Development and Implementation of Algorithms for LEDO-DFT

Andreas W. Götz

Lehrstuhl für Theoretische Chemie
Universität Erlangen

Seminar zur Theoretischen Chemie
February 14th, 2003
Goal of LEDO-DFT

A simplified DFT formalism in the framework of Kohn-Sham DFT exhibiting favourable scaling behaviour with respect to the number $N$ of AO basis functions.
Introduction

   Essentials of DFT
   Density Fitting (RI-J)

LEDO-DFT

   The LEDO Expansion
   The Expansion Basis
   The LEDO-DFT Formalism
   Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results

   Small molecules
   Larger Molecules
   Timings
Introduction
Essentials of DFT
Density Fitting (RI-J)

LEDO-DFT
The LEDO Expansion
The Expansion Basis
The LEDO-DFT Formalism
Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results
Small molecules
Larger Molecules
Timings
Density Functional Theory

Energy expression:

\[ E[\rho] = T_s[\rho] + E_{\text{ext}}[\rho] + E_h[\rho] + E_{\text{xc}}[\rho] \]

\[ = T_s[\rho] + \int V_{\text{ext}}(r) \rho(r) \, dr \]

\[ + \frac{1}{2} \int \rho(r_1) \rho(r_2) r_{12}^{-1} \, dr_1 \, dr_2 + \int \varepsilon_{\text{xc}}[\rho(r)] \rho(r) \, dr \]

- Kohn-Sham DFT (MO theory)

\[ \rho(r) = 2 \sum_{i}^{\text{occ}} |\psi_i(r)|^2 \]

- LCAO ansatz (introduce AO basis set, basis functions \( \{\phi_\mu\} \))

\[ \rho(r) = \sum_{\mu \nu} P_{\nu \mu} \phi_\mu(r) \phi_\nu(r) \]
Kohn-Sham DFT with LCAO Ansatz

Variational problem:
\[ \delta E = 0 \]

⇒ Secular matrix:
\[ F^{KS}_{\mu\nu} = T_{\mu\nu} + V^{\text{ext}}_{\mu\nu} + V^{h}_{\mu\nu} + V^{xc}_{\mu\nu} \]

<table>
<thead>
<tr>
<th>Scaling behaviour</th>
<th>formal</th>
<th>asymptotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{\mu\nu} ) = ( \langle \phi_\mu \mid - \frac{1}{2} \Delta \mid \phi_\nu \rangle )</td>
<td>( O(N^2) )</td>
<td>( O(N) )</td>
</tr>
<tr>
<td>( V^{\text{ext}}<em>{\mu\nu} ) = ( \langle \phi</em>\mu \mid V_{\text{ext}} \mid \phi_\nu \rangle )</td>
<td>( O(N^2) )</td>
<td>( O(N) )</td>
</tr>
<tr>
<td>( V^{h}<em>{\mu\nu} ) = ( \langle \phi</em>\mu \mid V_{h} \mid \phi_\nu \rangle = \sum_{\kappa\lambda} P_{\kappa\lambda} (\phi_\mu \phi_\nu \mid \phi_\kappa \phi_\lambda) )</td>
<td>( O(N^4) )</td>
<td>( O(N^2) )</td>
</tr>
<tr>
<td>( V^{xc}<em>{\mu\nu} ) = ( \langle \phi</em>\mu \mid V_{xc} \mid \phi_\nu \rangle )</td>
<td>( O(N^3) )</td>
<td>( O(N) )</td>
</tr>
</tbody>
</table>

**Bottleneck:** Evaluation of four-center ERIs (electron repulsion integrals)
Density Fitting / RI-J Approximation

Expand the electron density into atom-centered auxiliary basis \( \{ \Omega_\alpha \} \):

\[
\rho(r) \approx \tilde{\rho}(r) = \sum_\alpha c_\alpha \Omega_\alpha(r)
\]

- Density fitting in each SCF cycle necessary
- Approximate Hartree potential \( \tilde{V}_h(r_1) = \int \tilde{\rho}(r_2) r_{12}^{-1} \, dr_2 \)
- Only three-center ERIs have to be evaluated

<table>
<thead>
<tr>
<th>Scaling behaviour</th>
<th>formal</th>
<th>asymptotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{V}^h_{\mu\nu} = \langle \phi_\mu</td>
<td>\tilde{V}_h</td>
<td>\phi_\nu \rangle = \sum_\alpha c_\alpha (\phi_\mu \phi_\nu</td>
</tr>
<tr>
<td>( V^{xc}<em>{\mu\nu} = \langle \phi</em>\mu</td>
<td>V_{xc}</td>
<td>\phi_\nu \rangle )</td>
</tr>
</tbody>
</table>

**Bottleneck**: Evaluation of exchange correlation contribution
Introduction
   Essentials of DFT
   Density Fitting (RI-J)

LEDO-DFT
   The LEDO Expansion
   The Expansion Basis
   The LEDO-DFT Formalism
   Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results
   Small molecules
   Larger Molecules
   Timings
The LEDO Expansion

LEDO: **Limited Expansion of Diatomic Overlap**

Expand diatomic overlap densities \( \phi_{\mu A} \phi_{\nu B} \) into expansion functions \( \{ \Omega_{p A}, \Omega_{q B} \} \) located on atoms \( A \) and \( B \):

\[
\phi_{\mu A}(r)\phi_{\nu B}(r) \approx \sum_{p A} f_{p A,\mu A \nu B} \Omega_{p A}(r) + \sum_{q B} f_{q B,\mu A \nu B} \Omega_{q B}(r)
\]

\[
\rho_k \approx \tilde{\rho}_k = \sum_{p} \Omega_{p} f_{pk} = \Omega^\dagger f_k
\]

LEDO expansion coefficients depend on the

- molecular structure
- AO basis set \( \{ \phi_\mu \} \)
- LEDO expansion basis set \( \{ \Omega_p \} \)

\( \Rightarrow \) Fitting of overlap densities only once before start of SCF

---

\(^1\) Billingsley, Bloor *JCP* (55) 1971
Determination of the Expansion Coefficients

- Minimize the deviation between $\rho_k$ and $\tilde{\rho}_k$
- Employ the Coulomb norm:

$$\Delta_k = (\rho_k - \tilde{\rho}_k | \rho_k - \tilde{\rho}_k)$$

Minimize wrt. the expansion coefficients $f_k$:

$$\nabla f_k \Delta_k = 0$$

$\Rightarrow$ System of linear equations:

$$\sum_q (\Omega_p | \Omega_q) f_{qk} = (\Omega_p | \rho_k)$$

- Matrix notation: $W_{pq} = (\Omega_p | \Omega_q)$, $F_{qk} = f_{qk}$, $Y_{pk} = (\Omega_p | \rho_k)$

$$WF = Y$$
The Problem of Near Linear Dependences

$W$: positive definite for linear independent expansion functions $\Omega_p$

- Method of choice for solution: Cholesky factorization
- Problem: Numerical noise due to roundoff errors
  $\Rightarrow$ Near linear dependences lead to singular $W$

Solutions to the problem:

a) Perturbative approximation$^2$
   - Shift diagonal elements of $W$: multiply with $(1 + \delta)$
   - Standard linear algebra packages can be used

b) Modified Cholesky factorization$^3$
   - Eliminate linear dependent vectors from $W$
   - Sorting algorithm computational expensive, no standard linear algebra package can be used

---

$^2$Eichkorn, Ahlrichs *CPL* (242) 1995
$^3$Beebe, Linderberg *IJQC* (12) 1977
The Problem of Near Linear Dependences

Solutions to the problem:

c) *a priori* elimination

- Determine near linear dependent expansion functions \( \Omega_{pA} \) for an atom \( A \) *a priori* with modified Cholesky decomposition
- In practice, no additional linear dependences are encountered for atom pairs (with non-zero two-center block of \( \mathbf{W} \))
- Reduced dimensions (computational cost!)
- Standard linear algebra packages can be used
The LEDO Expansion Basis

The LEDO expansion basis consists of all one-center overlap densities:

$$\{\Omega_{\mu A}\} = \{\phi_{\mu A} \phi_{\mu A}'\}$$

Not sufficient for an accurate fit of all diatomic overlap densities $\rho_{\mu A \nu B}$

Solution: Employ additional auxiliary functions

a) Auxiliary orbitals $\chi_{\mu A}$
   - Lead to additional orbital products $\chi_{\mu A}' \phi_{\mu A}$ and $\chi_{\mu A}' \chi_{\mu A}$
   - Straightforward implementation
   - Increase of near linear dependent expansion functions in a rather uncontrollable fashion

b) Auxiliary functions $\Lambda_{r A}$
   - Better control of the quality of the LEDO expansion basis
   - Near linear dependences avoided from the outset
   - More difficult to determine than auxiliary orbitals
Optimization of the LEDO Expansion Basis

\[(\tilde{\rho}_k | \tilde{\rho}_k) = (\rho_k | \tilde{\rho}_k)\]

⇒ The Coulomb norm can be expressed as:

\[\Delta_k = (\rho_k | \rho_k) - (\tilde{\rho}_k | \rho_k) = (\rho_k | \rho_k) - \sum_p f_{pk}(\Omega_p | \rho_k)\]

▶ \(\Delta_k\) is always positive ⇒ approximate are smaller than exact integrals
▶ \(\Delta_k = 0 \iff \tilde{\rho}_k(r) = \rho_k(r)\)
▶ Use \(\Delta_k\) as a measure for the quality of the LEDO fit

Define the norm

\[\Delta_{l_A m_B} = \sum_{i_{l_A} j_{m_B}} \Delta_{i_{l_A} j_{m_B}}\]

▶ \(i_{l_A}, j_{m_B}\) are basis functions \(\phi_{\mu_A}, \phi_{\nu_B}\) belonging to shells \(l_A, m_B\)
▶ \(\Delta_{l_A m_B}\) is rotational invariant
Optimization of the LEDO Expansion Basis

Define the norms

\[ \Delta l_A m_B = \sum_{l_A} \sum_{m_B} \Delta l'_{A} m'_{B}, \quad \Delta^{AB} = \sum_{l_A} \sum_{m_B} \Delta l_A m_B \]

- Optimize the exponents \( \alpha^p_i \) and contraction coefficients \( c^p_i \) of the expansion functions \( \Omega_p \) according to

\[ \frac{\partial \Delta^{AB}}{\partial \alpha^p_i} = 0 \]

and

\[ \frac{\partial \Delta^{AB}}{\partial c^p_i} = 0 \]

- The derivatives can be easily obtained numerically
The LEDO-DFT Formalism

Use the approximate diatomic overlap densities \( \tilde{\rho}_{\mu A \nu B} \) in the expression for the electron density:

\[
\rho(\mathbf{r}) = \sum_{\mu \nu} P_{\nu \mu} \phi_{\mu}(\mathbf{r}) \phi_{\nu}(\mathbf{r}) \approx \tilde{\rho}(\mathbf{r}) = \sum_{A, p_A} N_{p_A} \Omega_{p_A}(\mathbf{r})
\]

- The expression for the approximate electron density \( \tilde{\rho} \) consists only of one-center terms
- The vector \( \mathbf{N} \) contains
  a) the one-center elements of the density matrix \( \mathbf{P} \)
  b) products of the two-center elements of the density matrix \( \mathbf{P} \) with the LEDO expansion coefficients

\[ ^4 \text{Kollmar, Hess Mol. Phys. (100) 2002} \]
The LEDO-DFT Formalism

Replace $\rho$ by $\tilde{\rho}$ in all energy terms of the Kohn-Sham formalism which are *explicit* functionals of the electron density (or its gradient)

$$\rho(r) \rightarrow \tilde{\rho}(r) \quad \Rightarrow \quad E \rightarrow \tilde{E}$$

Variational problem:

$$\delta \tilde{E} = 0$$

$\Rightarrow$ Secular matrix:

$$\tilde{F}_{\mu\nu}^{KS} = T_{\mu\nu} + \tilde{V}_{\mu\nu}^{ext} + \tilde{V}_{\mu\nu}^{h} + \tilde{V}_{\mu\nu}^{xc}$$

- Approximate Hartree potential

$$\tilde{V}_{h}(r_1) = \int \tilde{\rho}(r_2)r_1^{-1} dr_2$$

- Approximate exchange-correlation potential

$$\tilde{V}_{xc}(r) = \frac{\delta E_{xc}[\tilde{\rho}(r)]}{\delta \tilde{\rho}(r)}$$
The LEDO-DFT Formalism

- Only one-center matrix elements of the potential contributions \( \tilde{V}^{\text{ext}}_{pA} \), \( \tilde{V}^{h}_{pA} \) and \( \tilde{V}^{\text{xc}}_{pA} \) have to be evaluated explicitly.
- The two-center matrix elements \( \tilde{V}^{\text{ext}}_{\mu A \nu B} \), \( \tilde{V}^{h}_{\mu A \nu B} \) and \( \tilde{V}^{\text{xc}}_{\mu A \nu B} \) are obtained from the LEDO expansion coefficients.

### Scaling behaviour

<table>
<thead>
<tr>
<th></th>
<th>formal</th>
<th>asymptotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{V}^{\text{ext}}<em>{pA} ) ( \equiv \int V</em>{\text{ext}}(r) \Omega_{pA}(r) , dr )</td>
<td>( \mathcal{O}(N) )</td>
<td>( \mathcal{O}(N) )</td>
</tr>
<tr>
<td>( \tilde{V}^{h}<em>{pA} ) ( \equiv \int \tilde{V}</em>{h}(r) \Omega_{pA}(r) , dr = \sum_{B,qB} N_{qB}(\Omega_{pA}</td>
<td>\Omega_{qB}) )</td>
<td>( \mathcal{O}(N^2) )</td>
</tr>
<tr>
<td>( \tilde{V}^{\text{xc}}<em>{pA} ) ( \equiv \int \tilde{V}</em>{\text{xc}}(r) \Omega_{pA}(r) , dr )</td>
<td>( \mathcal{O}(N^2) )</td>
<td>( \mathcal{O}(N) )</td>
</tr>
<tr>
<td>( \tilde{V}<em>{\mu A \nu B} \equiv \sum</em>{A,pA} f_{pA,\mu A \nu B} \tilde{V}<em>{pA} + \sum</em>{B,qB} f_{qB,\mu A \nu B} \tilde{V}_{qB} )</td>
<td>( \mathcal{O}(N^2) )</td>
<td>( \mathcal{O}(N) )</td>
</tr>
</tbody>
</table>

- In practice no approximation of \( V^{\text{ext}} \).
Analytical gradients for LEDO-DFT

Derivative of the approximated electronic energy $\tilde{E}$ wrt. a perturbation $\alpha$:

$$
\frac{\partial \tilde{E}}{\partial \alpha} = \sum_{A,B,\mu_A,\nu_B} P_{\nu_B \mu_A} \langle \phi_{\mu_A} | - \frac{1}{2} \Delta | \phi_{\nu_B} \rangle^\alpha 
+ \sum_{A,p_A} N_{p_A} \left\{ \left( \int V_{\text{ext}}(r) \Omega_{p_A}(r) \, d^3r \right)^{\alpha} + \left( \int \tilde{V}_{\text{xc}}(r) \Omega_{p_A}(r) \, d^3r \right)^{(\alpha)} \right\}
+ \frac{1}{2} \sum_{A,B,p_A,q_B} N_{p_A} N_{q_B} (\Omega_{p_A} | \Omega_{q_B})^\alpha 
+ \sum_{A,p_A} N_{p_A} \tilde{V}_{p_A} 
+ 4 \sum_{i} \varepsilon_i G_{ii}^{\alpha} 
= T^\alpha + \tilde{V}_{\text{ext}}^\alpha + \tilde{V}_{\text{xc}}^\alpha + \tilde{V}_{h}^\alpha + V_{\text{LEDO}}^\alpha + G^\alpha
$$

Formal scaling behaviour: $O(N^2)$
Determination of Derivatives of the Expansion Coefficients

- The LEDO expansion coefficients $f_{pk}$ depend on the structure of the molecule

- $N^\alpha$ contains derivatives $f_{pk}^\alpha$ of the LEDO expansion coefficients

⇒ New system of linear equations:

$$\sum_q (\Omega_p | \Omega_q) f_{qk}^\alpha = (\Omega_p | \rho_k)^\alpha - \sum_q (\Omega_p | \Omega_q)^\alpha f_{qk}$$

$$WF^\alpha = Y^\alpha - W^\alpha F$$

This system of linear equations

- differs only in the inhomogenity from the one for the determination of the LEDO expansion coefficients $f_{pk}$

- contains derivatives of two-center ERIs

- contains the LEDO expansion coefficients $f_{pk}$
Introduction

Essentials of DFT
Density Fitting (RI-J)

LEDO-DFT

The LEDO Expansion
The Expansion Basis
The LEDO-DFT Formalism
Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results

Small molecules
Larger Molecules
Timings
Framework of the Implementation

- Program package TURBOMOLE 5.1
  - SCF: program RIDFT
  - Gradient: program RDGRAD

- LEDO approximation for the Coulomb and XC operators
- Use of auxiliary orbitals $\chi_{\mu A}$, no auxiliary functions $\Lambda_{rA}$
- No molecular point group symmetry
- Prescreening of overlap densities and ERIs
- Translational invariance of the ERIs fully exploited
- In-core algorithm

- Programming language: FORTRAN 77
- Memory management: dynamical
  (FORTRAN 90 ALLOCATE and MEMMGR by B.A. Hess)
- Extensive use of LAPACK and BLAS linear algebra routines
- Communication between RIDFT and RDGRAD via files
Flow Chart of the SCF Program

- calculate $T$ and $V_{\text{ext}}$
- calculate $W$ and $Y$
- loop over atom pairs $A > B$
  - solve $W_{AB}F_{AB} = Y_{AB}$
  - form vector $N$
  - form vector $N$
  - form $\tilde{V}^h_A = \sum_B W_{AB}N_B$
  - calculate $\tilde{V}_{xc}^A$ from $\tilde{\rho} = \Omega^\dagger N$
  - form $\tilde{V}^h,xc_{AB} = \tilde{V}^h,xc_A F_{AB} + \tilde{V}^h,xc_B F_{AB}$
- convergence?
- write $W_{AB}^{-1}$ and $F_{AB}$ to disk
- write vector $N$ and $\tilde{V}^h,xc_A$ to disk
- solve secular equations
- guess for initial $P$
- get new $P$
Flow Chart of the SCF Program

- calculate $T$ and $V_{\text{ext}}$
- calculate $W$ and $Y$
- loop over atom pairs $A > B$
  - solve $W_{AB}F_{AB} = Y_{AB}$
  - write $W_{AB}^{-1}$ and $F_{AB}$ to disk
- form vector $N$
  - form $\tilde{V}^h_A = \sum_B W_{AB}N_B$
  - calculate $\tilde{V}^{xc}_A$ from $\tilde{\rho} = \Omega^iN$
  - form $\tilde{V}_{AB}^{h,xc} = \tilde{V}_A^{h,xc}F_{AB} + \tilde{V}_B^{h,xc}F_{AB}$
- solve secular equations
- convergence?
  - evaluate norm $\Delta_k, \Delta_{l_{AmB}}$
  - calculate integrals $(\rho_k|\rho_k)$
- guess for initial $P$
- get new $P$
- write vector $N$ and $\tilde{V}_A^{h,xc}$ to disk

Andreas Götz, Universität Erlangen
Development and Implementation of Algorithms for LEDO-DFT
Flow Chart of the Gradient Program

- calculate $T^\alpha$, $V^\alpha_{ext}$, $G^\alpha$
- read vector $N$ from disk
- loop over atom pairs $A > B$
  - calculate $W^\alpha_{AB}$
  - form contribution to gradient: $\tilde{V}^\alpha_{h,AB} = N_A^\dagger W^\alpha_{AB} N_B$
- read LEDO expansion coefficients $F_{AB}$ from disk
- form contribution to inhomogeneity: $-W^\alpha_{AB}F_{AB}$
- calculate $Y^\alpha_{AB}$ and form inhomogeneity: $Y^\alpha_{AB} - W^\alpha_{AB}F_{AB}$
- read factorized coefficient matrix $W^{-1}_{AB}$ from disk
- solve for derivatives of LEDO expansion coefficients: $F^\alpha_{AB} = W^{-1}_{AB}(Y^\alpha_{AB} - W^\alpha_{AB}F_{AB})$
- form contribution of atom pair $A > B$ to $N^\alpha$
- read $\tilde{V}^h_{A,xc}$, form $V^\alpha_{LEDO} = N^\alpha \dagger \tilde{V}^h_{A,xc}$
- calculate $\tilde{V}^\alpha_{xc}$ from $\tilde{\rho} = \Omega^\dagger N$
Introduction
   Essentials of DFT
   Density Fitting (RI-J)

LEDO-DFT
   The LEDO Expansion
   The Expansion Basis
   The LEDO-DFT Formalism
   Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results
   Small molecules
   Larger Molecules
   Timings
Auxiliary Orbitals: General observations

A good LEDO expansion basis should minimize the norm $\Delta_{l_A m_B}$ for all combinations of shell pairs that can appear in molecular calculations.

Uncritical are shell pairs involving

- core shells
- valence shells with angular momentum $l$ if the basis set contains another shell with
  1. $l' > l$
  2. similar radial decay properties

Most critical are shell pairs involving

- shells with high $l$

⇒ solution: include auxiliary orbitals

- 1 uncontracted Gaussian for each critical shell with
  1. $l' = l + 1$
  2. similar radial decay properties
SVP basis set for carbon

split valence + polarization function: $[3s2p1d]$

<table>
<thead>
<tr>
<th>Shell Type</th>
<th>Exponents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core shell</td>
<td>1s</td>
</tr>
<tr>
<td>Valence shells</td>
<td>2s (compact), 3s (diffuse), 4p (compact)</td>
</tr>
<tr>
<td>Polarization function</td>
<td>6d (compact)</td>
</tr>
</tbody>
</table>

⇒ need two uncontracted Gaussians as auxiliary orbitals:

1. $d$ (diffuse) for the $5p$ shell
2. $f$ (compact) for the $6d$ shell

Homonuclear atom pair, two-step procedure: optimize exponents $\zeta$ of

1. $d$ type auxiliary orbital, use the norm $\Delta^{5p5p}$
2. $f$ type auxiliary orbital, use the norm $\Delta^{6d6d} = \Delta^{CC}$
Homonuclear diatomic charge distributions

CC, interatomic distance = 120 pm (≡ triple bond)

\[
\begin{align*}
\Delta^{5p5p} : & \quad 7.9 \times 10^{-4} \longrightarrow 1.5 \times 10^{-5} \\
\text{with } & \quad \zeta(d) = 0.16 \ (\zeta(5p) \approx 0.15) \\
\Delta^{CC} : & \quad 1.6 \times 10^{-2} \longrightarrow 1.6 \times 10^{-4} \\
\text{with } & \quad \zeta(f) = 0.85 \ (\zeta(6d) = 0.8)
\end{align*}
\]

⇒ improvement of the fit: two orders of magnitude
Homonuclear diatomic charge distributions

CC, interatomic distance = 120 pm

\[ 2.0 \times 10^{-7} \leq \Delta^{5p5p} \leq 7.5 \times 10^{-6} \]

\[ 1.5 \times 10^{-4} \leq \Delta^{CC} \leq 4.1 \times 10^{-4} \]

- \( \Delta^{5p5p} \) depends mainly on \( \zeta(d) \)
- \( \Delta^{CC} \) depends mainly on \( \zeta(f) \)

- flat surface in proximity of minima

\( \Rightarrow \) independent, consecutive optimization of \( \zeta(d) \) and \( \zeta(f) \) justified
Distance dependence

Andreas Götz, Universität Erlangen
Development and Implementation of Algorithms for LEDO-DFT
Heteronuclear diatomic charge distributions

CN, interatomic distance = 122 pm (\(\equiv\) double bond)

\[2.3 \times 10^{-5} \leq \Delta^{5p5p} \leq 5.0 \times 10^{-5}\]

\[2.0 \times 10^{-4} \leq \Delta^{CN} \leq 5.8 \times 10^{-4}\]

- flat surface in proximity of minima
- no strong atom-pair dependence of best exponents

\(\Rightarrow\) use auxiliary orbitals optimized for homonuclear atom pairs
Auxiliary orbitals for H, Li–F, Na–Cl for SVP basis set

- Similar behaviour of distance dependence for all elements
- Changes are regular across the periodic table
- Reduction of the norm $\Delta^{AA}$ of at least two orders of magnitude

![Graph showing the number of expansion functions for different elements. Red and green dots represent full expansion basis and a priori elimination, respectively.](image)

- a-priori elimination reduces the size of the expansion basis by 23–57%
  ⇒ reduces the computational effort for the factorization of $W$ by 55–92%
Introduction
Essentials of DFT
Density Fitting (RI-J)

LEDO-DFT
The LEDO Expansion
The Expansion Basis
The LEDO-DFT Formalism
Analytical Gradients for LEDO-DFT

Implementation

Auxiliary Orbitals

Numerical Results
Small molecules
Larger Molecules
Timings
Small Molecules

- All calculations have been performed with the BP86 functional
- Experimental or RI-DFT/SV(P) optimized structures as starting point for the optimizations
- Force constants have been obtained as numerical derivatives of the analytical gradients with SNF
- Test set of 142 small molecules
  - 315 bond distances
  - 282 bond angles
  - 85 dipole moments
  - 1446 harmonic frequencies
### Small Molecules

<table>
<thead>
<tr>
<th>Molecule</th>
<th>$\Delta E/10^{-3} \text{au}$</th>
<th>$\Delta d/\text{pm}$</th>
<th>$\Delta \gamma/^{\circ}$</th>
<th>$\Delta \mu/10^{-3} \text{au}$</th>
<th>$\Delta \nu/\text{cm}^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$O</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.1</td>
<td>-0.8</td>
<td>9</td>
</tr>
<tr>
<td>N$_2$H$_4$</td>
<td>0.01</td>
<td>0.05</td>
<td>0.1</td>
<td>-3.0</td>
<td>7</td>
</tr>
<tr>
<td>C$_3$H$_6$</td>
<td>0.11</td>
<td>0.08</td>
<td>0.1</td>
<td>-1.7</td>
<td>16</td>
</tr>
<tr>
<td>H$_3$CN$_2$CH$_3$</td>
<td>0.76</td>
<td>0.10</td>
<td>0.2</td>
<td>-0.1</td>
<td>8</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.67</td>
<td>0.12</td>
<td>0.2</td>
<td>5.4</td>
<td>6</td>
</tr>
<tr>
<td>Cysteine</td>
<td>1.08</td>
<td>0.22</td>
<td>0.3</td>
<td>-16.6</td>
<td>43</td>
</tr>
<tr>
<td>Li$_2$O</td>
<td>-0.77</td>
<td>0.06</td>
<td>1.9</td>
<td>-54.9</td>
<td>15</td>
</tr>
<tr>
<td>Si$_2$H$_6$</td>
<td>0.83</td>
<td>1.37</td>
<td>0.0</td>
<td>—</td>
<td>36</td>
</tr>
<tr>
<td>Si(CH$_3$)$_3$Cl</td>
<td>1.06</td>
<td>0.20</td>
<td>0.1</td>
<td>-20.3</td>
<td>43</td>
</tr>
<tr>
<td>P(CH$_3$)$_3$</td>
<td>1.90</td>
<td>0.13</td>
<td>0.9</td>
<td>-6.5</td>
<td>98</td>
</tr>
<tr>
<td>PO(CH$_3$)$_3$</td>
<td>1.07</td>
<td>0.50</td>
<td>0.3</td>
<td>-55.2</td>
<td>19</td>
</tr>
<tr>
<td>P$<em>4$O$</em>{10}$</td>
<td>11.04</td>
<td>0.14</td>
<td>0.4</td>
<td>—</td>
<td>7</td>
</tr>
</tbody>
</table>

...
Small Molecules

<table>
<thead>
<tr>
<th></th>
<th>$\Delta E/10^{-3}\text{au}$</th>
<th>$\Delta d/\text{pm}$</th>
<th>$\Delta \gamma/^{\circ}$</th>
<th>$\Delta \mu/10^{-3}\text{au}$</th>
<th>$\Delta \nu/\text{cm}^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>max</td>
<td>11.03</td>
<td>1.37</td>
<td>1.9</td>
<td>55.2</td>
<td>98</td>
</tr>
<tr>
<td>RMS</td>
<td>1.33</td>
<td>0.15</td>
<td>0.3</td>
<td>15.2</td>
<td>10</td>
</tr>
</tbody>
</table>

- Energy differences are of the order of $10^{-3}\text{au}$
- Errors in
  - bond distances are of the order of 0.1 pm
  - bond distances involving hydrogen never exceed 0.1 pm
  - bond angles are of the order of 0.1°
- Larger errors usually appear for weak bonds and soft angles
- Dipole moments and harmonic frequencies are in general in very good agreement

⇒ the performance of LEDO-DFT with the optimized auxiliary orbitals is satisfactory for the molecules under consideration
Larger Molecules — Linear Alkanes

- MM3 optimized structures as starting point for the optimizations

<table>
<thead>
<tr>
<th>Chain length</th>
<th>$\Delta E$</th>
<th>$\Delta d$/pm</th>
<th>$\Delta \gamma$/°</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.54</td>
<td>0.09</td>
<td>0.4</td>
</tr>
<tr>
<td>10</td>
<td>5.24</td>
<td>0.12</td>
<td>0.5</td>
</tr>
<tr>
<td>15</td>
<td>12.42</td>
<td>0.12</td>
<td>0.5</td>
</tr>
<tr>
<td>20</td>
<td>21.42</td>
<td>0.12</td>
<td>0.5</td>
</tr>
<tr>
<td>25</td>
<td>31.91</td>
<td>0.14</td>
<td>0.5</td>
</tr>
<tr>
<td>30</td>
<td>43.06</td>
<td>0.15</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- Errors in
  - total energies increase relatively fast
  - structure parameters increase slowly / remain almost stable with increasing system size
Larger Molecules — Projection technique

The numerical errors introduced by the LEDO approximation can lead to problems with the SCF convergence for larger molecules.

⇒ A projection technique\(^5\) can be successfully employed

- Eliminate canonically orthogonolized orbitals (COOs) belonging to near-linear dependent AOs
- Condition number \(c\) of the overlap matrix \(S\) as a measure for the near-linear dependence
- A value \(c^{\text{max}} = 4000\) is about right
- Electronic energy is raised, however effects are small since eliminated COOs contribute mainly to virtual space
- The use of this technique is inexpensive and requires only minor modifications to the program

Larger Molecules — Projection technique

<table>
<thead>
<tr>
<th></th>
<th>exact DFT</th>
<th></th>
<th>LEDO-DFT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>$\Delta E / 10^{-3} \text{ au}$</td>
<td>$\Delta d / \text{pm}$</td>
</tr>
<tr>
<td>$\text{C}_6\text{H}_2$</td>
<td>3</td>
<td>3.99</td>
<td>0.09</td>
</tr>
<tr>
<td>$\text{C}_8\text{H}_2$</td>
<td>5</td>
<td>6.65</td>
<td>0.14</td>
</tr>
<tr>
<td>$\text{C}_{20}\text{H}_2$</td>
<td>14</td>
<td>18.87</td>
<td>0.11</td>
</tr>
<tr>
<td>$\text{C}_{20}$ (ring)</td>
<td>13</td>
<td>10.65</td>
<td>0.16</td>
</tr>
<tr>
<td>$\text{C}_{20}$ (bowl)</td>
<td>14</td>
<td>4.19</td>
<td>0.12</td>
</tr>
<tr>
<td>$\text{C}_{20}$ (cage)</td>
<td>14</td>
<td>9.41</td>
<td>0.31</td>
</tr>
</tbody>
</table>

M: number of eliminated COOs

- Errors in bond distances due to the projection technique remain very small in all cases
- Errors due to the LEDO approximation are of the same order

⇒ Application of the projection technique is justified
Timings for Linear Alkanes

Total CPU time for the SCF

Intel XEON / 2.0GHz

- exact DFT (semi-direct)
- RI-DFT (incore)
- LEDO-DFT (incore)
- XC
- XC (LEDO)
- diagonalization
Timings for Linear Alkanes

Total CPU time for the analytical gradient

Intel XEON / 2.0GHz

- exact DFT
- RI-DFT
- LEDO-DFT
- XC
- XC (LEDO)

Chain length vs. t/min for different methods.
Conclusions and Outlook

We have shown that

- LEDO-DFT shows a formal $O(N^2)$ scaling behaviour
- near linear dependence of the LEDO expansion basis is not a problem
- the numerical results obtained with LEDO-DFT and the optimized auxiliary orbitals are in good agreement with unapproximated DFT
- a projection technique can be applied to guarantee SCF convergence
- the speed of both the SCF and the analytical gradient can compete with the RI-J approximation

It is highly desirable to have

- a semi-direct algorithm for very large systems
- an implementation based on auxiliary functions
  - better fits with less expansion functions are likely
  - near linear dependences avoided from the outset
  - good starting points for auxiliary bases are available
Acknowledgements

- Prof. Dr. Bernd A. Heß (†)
- Prof. Dr. Andreas Görling
- Dr. Christian Kollmar
- Dr. Nico van Eikema Hommes

Thank you for your attention.