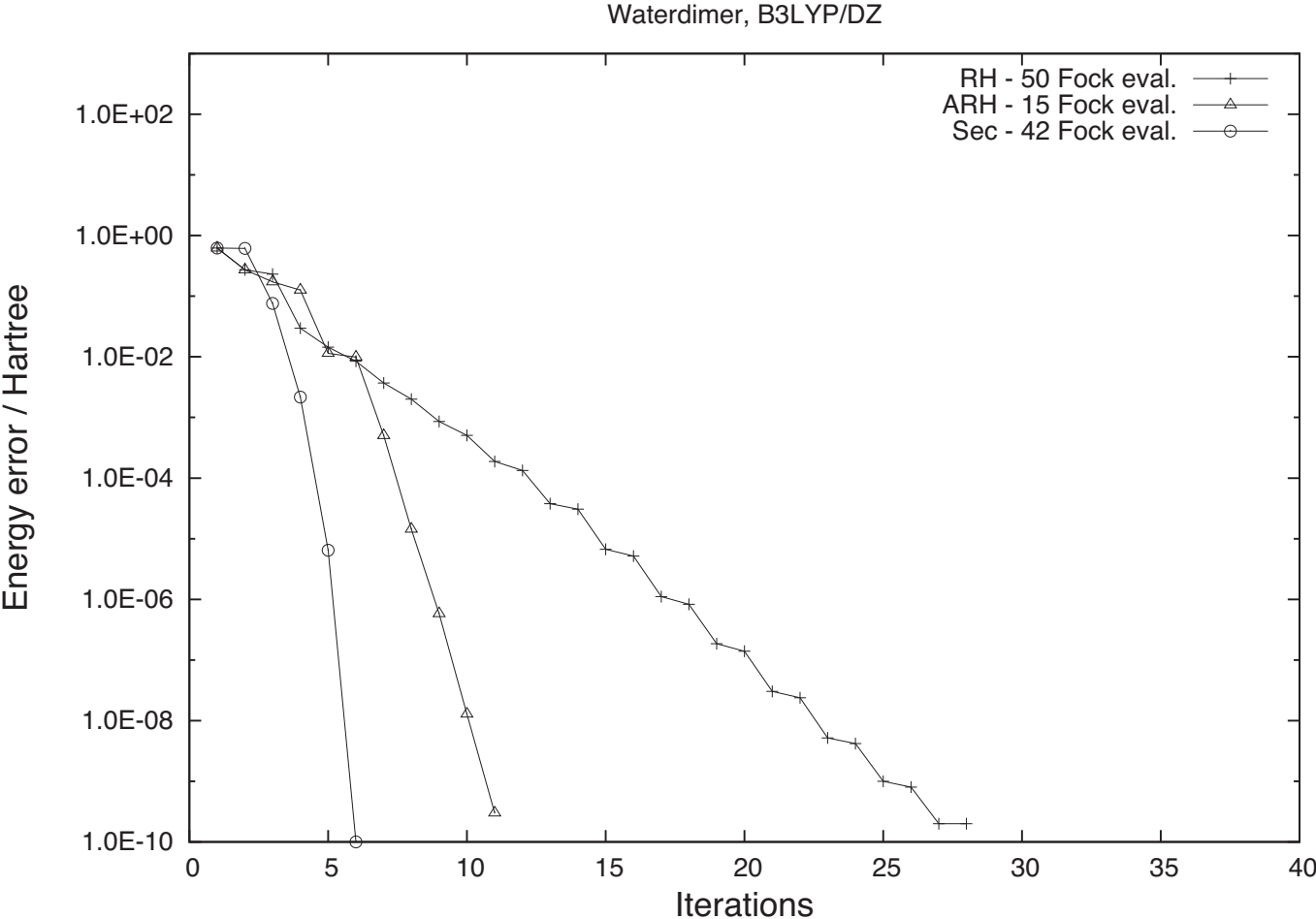
A comparison of RH, augmented RH, and SO optimizations

- water HF/STO-3G
 - 20 (RH), 9 (ARH), and 24 (SO) Fock-matrix evaluations



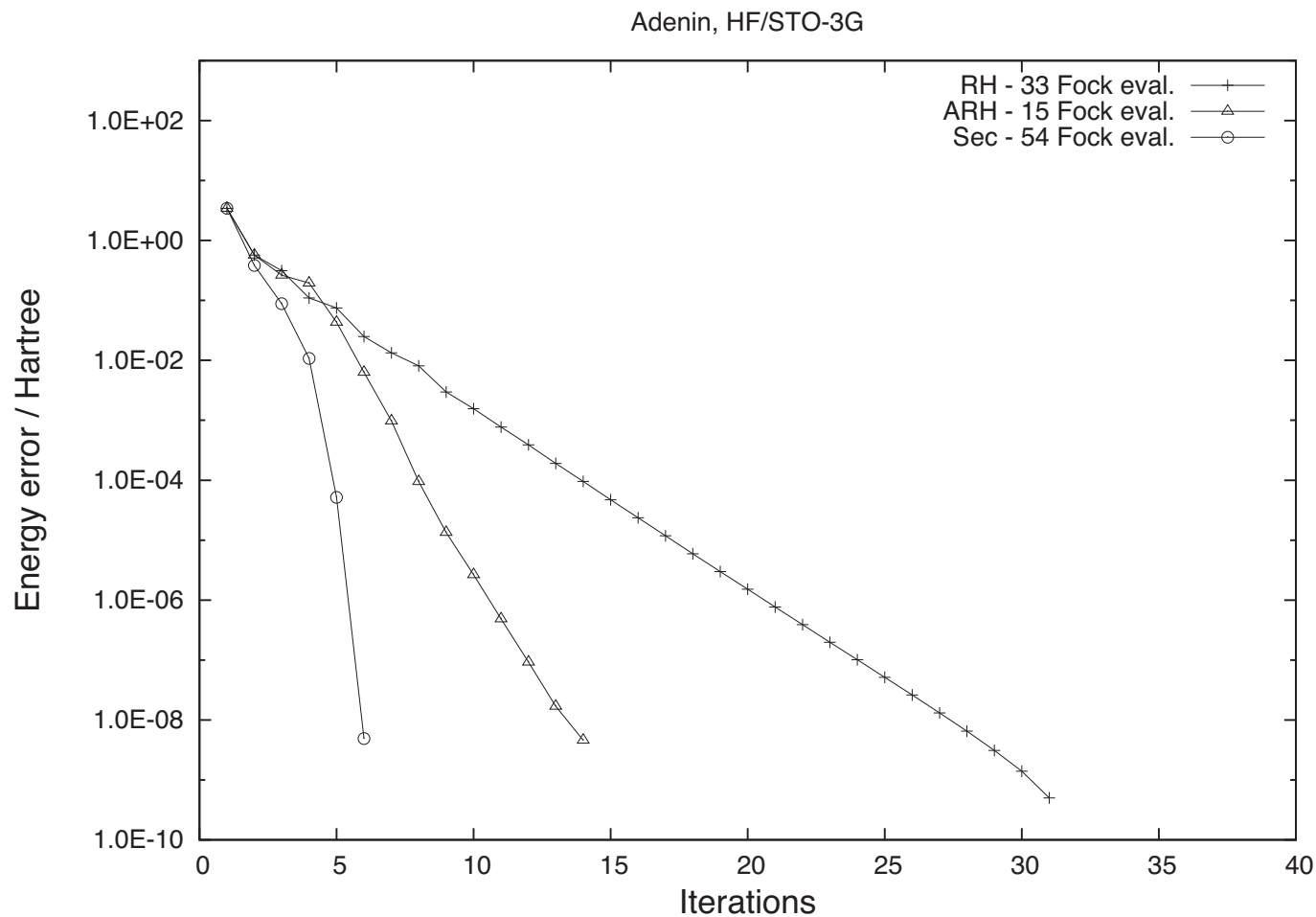
A comparison of RH, augmented RH, and SO optimizations

- water dimer B3LYP/DZ
 - 50 (RH), 15 (ARH), and 42 (SO) KS-matrix evaluations



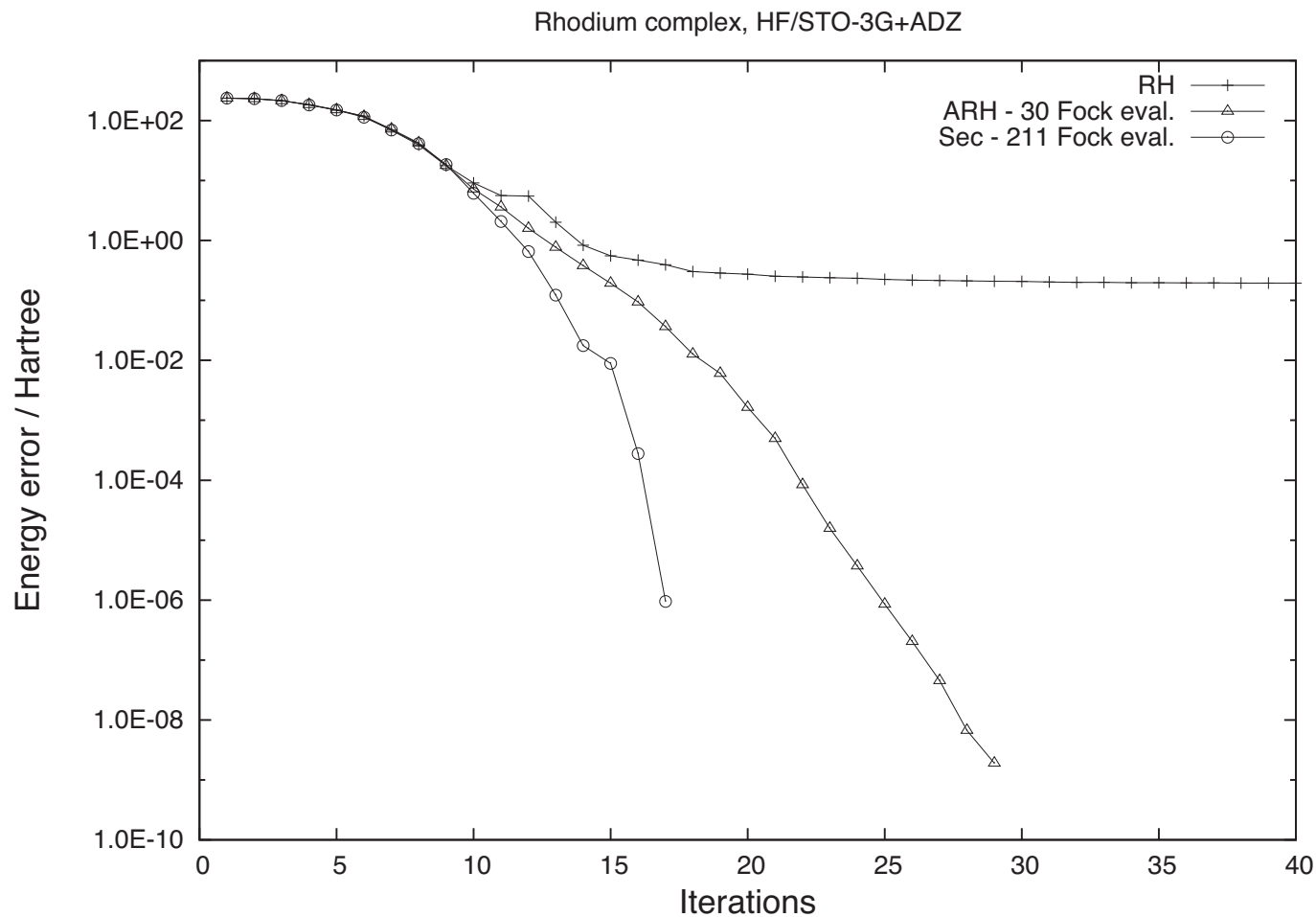
A comparison of RH, augmented RH, and SO optimizations

- adenine HF/STO-3G
 - 33 (RH), 15 (ARH), and 54 (SO) Fock-matrix evaluations



A comparison of RH, augmented RH, and SO optimizations

- rhodium complex HF/STO-3G+ADZ
 - ∞ (RH), 30 (ARH), and 211 (SO) Fock-matrix evaluations



Relationship to the RH/DIIS method

- The ARH Newton step is given by

$$(\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) + \sum_{i=1}^{n-1} (\mathbf{F}_{in}^{\text{ov}} - \mathbf{F}_{in}^{\text{vo}}) P_i(\mathbf{X}) = \mathbf{F}_n^{\text{vo}} - \mathbf{F}_n^{\text{ov}}$$

- Hessian exact in the directions of all previous density matrices \mathbf{D}_{in}

- Let us now rewrite the ARH equations slightly

$$(\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) \mathbf{X} + \mathbf{X} (\mathbf{F}_n^{\text{vv}} - \mathbf{F}_n^{\text{oo}}) = \sum_{i=1}^n c_i(\mathbf{X}) (\mathbf{F}_i^{\text{ov}} - \mathbf{F}_i^{\text{vo}}), \quad \sum_{i=1}^n c_i(\mathbf{X}) = 1$$

- the ARH step is therefore the RH step with a modified, averaged gradient
- reminiscent of RH diagonalization with DIIS-averaged Fock matrix!

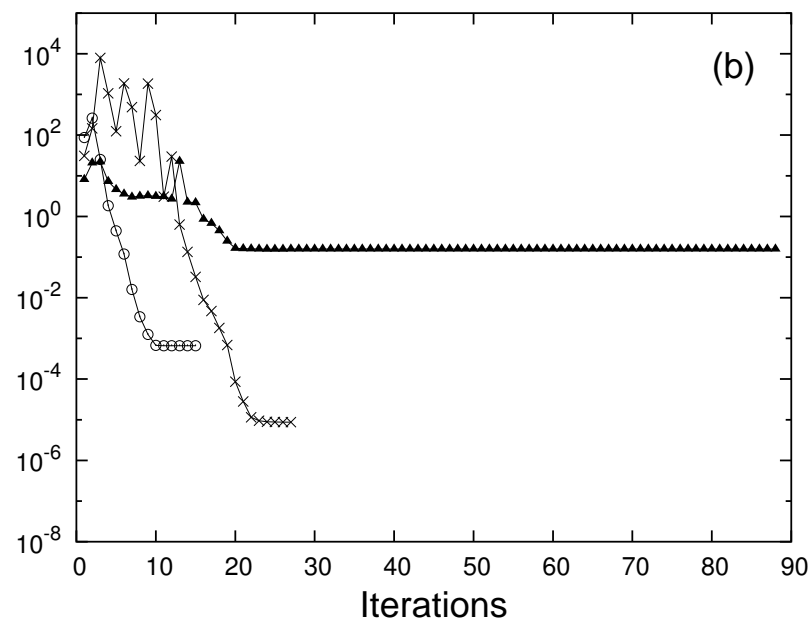
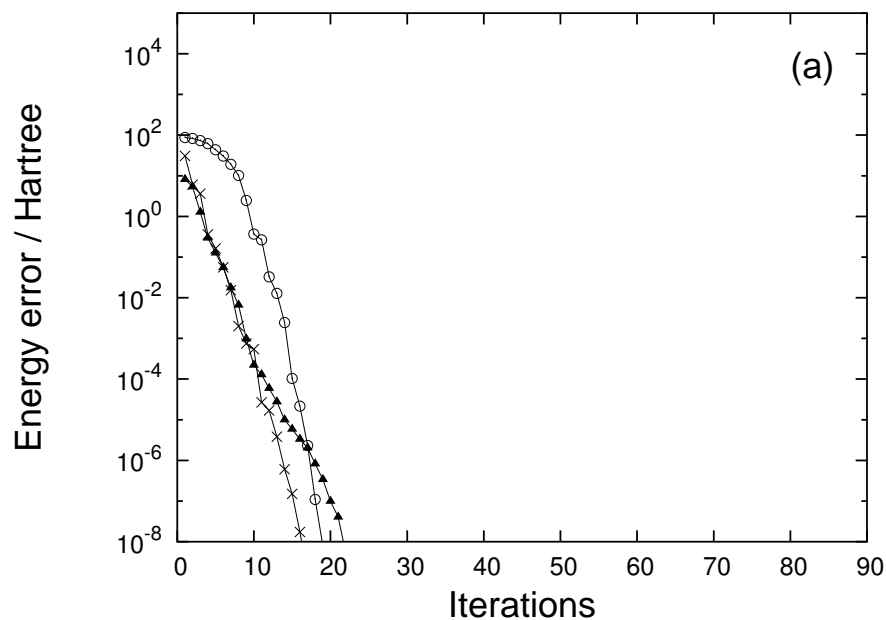
- RH/DIIS obtained by solving the ARH Newton equations in two steps

1. neglect the left-hand side and solve the resulting equations for c_i
2. introduce an averaged right-hand side in RH equations

- However, unlike ARH, RH/DIIS neglects the coupling between the two steps

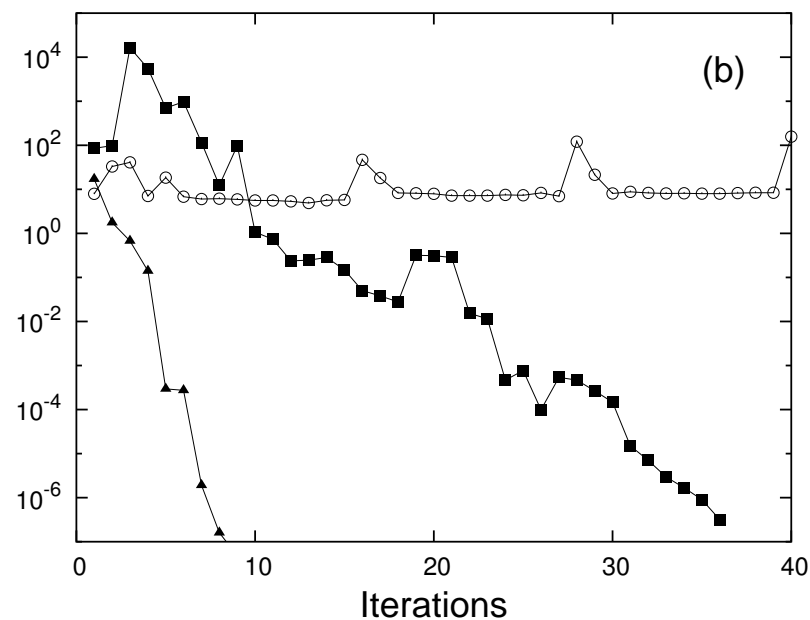
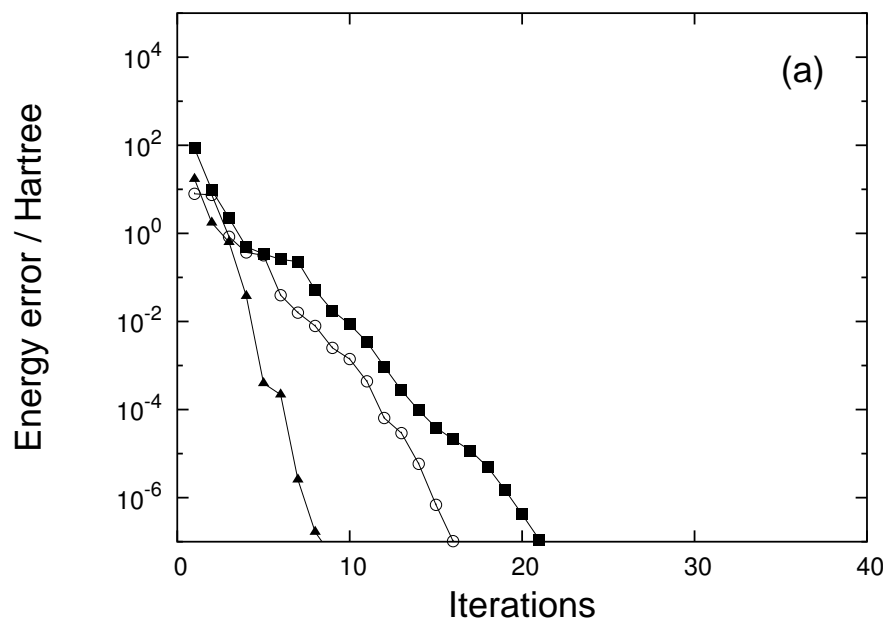
The ARH method compared with RH/DIIS (I)

- Convergence for a variety of molecules
 - cadmium–imidazole complex (circles) B3LYP/3-21G (10/89)
 - 29-residue polyalanine peptide (crosses) B3LYP/6-31G (292/1599)
 - model B12 vitamin (triangles) HF/AhlrichsVDZ (74/428)
- Energy-error on a logarithmic scale (ARH left, RH/DIIS right)
 - RH/DIIS converges to a saddle point, ARH always to a minimum



The ARH method compared with RH/DIIS (II)

- Convergence of B3LYP calculations for a variety of molecules
 - cluster of 51 water molecules (full triangles) B3LYP/cc-pVTZ (153/2958)
 - model B12 vitamin (empty circles) BP86/AhrlrichsVDZ (74/428)
 - insulin (full squares) B3LYP/6-21G (786/4417)
- Energy-error on a logarithmic scale (ARH left, RH/DIIS right)
 - ARH converges smoothly, RH/DIIS oscillates (and diverges)



Part III: Response theory

- We consider a system described by the one-electron density matrix

$$\mathbf{D}(\mathbf{X}) = \exp(-\mathbf{X})\mathbf{D}\exp(\mathbf{X}) \quad (\mathbf{X} = \mathbf{0} \text{ for unperturbed system})$$

and define the Hessian and metric operators in terms of their transformations

$$\mathbf{E}^{[2]}(\mathbf{X}) = (\mathbf{F}^{\text{vv}} - \mathbf{F}^{\text{oo}})\mathbf{X} + \mathbf{X}(\mathbf{F}^{\text{vv}} - \mathbf{F}^{\text{oo}}) + \mathbf{G}^{\text{vo}}([\mathbf{D}, \mathbf{X}]) - \mathbf{G}^{\text{ov}}([\mathbf{D}, \mathbf{X}])$$

$$\mathbf{S}^{[2]}(\mathbf{X}) = \mathbf{X}^{\text{ov}} - \mathbf{X}^{\text{vo}}$$

- When perturbed by \mathbf{V}_ω of frequency ω , the system responds to first order as

$$\mathbf{E}^{[2]}(\mathbf{X}_\omega) - \omega\mathbf{S}^{[2]}(\mathbf{X}_\omega) = [\mathbf{D}, \mathbf{V}_\omega] \quad \leftarrow \text{linear response matrix equation}$$

– perturbed density matrix:

$$\mathbf{D}_\omega = [\mathbf{D}, \mathbf{X}_\omega]$$

– perturbed expectation values: $\langle\langle \hat{A}; \hat{V}_\omega \rangle\rangle_\omega = \text{Tr } \mathbf{A} [\mathbf{D}, \mathbf{X}_\omega]$

- In the absence of a perturbation $\mathbf{V}_\omega = \mathbf{0}$, we obtain an eigenvalue equation

$$\mathbf{E}^{[2]}(\mathbf{X}_n) = \omega_n\mathbf{S}^{[2]}(\mathbf{X}_n) \quad \leftarrow \text{RPA matrix eigenvalue equation}$$

– transition density matrix:

$$\mathbf{D}_{0n} = [\mathbf{D}, \mathbf{X}_n]$$

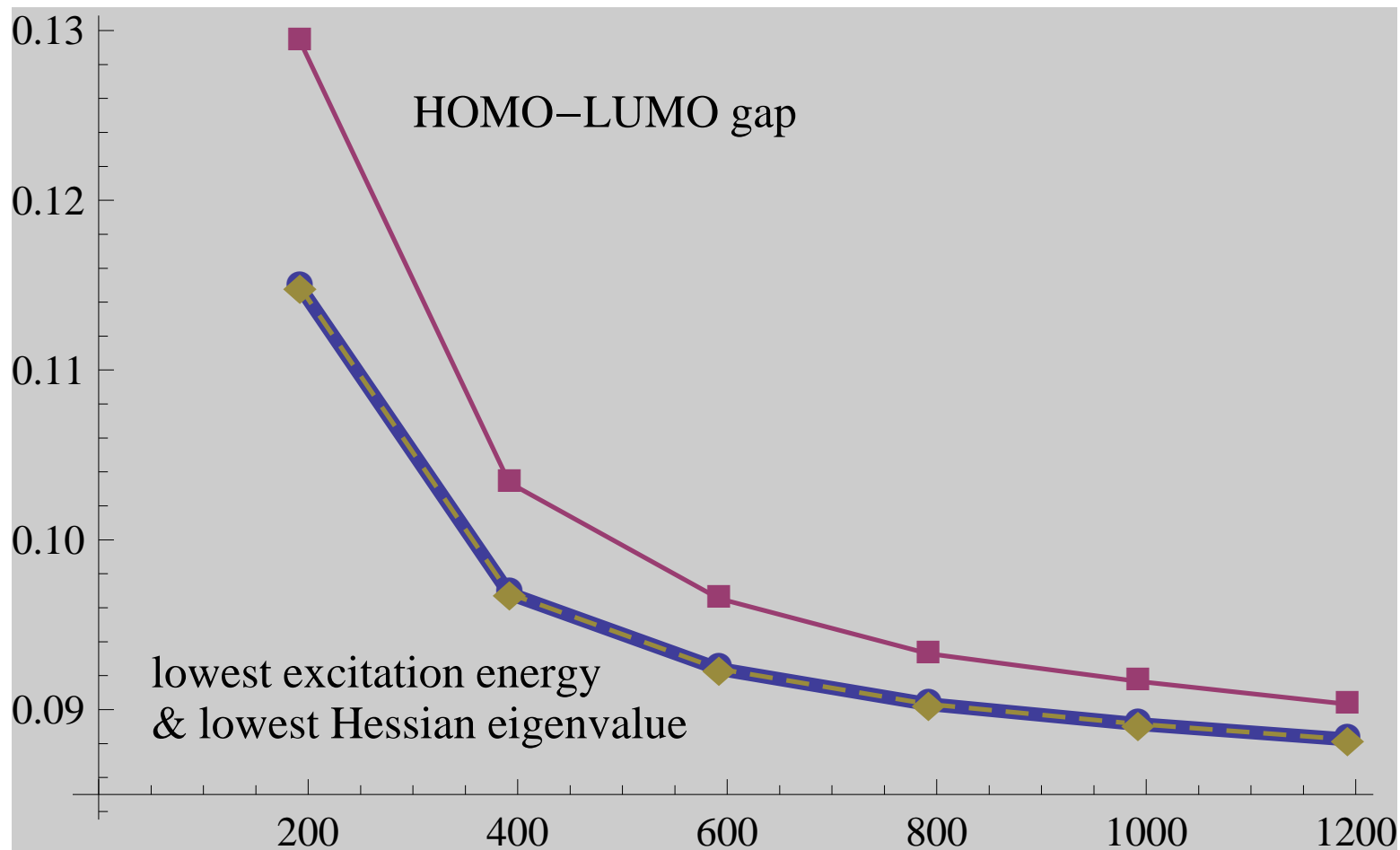
– transition moments:

$$\langle 0 | \hat{A} | n \rangle = \text{Tr } \mathbf{A} [\mathbf{D}, \mathbf{X}_n]$$

- Coriani *et al.*, J. Chem. Phys. **126**, 154108 (2007)

– Ochsenfeld, Head-Gordon, Weber, Niklasson, and Challacombe (static properties)

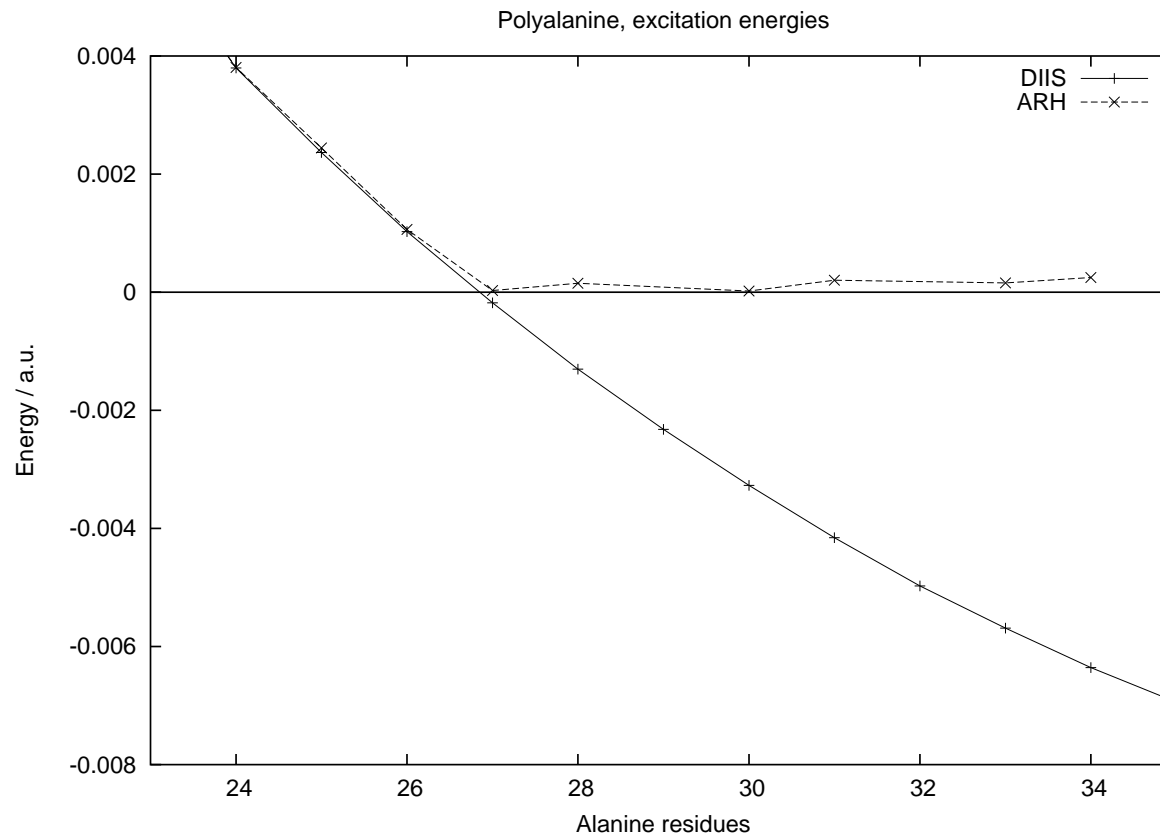
CAMB3LYP/6-31G excitation energies of alanine residue peptides



- lowest excitation energy almost identical with lowest Hessian eigenvalue
- CAMB3LYP has 19% short-range and 65% long-range exact exchange

B3LYP/6-31G excitation energies of alanine residue peptides

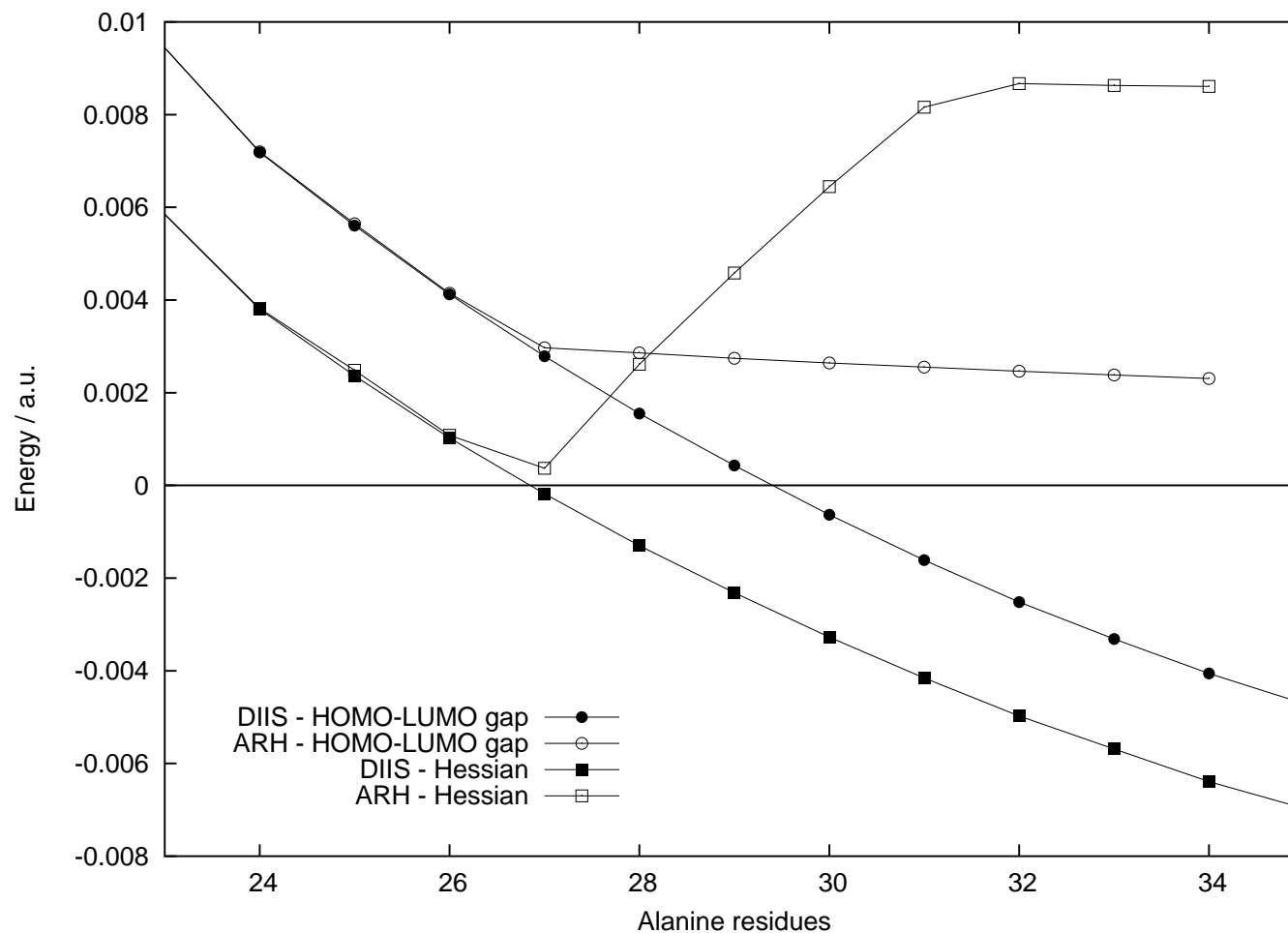
- B3LYP (with 20% exact exchange) behaves differently, in an unphysical manner
 - beyond 27 alanine residues, the lowest excitation energy becomes very small
 - RH/DIIS converges to an excited state (first-order saddle point)



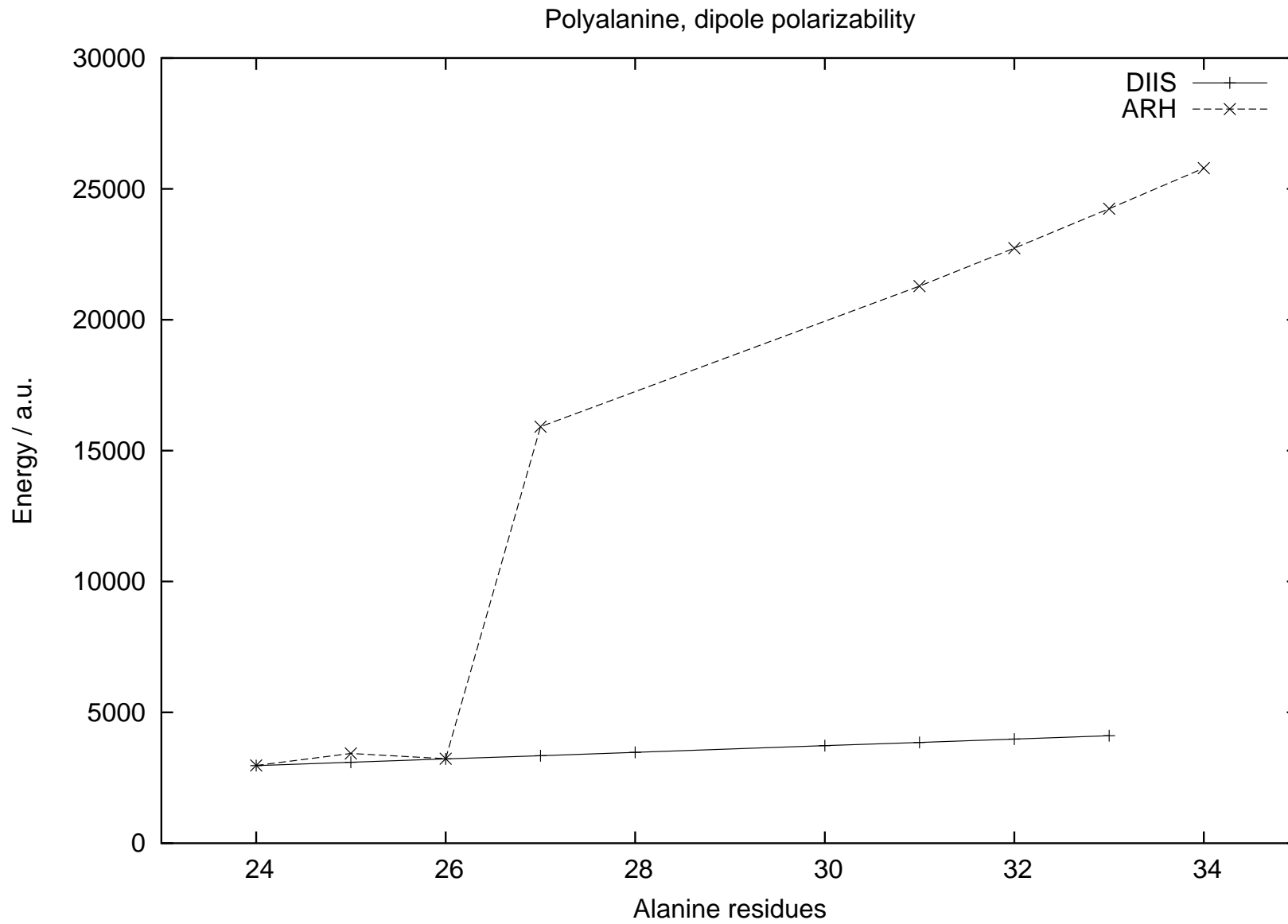
- for LDA and GGA functionals, the crossing occurs even earlier
 - local DFT functionals unable to account for long-range exchange in large molecules

B3LYP/6-31G polyaniline HOMO–LUMO gap and lowest Hessian eigenvalue

- The B3LYP energy minimum behaves in a strange manner
 - the lowest Hessian eigenvalue goes through a minimum for 27 residues
 - the RH/DIIS solution behaves smoothly with increasing chain length



B3LYP/6-31G polyaniline polarizabilities

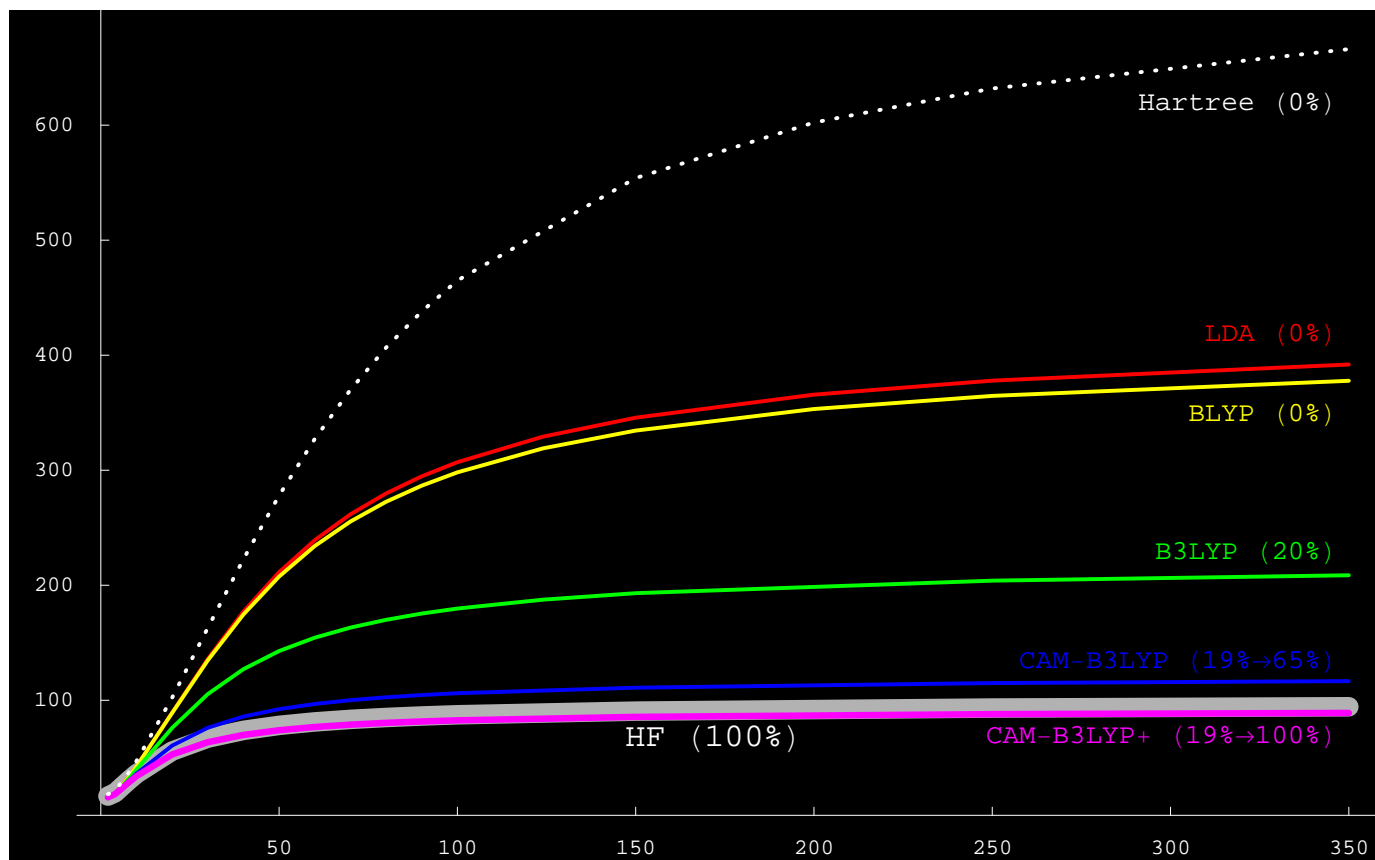


Summary

- We have considered some computational tasks in SCF theory
 - orthogonalization of the atomic orbitals
 - minimization of the SCF energy in an orthonormal basis
- Several approaches may be taken to the SCF energy minimization
 - full second-order Newton optimization
 - Roothaan–Hall (RH) Newton minimization (equivalent to diagonalization)
 - augmented RH (ARH) (reuses old Fock matrices)
- The ARH method is superior to RH/DIIS
 - requires one Fock/KS-matrix evaluation at each iteration
 - reuses efficiently information present in old Fock/KS matrices
 - converges by design to the ground state, smoothly and robustly
- For sufficiently sparse AO matrices, all tasks can be carried out in linear time
 - the Fock/KS matrix is reevaluated no more often than in MO theory
 - the remaining work is rich in matrix multiplication
 - well suited to parallelization

Static polarizabilities

- To illustrate, we have calculated longitudinal polarizabilities of linear alkene chains
 - HF and DFT α/N in 6-31G basis, plotted against the number of carbons N



- poor long-range exchange leads to overpolarization

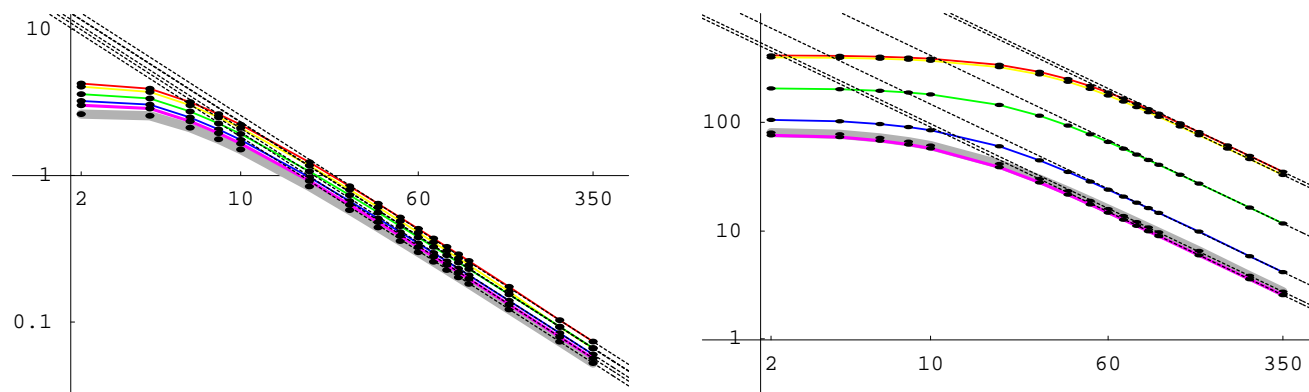
Asymptotic behaviour of group polarizabilities

- How does the group polarizability converge towards the infinite limit?

$$\bar{\alpha}_{\infty} - \bar{\alpha}_N = eN^{-1} + \mathcal{O}(N^{-2}) \quad \text{Kudin *et al.*, JCP **122**, 134907 (2005)}$$

– this behaviour is universal, holding at all levels of theory

- Log–log plots of $\bar{\alpha}_{\infty} - \bar{\alpha}_N$ for alkanes and alkenes:



– limit obtained by extrapolation $\bar{\alpha}_{\infty} = (\alpha_N - \alpha_M)/(N - M)$

– straight lines of slope -1 superimposed through the points at $N = 350$

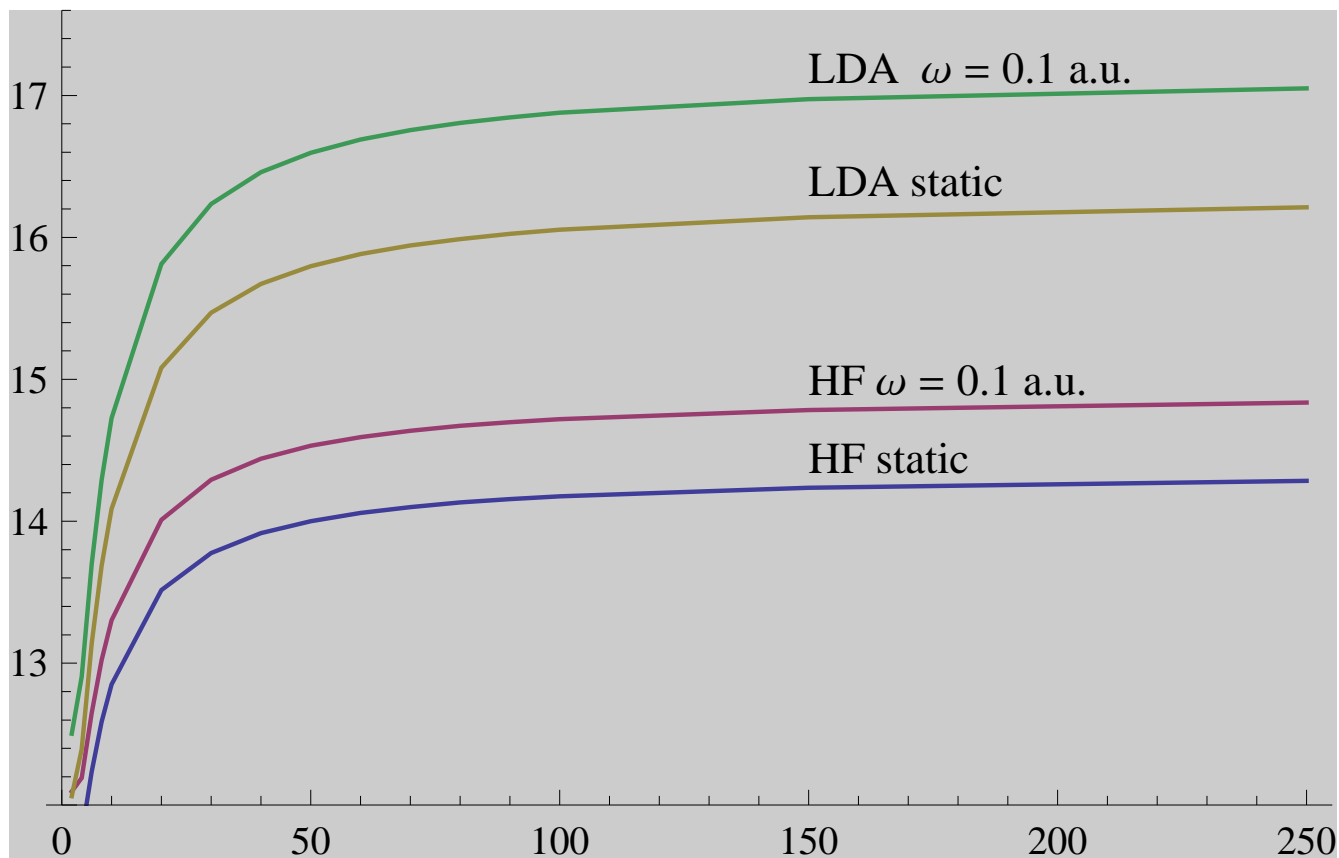
- The asymptotic region is reached with $\text{C}_{30}\text{H}_{62}$ (alkanes) and $\text{C}_{60}\text{H}_{62}$ (alkenes)

– alkane $\bar{\alpha}_{\infty}$ predicted to within 1% from $\text{C}_{30}\text{H}_{62}$

– alkene $\bar{\alpha}_{\infty}$ predicted to within 1% from $\text{C}_{60}\text{H}_{62}$ for HF and from $\text{C}_{150}\text{H}_{152}$ for LDA

Frequency-dependent polarizabilities

- HF and LDA longitudinal polarizabilities in linear alkane chains
 - α/N in 6-31G basis, plotted against the number of carbons N



- long-range exchange is less important for the saturated alkanes
- LDA overestimates dispersion as well as the static limit

Summary

- We have considered several computational tasks in self-consistent field theory
 - orthogonalization of the atomic orbitals
 - solution of the Roothaan–Hall eigenvalue problem
 - linear response theory
- For sufficiently sparse AO matrices, all these tasks can be carried out in linear time
 - the Fock/KS matrix is reevaluated no more often than in MO theory
 - the remaining work is rich in matrix multiplication
 - well suited to parallelization
- We have not considered the global convergence of the SCF method
 - standard methods may not converge to the ground state
 - better global optimization schemes must be developed
 - standard functionals often perform poorly for large systems
- Standard functionals may not be suitable for large systems
 - local functionals do not recover long-range correlation and exchange effects
 - the inclusion of exact exchange important