Numerical Solution of the Hartree-Fock Equation by Multilevel Tensor-structured Methods

Venera Khoromskaia
TU Berlin, December 2010

Max-Planck-Institute for Mathematics in the Sciences, Leipzig

Outline of the talk

1. The Hartree-Fock equation
2. Tensor-structured (TS) methods in $\mathbb{R}^d$
   - Basic rank-structured formats
   - Best Tucker approximation (BTA)
   - Tensor product operations
3. Linear complexity algorithms: Multigrid accelerated (MGA) BTA
   - MGA BTA for canonical tensors
   - MGA BTA for full format tensors
4. Accurate TS computation of 3D and 6D integral operators
   - TS Computation of the Coulomb matrix
   - Computation of the exchange matrix
5. TS numerical solution of the Hartree-Fock equation
   - 3D nonlinear EVP solver by TS methods
   - Numerical examples for moderate size molecules

The Hartree-Fock equation

nonlinear eigenvalue problem in electronic structure calculations,

$$\mathcal{F}\varphi_i(x) = \lambda_i \varphi_i(x), \quad i = 1, ..., N_{orb}, \quad x \in \mathbb{R}^3$$

with the Fock operator

$$\mathcal{F} := -\frac{1}{2}\Delta - V_c + V_H - \mathcal{K},$$

the Hartree potential

$$V_H(x) := \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|}\, dy, \quad \rho(y) = 2\sum_{i=1}^{N_{orb}} (\varphi_i(y))^2$$

The exchange operator

$$(\mathcal{K}\varphi)(x) := \int_{\mathbb{R}^3} \frac{\tau(x,y)}{|x-y|}\, \varphi(y) dy, \quad \tau(x,y) = \sum_{i=1}^{N_{orb}} \varphi_i(x)\varphi_i(y).$$

Numerical challenges:
high accuracy
evaluation of 3D and 6D integrals,
strong nonlinearity.

Multidimensional tensors

A tensor of order d is a multidimensional array over a d-tuple index set $\mathcal{I} = I_1 \times \ldots \times I_d$,

$$A = [a_{i_1 \ldots i_d} : i_\ell \in I_\ell] \in \mathbb{R}^{\mathcal{I}}, \qquad I_\ell = \{1, ..., n_\ell\},\ \ell = 1, \ldots, d.$$

A tensor A is an element of the linear space

$$\mathbb{V}_{\mathbf{n}} = \otimes_{\ell=1}^{d} \mathbb{V}_\ell, \quad \mathbb{V}_\ell = \mathbb{R}^{I_\ell},$$

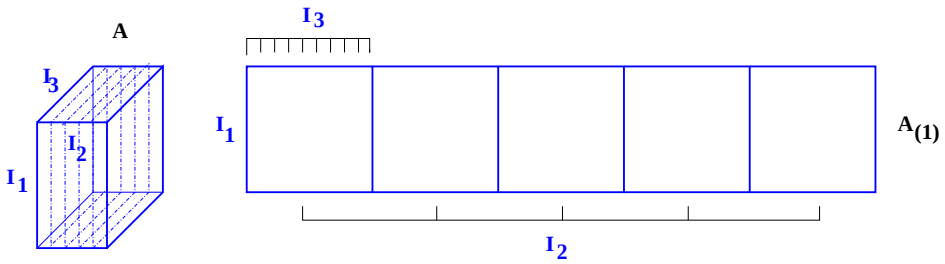
equipped with the Euclidean *scalar product* $\langle \cdot, \cdot \rangle : \mathbb{V}_{\mathbf{n}} \times \mathbb{V}_{\mathbf{n}} \to \mathbb{R}$, defined as

$$\langle A, B \rangle := \sum_{(i_1, \ldots, i_d) \in \mathcal{I}} a_{i_1 \ldots i_d} b_{i_1 \ldots i_d} \quad \text{for } A,\ B \in \mathbb{V}_{\mathbf{n}}.$$

Definition 1. The *unfolding of a tensor* along mode ℓ is a matrix

$$A_{(\ell)} = [a_{ij}] \in \mathbb{R}^{n_\ell \times (n_{\ell+1} \cdots n_d n_1 \cdots n_{\ell-1})}$$

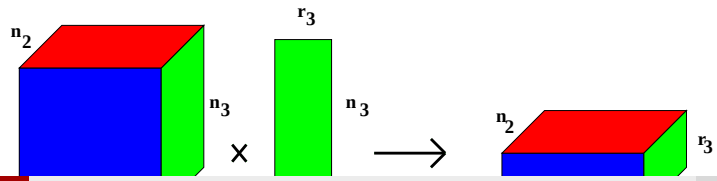
whose columns are the respectively ordered fibers along ℓ-mode.



Definition 2. Given a tensor $A \in \mathbb{R}^{I_1 \times \ldots I_d}$ and a matrix $M \in \mathbb{R}^{J_\ell \times I_\ell}$, we define the mode-$\ell$ tensor-matrix product by

$$B = A \times_\ell M \in \mathbb{R}^{I_1 \times \ldots \times I_{\ell-1} \times J_\ell \times I_{\ell+1} \ldots \times I_d},$$

where $b_{i_1,\ldots,i_{\ell-1},j_\ell,i_{\ell+1},\ldots,i_d} = \sum_{i_\ell=1}^{n_\ell} a_{i_1,\ldots,i_{\ell-1},i_\ell,i_{\ell+1},\ldots,i_d} m_{j_\ell,i_\ell}, \quad j_\ell \in J_\ell.$



Rank-structured separable representation of higher order tensors

$A = [a_{i_1 \ldots i_d} : i_\ell \in I_\ell] \in \mathbb{R}^{I_1 \times \ldots \times I_d}$, $I_\ell = \{1, ..., n_\ell\}$, $\ell = 1, \ldots, d$,
for $n_\ell = n$, storage $N = n^d$ –"curse of dimensionality".

Tensor product of vectors $u^{(\ell)} = \{u^{(\ell)}_{i_\ell}\}_{i_\ell=1}^{n} \in \mathbb{R}^{I_\ell}$ forms the canonical rank-1 tensor

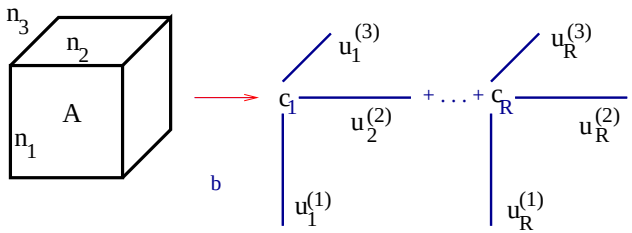
$$A_{(1)} \equiv [u_{\mathbf{i}}]_{\mathbf{i} \in \mathcal{I}} = u^{(1)} \otimes \ldots \otimes u^{(d)},$$

with entries $u_{\mathbf{i}} = u^{(1)}_{i_1} \cdots u^{(d)}_{i_d}$. Storage: $dn << n^d$.

Definition 3. The canonical format, $\mathcal{C}_{R,\mathbf{n}}$

$$A_{(R)} = \sum_{k=1}^{R} c_k u^{(1)}_k \otimes \ldots \otimes u^{(d)}_k, \quad c_k \in \mathbb{R}, \tag{1}$$

with normalised $u^{(\ell)}_k \in \mathbb{R}^n$ and the canonical rank $\leq R$.



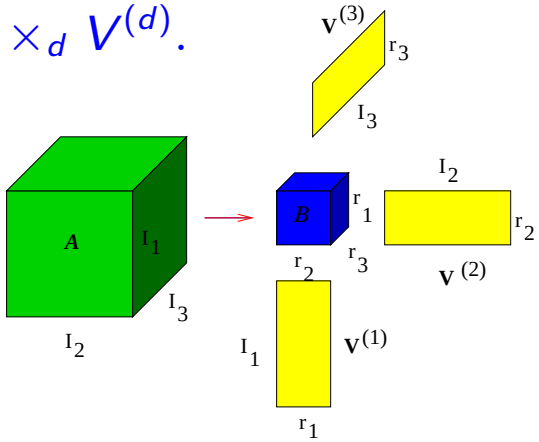
Tucker tensor format

Definition 4. Given the rank parameter $\mathbf{r} = (r_1, \ldots, r_d)$ we denote by $\mathcal{T}_\mathbf{r}$ the subset of tensors in $\mathbb{V}_\mathbf{n}$ represented in the Tucker format [Tucker, De Lathauwer].

$$A_{(\mathbf{r})} = \sum_{\nu_1=1}^{r_1} \cdots \sum_{\nu_d=1}^{r_d} \beta_{\nu_1 \ldots \nu_d} \, v_{\nu_1}^{(1)} \otimes \ldots \otimes v_{\nu_d}^{(d)}, \tag{2}$$

with orthonormal $v_{\nu_\ell}^{(\ell)} \in \mathbb{V}_\ell = \mathbb{R}^{I_\ell}$ $(1 \le \nu_\ell \le r_\ell)$, which form the basis of r_ℓ-dimensional subspaces $\mathbb{T}_\ell = \mathrm{span}\{v_\nu^{(\ell)}\}_{\nu=1}^{r_\ell}$ $(\ell = 1, ..., d)$.
The coefficients tensor $\boldsymbol{\beta} = [\beta_{\nu_1,\ldots,\nu_d}] \in \mathbb{B}_\mathbf{r} = \mathbb{R}^{r_1 \times \ldots \times r_d}$, is called the *core tensor*.

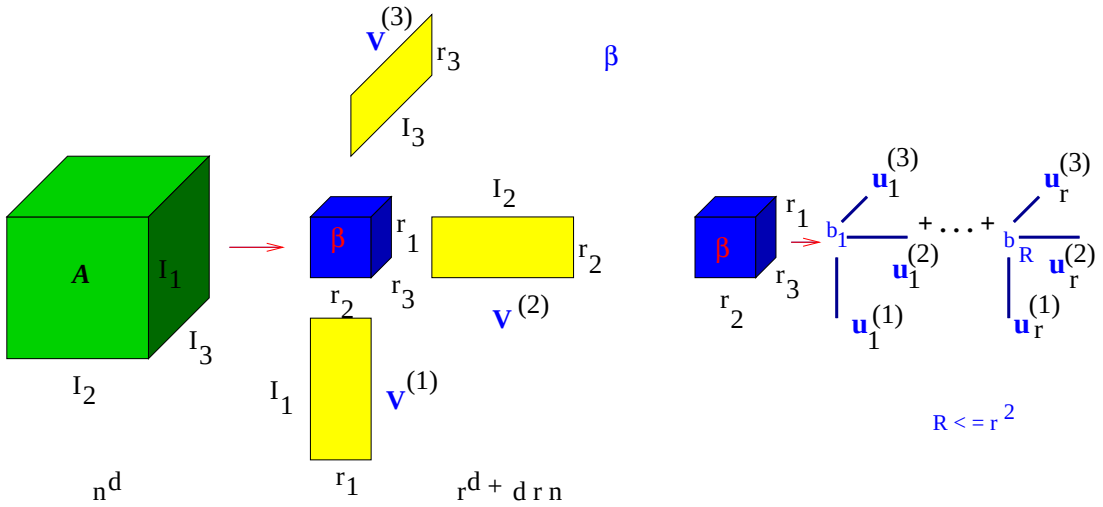
Using side matrices $V^{(\ell)} = [v_1^{(\ell)} \ldots v_{r_\ell}^{(\ell)}] \in \mathbb{R}^{n_\ell \times r_\ell}$ and tensor-by-matrix contracted product,
$A_{(\mathbf{r})} = \boldsymbol{\beta} \times_1 V^{(1)} \times_2 V^{(2)} \times_3 \ldots \times_d V^{(d)}$.



Storage: $r^d + drn \ll n^d$

Mixed Tucker-canonical (TC) tensor format

$$A_{(\mathbf{r})} = \left(\sum_{k=1}^{R} b_k u_k^{(1)} \otimes \ldots \otimes u_k^{(d)} \right) \times_1 V^{(1)} \times_2 V^{(2)} \times_3 \ldots \times_d V^{(d)}.$$



Advantages of the Tucker tensor decomposition:

1. Robust algorithm for approximating full format tensors of size n^d.
2. Rank reduction of the rank-R canonical tensors with large R.

3. Efficient for 3D tensors since r^3 is small.

Tucker tensor approximation

$$A_0 \in \mathbb{V}_{\mathbf{n}}: \quad f(A) := \|A - A_0\|^2 \to min \quad \text{over} \quad A \in \mathcal{T}_{\mathbf{r}}. \tag{3}$$

The minimisation problem (3) is equivalent to the maximisation problem

$$g(V^{(1)}, ..., V^{(d)}) := \left\| A_0 \times_1 V^{(1)\,T} \times ... \times_d V^{(d)\,T} \right\|^2 \to max \tag{4}$$

over the set of orthogonal matrices $V^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$.

For given $V^{(\ell)}$, the core $\boldsymbol{\beta}$ that minimises (3) is

$$\boldsymbol{\beta} = A_0 \times_1 V^{(1)\,T} \times ... \times_d V^{(d)\,T} \in \mathbb{R}^{r_1 \times ... \times r_d}. \tag{5}$$

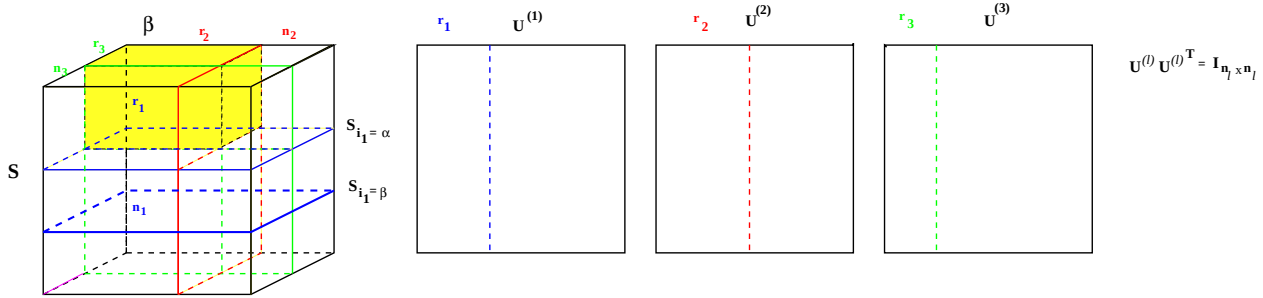
Algorithm BTA ($\mathbb{V}_{\mathbf{n}} \to \mathcal{T}_{\mathbf{r},\mathbf{n}}$), by ALS iteration.
Complexity for $d = 3$: $W_{F\to T} = O(n^4 + n^3 r + n^2 r^2 + n r^3)$.

Full format tensor to Tucker

Theorem 1. (dth order SVD, higher order SVD (HOSVD), [De Lathauwer et al, 2000]).
Every real $n_1 \times n_2 \times \ldots \times n_d$-tensor A can be written as the product

$$A = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \ldots \times_d U^{(d)},$$

1. $U^{(\ell)} = [U_1^{(\ell)} U_2^{(\ell)} \ldots U_{n_\ell}^{(\ell)}]$, $U^{(\ell)\,T} U^{(\ell)} = I_{n_\ell \times n_\ell}$,
2. Subtensors of $\mathcal{S}$, $\mathcal{S}_{i_\ell=\alpha}$, for fixed ℓth index, have the properties of
(i) all-orthogonality: $\langle \mathcal{S}_{i_\ell=\alpha}, \mathcal{S}_{i_\ell=\beta} \rangle = 0$ when $\alpha \neq \beta$,
(ii) ordering: $\|\mathcal{S}_{i_\ell=1}\| \geq \|\mathcal{S}_{i_\ell=2}\| \geq \ldots \geq \|\mathcal{S}_{i_\ell=n_\ell}\| \geq 0$ for all positive values of ℓ.
The Frobenius norms $\|\mathcal{S}_{i_\ell=i}\|$, $\sigma_i^{(\ell)}$, are ℓ-mode singular values of $A_{(\ell)}$ and the vector $U_i^{(\ell)}$ is an ith ℓ-mode left singular vector of $A_{(\ell)}$.



Truncated HOSVD

Theorem 2

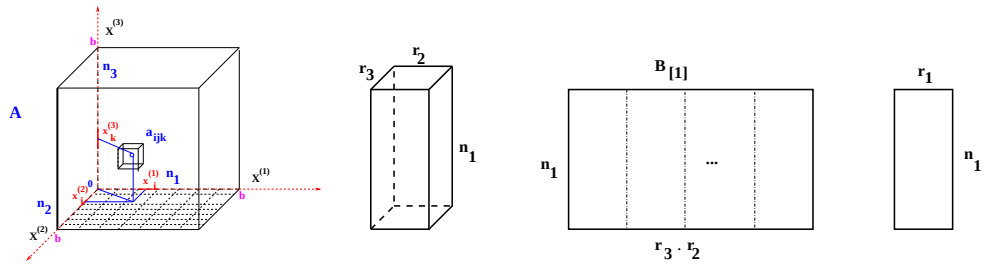
(Approximation by HOSVD, [De Lathauwer et al, 2000]).
Let the HOSVD of A be given as in Theorem 1 and let the ℓ-mode rank of A, $\mathrm{rank}(A_{(\ell)})$, be equal to R_ℓ $(\ell = 1, ..., d)$. Define a tensor $\tilde{A}$ by discarding the smallest ℓ-mode singular values $\sigma^{(\ell)}_{r_\ell+1}, \sigma^{(\ell)}_{r_\ell+2}, ..., \sigma^{(\ell)}_{R_\ell}$ for given values of r_ℓ $(\ell = 1, ..., d)$, i.e., set the corresponding parts of $\mathcal{S}$ equal to zero. Then we have

$$\|A - \tilde{A}\|^2 \leq \sum_{i_1=r_1+1}^{R_1} {\sigma^{(1)}_{i_1}}^2 + \sum_{i_2=r_2+1}^{R_2} {\sigma^{(2)}_{i_2}}^2 + \cdots + \sum_{i_d=r_d+1}^{R_d} {\sigma^{(d)}_{i_d}}^2.$$

Best Tucker approximation of function related tensors

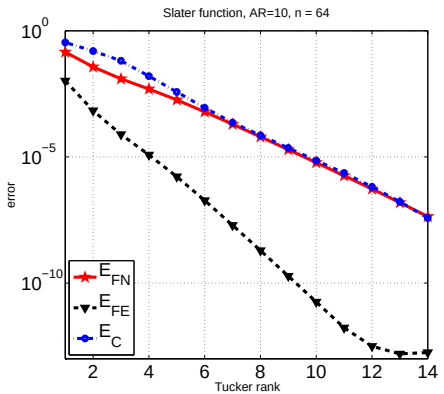
$f(x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^3$ is discretized in $[a, b]^3$. Sampling points: $x_{i_\ell}^{(\ell)} = a + (i_\ell - \frac{1}{2})(\frac{b-a}{n_\ell})$, $i_\ell = 1, 2, \ldots n_\ell$.

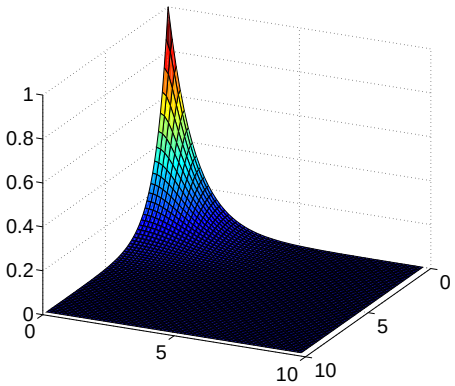
$\Rightarrow$ We generate a tensor $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with entries $a_{ijk} = f(x_i^{(1)}, x_j^{(2)}, x_k^{(3)})$.

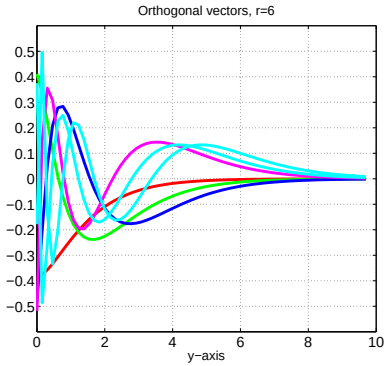


Example: Slater function $g(x) = \exp(-\alpha\|x\|)$ with $x = (x^{(1)}, x^{(2)}, x^{(3)})^T \in \mathbb{R}^3$

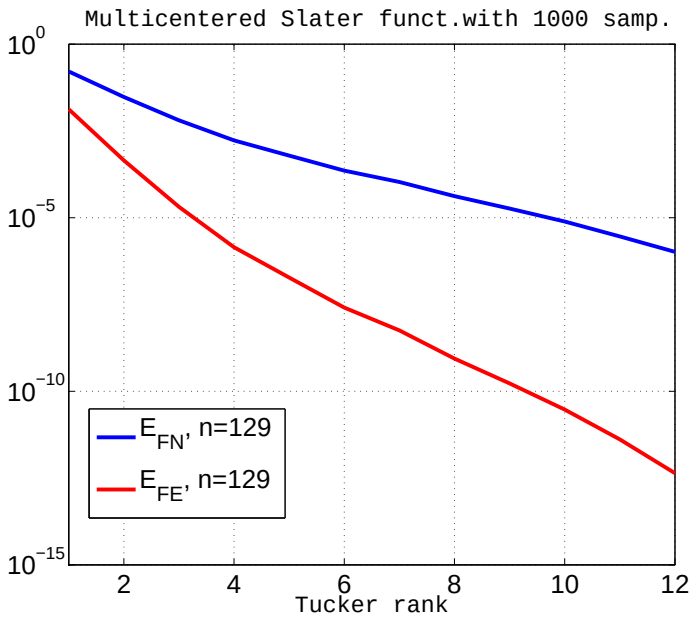
$$E_{FN} = \frac{\|A_0 - A_{(r)}\|}{\|A_0\|} \quad E_{FE} = \frac{\|A_0\| - \|A_{(r)}\|}{\|A_0\|}, \quad E_C := \frac{\max_{i\in\mathcal{I}} |a_{0,i} - a_{r,i}|}{\max_{i\in\mathcal{I}} |a_{0,i}|}.$$

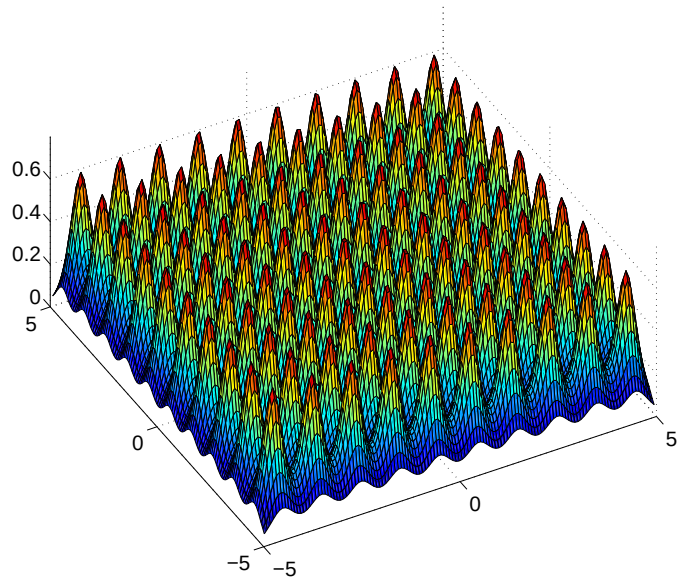






MGA BTA for 3D periodic structures in a box





The "multi-centered Slater potential" obtained by displacing a single Slater potential with respect to the $m \times m \times m$ spatial grid of size $H > 0$, (here $m = 10$)

$$g(x) = \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} e^{-\alpha \sqrt{(x_1 - iH)^2 + (x_2 - jH)^2 + (x_3 - kH)^2}}.$$

Canonical to Tucker approximation: RHOSVD

To reduce the ranks of canonical rank-R input tensors, we develop the two-level canonical-to-Tucker (C2T) approximation with the consequent Tucker-to-canonical (T2C) approx.:

$$\mathcal{C}_{R,\mathbf{n}} \quad \overset{I}{\to} \quad \mathcal{T}_{\mathcal{C}_R,\mathbf{r}} \quad \overset{II}{\to} \quad \mathcal{T}_{\mathcal{C}_{R'},\mathbf{r}}. \tag{6}$$

Definition Reduced HOSVD (RHOSVD).
Given $A^0 = \boldsymbol{\xi} \times_1 U^{(1)} \times_2 U^{(2)} \dots \times_d U^{(d)} \in \mathcal{C}_{R,\mathbf{n}}$, $\boldsymbol{\xi} = diag\{\xi_1, \dots, \xi_R\}$,
introduce the truncated SVD of the side-matrices $U^{(\ell)}$, $Z_0^{(\ell)} D_{\ell,0} {W_0^{(\ell)}}^T$, $(\ell = 1, ..., d)$,
where $D_{\ell,0} = \mathrm{diag}\{\sigma_{\ell,1}, \sigma_{\ell,2}, ..., \sigma_{\ell,r_\ell}\}$, while $Z_0^{(\ell)} \in \mathbb{R}^{n \times r_\ell}$ and ${W_0}^{(\ell)} \in \mathbb{R}^{R \times r_\ell}$ are the respective submatrices of $Z^{(\ell)}$ and $W^{(\ell)}$ in SVD of $U^{(\ell)}$.
Then the RHOSVD approximation of A^0 is given by

$$A^0_{(\mathbf{r})} = \boldsymbol{\xi} \times_1 \left[Z_0^{(1)} D_{1,0} {W_0^{(1)}}^T \right] \times_2 \left[Z_0^{(2)} D_{2,0} {W_0^{(2)}}^T \right] \cdots \times_d \left[Z_0^{(d)} D_{d,0} {W_0^{(d)}}^T \right]. \tag{7}$$

Note that $A^0_{(\mathbf{r})}$ in (7) is obtained by the projection of A^0 onto the matrices of left singular vectors $Z_0^{(\ell)}$.

Canonical to Tucker Approximation

Theorem [B.N. Khoromskij, V.Kh. (SISC, 34 (2009))]
Let $A \in \mathcal{C}_{R,\mathbf{n}}$. Then the minimisation problem

$$A \in \mathcal{C}_{R,\mathbf{n}} \subset \mathbb{V}_{\mathbf{n}} : \quad A_{(\mathbf{r})} = \operatorname{argmin}_{V \in \mathcal{T}_{\mathbf{r},\mathbf{n}}} \|A - V\|_{\mathbb{V}_{\mathbf{r}}}, \tag{8}$$

is equivalent to

$$[Z^{(1)}, ..., Z^{(d)}] = \operatorname{argmax}_{V^{(\ell)} \in \mathcal{M}_\ell} \left\| \sum_{\nu=1}^{R} \xi_\nu \left(V^{(1)\,T} u_\nu^{(1)} \right) \otimes ... \otimes \left(V^{(d)\,T} u_\nu^{(d)} \right) \right\|_{\mathbb{V}_{\mathbf{r}}}^2,$$

$V^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$ – orthogonal. Init. guess: RHOSVD $Z_0^{(\ell)} \Rightarrow$ the truncated SVD of $U^{(\ell)} = [u_1^{(\ell)} ... u_R^{(\ell)}]$, under the compatibility condition

$$r_\ell \le rank(U^{(\ell)}) \quad \text{with} \quad U^{(\ell)} = [u_1^{(\ell)} ... u_R^{(\ell)}] \in \mathbb{R}^{n \times R},$$

Error bounds for RHOSVD

$$\|A - A_{(\mathbf{r})}^0\| \le \|\boldsymbol{\xi}\| \sum_{\ell=1}^{d} \Big(\sum_{k=r_\ell+1}^{\min(n,R)} \sigma_{\ell,k}^2 \Big)^{1/2}, \qquad \text{where } \|\boldsymbol{\xi}\|^2 = \sum_{\nu=1}^{R} \xi_\nu^2.$$

Canonical to Tucker Approximation

The maximizer $Z^{(\ell)} = [z_1^{(\ell)} ... z_{r_\ell}^{(\ell)}] \in \mathbb{R}^{n \times r_\ell}$ is obtained by ALS iteration using the matrix unfolding

$$B_{(q)} \in \mathbb{R}^{n \times \bar{r}_q}, \quad \bar{r}_q = r_1 ... r_{q-1} r_{q+1} ... r_d. \tag{9}$$

The minimiser in (8) is then

$$A_{(\mathbf{r})} = \boldsymbol{\beta} \times_1 Z^{(1)} \times_2 Z^{(2)} \times_3 \ldots \times_d Z^{(d)}.$$

$$\boldsymbol{\beta} = \sum_{\nu=1}^{R} \xi_\nu (Z^{(1)\,T} u_\nu^{(1)}) \otimes ... \otimes (Z^{(d)\,T} u_\nu^{(d)}) \in \boldsymbol{\mathcal{C}}_{R,\mathbf{r}}.$$

Complexity for $d = 3$, $n \geq R$: $W_{C \to T} = O(nR^2 + nr^4)$.

C2T $\Rightarrow$ T2C: $R \to R' \leq r^2 \ll R$

Multilinear operations in $\mathcal{C}_{R,\mathbf{n}}$ tensor format

$$A_1 = \sum_{k=1}^{R_1} c_k u_k^{(1)} \otimes \ldots \otimes u_k^{(d)}, \quad A_2 = \sum_{m=1}^{R_2} b_m v_m^{(1)} \otimes \ldots \otimes v_m^{(d)}.$$

Euclidean scalar product (complexity $O(dR_1R_2n) \ll n^d$),

$$\langle A_1, A_2 \rangle := \sum_{k=1}^{R_1} \sum_{m=1}^{R_2} c_k b_m \prod_{\ell=1}^{d} \left\langle u_k^{(\ell)}, v_m^{(\ell)} \right\rangle.$$

Hadamard product of A_1, A_2

$$A_1 \odot A_2 := \sum_{k=1}^{R_1} \sum_{m=1}^{R_2} c_k b_m \left(u_k^{(1)} \odot v_m^{(1)} \right) \otimes \ldots \otimes \left(u_k^{(d)} \odot v_m^{(d)} \right).$$

Convolution, for $d = 3$ ($O(R_1R_2n \log n) \ll n^3 \log n$ (3D FFT))

$$A_1 * A_2 = \sum_{k=1}^{R_1} \sum_{m=1}^{R_2} c_k b_m (u_m^{(1)} * v_k^{(1)}) \otimes (u_m^{(2)} * v_k^{(2)}) \otimes (u_m^{(3)} * v_k^{(3)})$$

In electronic structure calculations we need large $n \times n \times n$ 3D Cartesian grids, and robust rank reduction for large R.

The conventional BTA algorithm ($d = 3$) $\Rightarrow$ small n, moderate R:
(a) For full format tensors

$$W_{F2T} = O(n^{d+1}) \quad \Rightarrow \quad W_{F2T} = O(n^4)$$

(b) for the canonical rank-R input tensors

$$W_{C2T} = O(Rn \min\{R, n\} + r^{d-1} n \min\{r^{d-1}, n\}),$$

$$\Rightarrow W_{C2T} = O(nR^2 + nr^4).$$

$\Rightarrow$ SVD of side matrices of the sizes $\geq 10^4 \times 10^4$ is computationally unfeasible.

The answer to the problem:

Multigrid accelerated best Tucker approximation:

$$W_{F2T} = O(n^3), \quad \text{for full size tensors}$$

$$W_{C2T} = O(rRn), \quad \text{for canonical rank-R tensors.}$$

MGA Tucker approximation for $\mathcal{C}_{R,\mathbf{n}}$ tensors

1. Sequence of nonlinear appr. problems for $A = A_n$,$n = n_m := n_0 2^m$, $m = 0, 1, \ldots, M$, on a sequence of refined grids $\omega_{\mathbf{3},n_m}$.
2. $Z_0^{(q)}$ on grid $\omega_{\mathbf{3},n_m} \rightarrow$ by linear interp. of $Z^{(q)} \in \mathbb{R}^{n_{m-1} \times r_q}$ from $\omega_{\mathbf{3},n_{m-1}}$.
3. The *restricted ALS iteration*, is based on "most important fibers" (MIFs) of unfolding matrices.
   Positions of MIFs are extracted at the coarsest grid

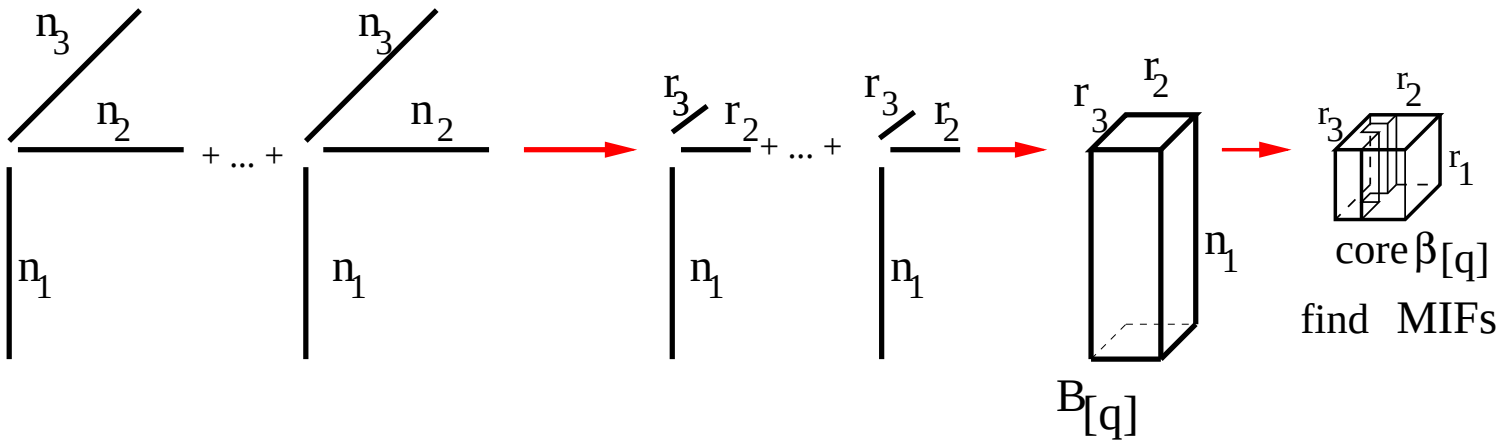
$$\boldsymbol{\beta}_{(q)} = {Z^{(q)}}^T B_{(q)} \in \mathbb{R}^{r_q \times \overline{r}_q}. \tag{10}$$

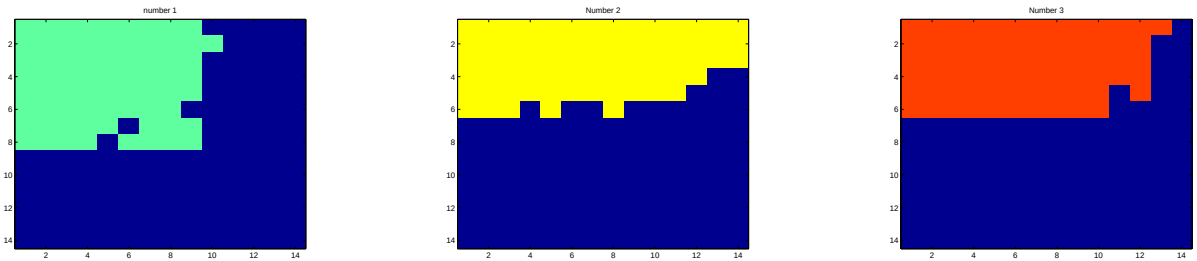
   Location of MIFs corresponds to pr columns in $\boldsymbol{\beta}_{(q)}$, $p \ll r$ with maximal Euclidean norms (maximal energy principle).
4. Complexity for $d = 3$: $O(Rrn_M + p^2 r^2 n_M)$.
   (Linear w.r.t. n and R !)

MGA C2T ($\mathcal{C}_{R,\mathbf{n}}$)

[B.N. Khoromskij, V.Kh. (SISC, 34 (2009))]

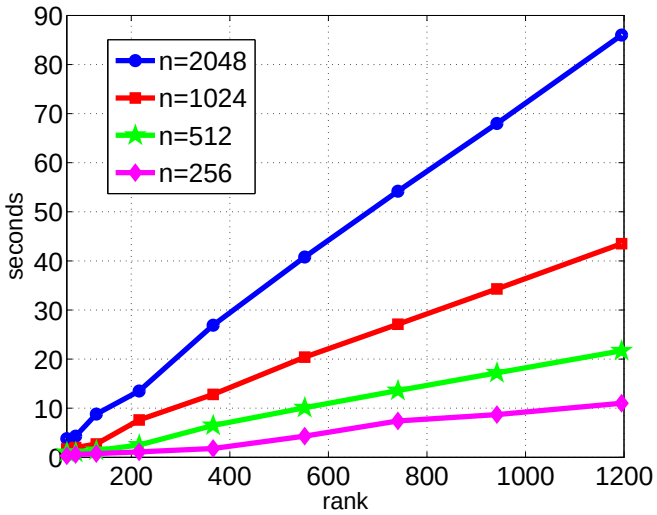
$B_{(q)} \to$ unfolding of $(B_{[q]})$, $\quad \beta_{(q)} = {Z^{(q)}}^T B_{(q)}$; $n_0 = 64$, $n_M = 16384$

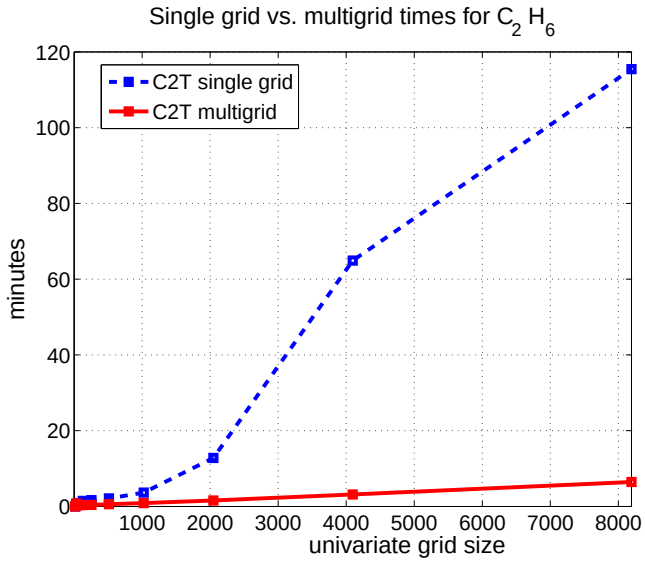




MIFs for H_2O, $p = 4$.

MGA Tucker vs. single grid BTA





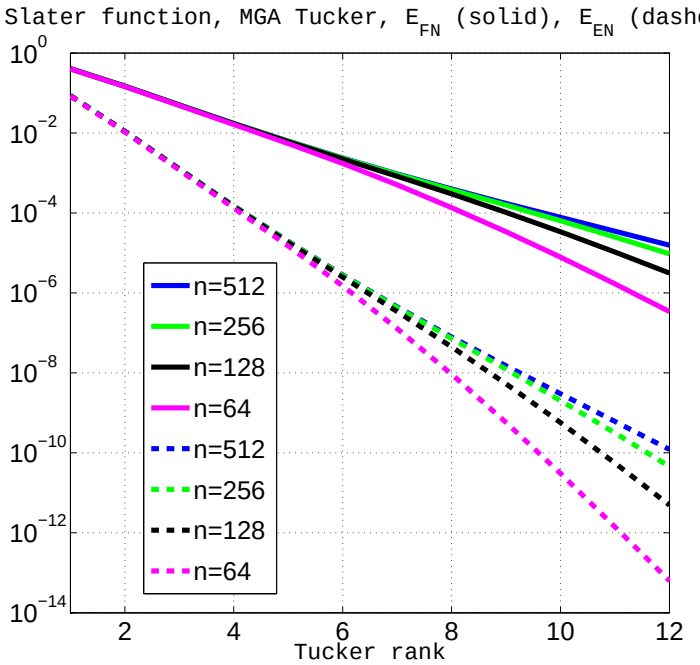
Linear scaling in R and in n (left) for C2T algorithm.

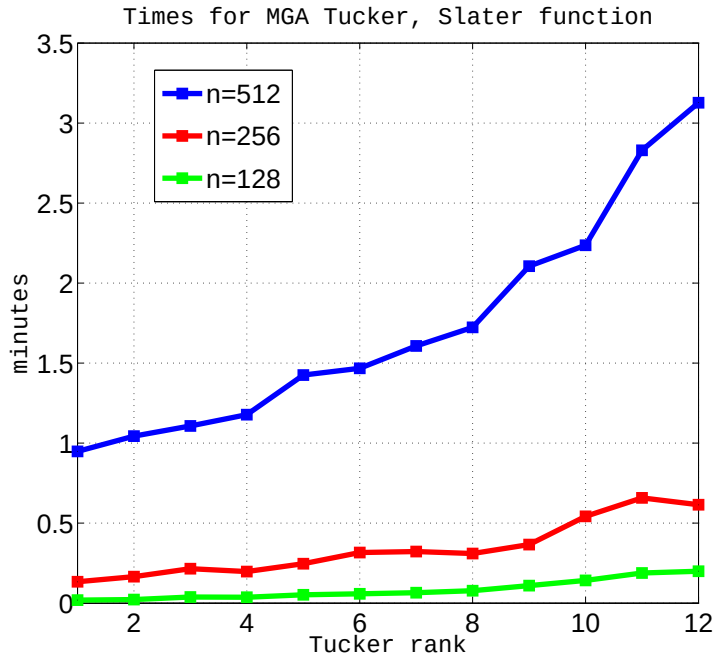
Multigrid vs. single grid times for the C2T approximation of V_H for C_2H_6 (right).

MGA best Tucker decomposition for full format tensors,

[V.Kh. '10, in progress]

$$W_{F\to T} = O(n^3) \text{ for MGA BTA, instead of } O(n^4) \text{ for BTA.}$$

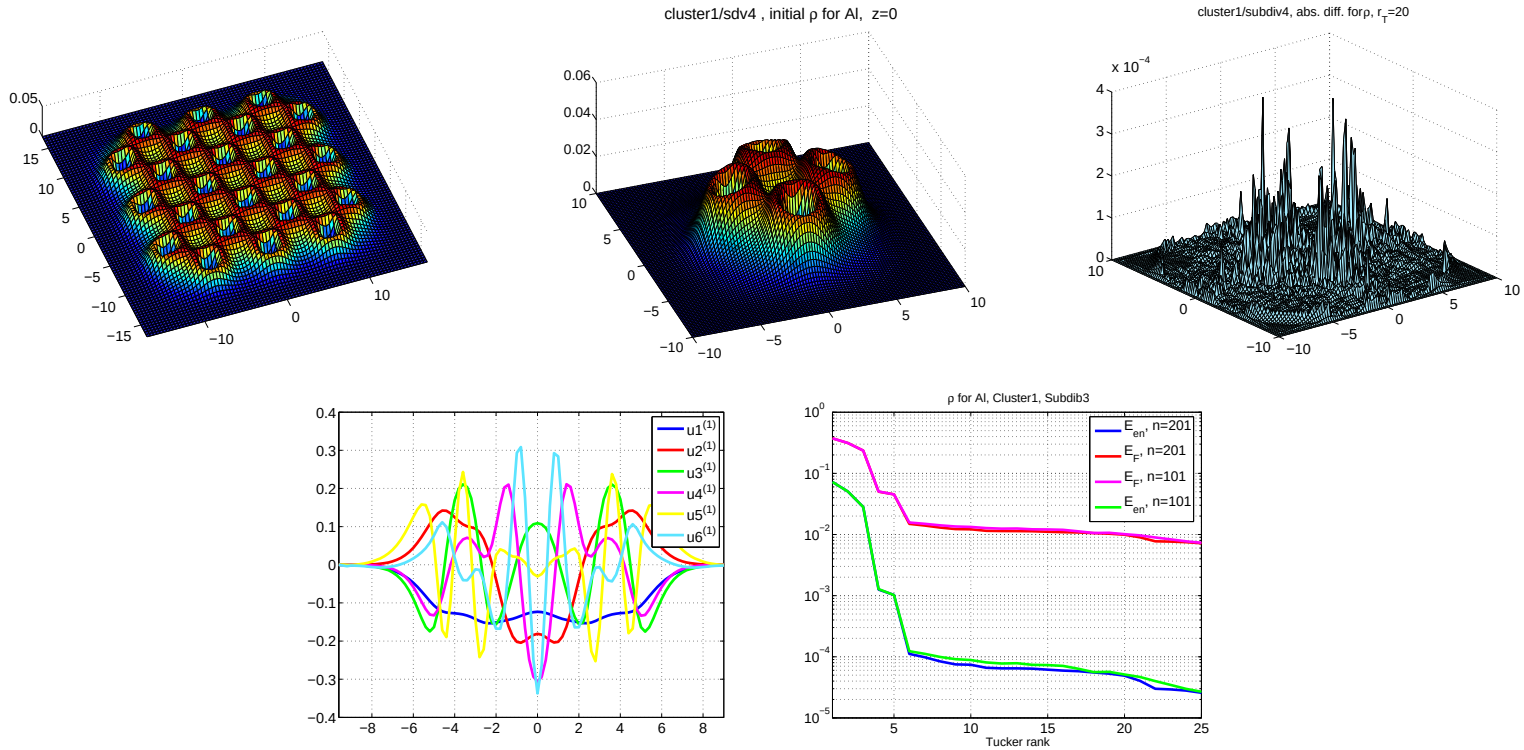




Full size tensor with number of entries $\leq 512^3$, times for Matlab.

Other examples of MGA BTA for full format tensors (Aluminium Clusters)

[T. Blesgen, V. Gavini and V.Kh., '09]



FE (coarse graining) $\Rightarrow$ computed on large supercomputers (80 processors) *; (clusters: 14, 172, 365 and 666 atoms).

FE data* $\Rightarrow$ 3D Cartesian grid $\Rightarrow$ Tucker decomposition

* [V. Gavini, J. Knap, K. Bhattacharya, and M. Ortiz,'07]

Other grid-based approaches and tensor formats

- Plane waves (BigDFT)
  Groups of R. Schneider, S. Goedicker.
- Wavelets - canonical format
  G. Beylkin, M. Mohlenkamp,
  R. Chinnamsetty, M. Espig, H.-J. Flad, W. Hackbusch, B. Khoromskij.
- 3D Adaptive cross - Tucker format
  I. Oseledets, D. Savostyanov, E. Tyrtyshnikov.
- Hierarchical dimension splitting, hierarchical SVD of tensors
  W. Hackbusch, L. Grasedyck, S. Kühn.
- TT, TC, and QTT formats
  I. Oseledets, E. Tyrtyshnikov, B. Khoromskij.
- Best TT approximation
  R. Schneider, T. Rohwedder, S. Holtz.

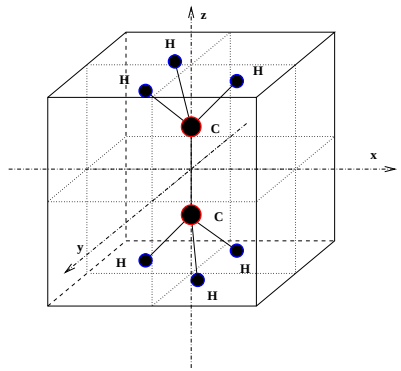
$O(n)$ computation of the Hartree potential on huge 3D grids

V_H is computed on large $n \times n \times n$ Cartesian grids, with $n \sim 10^3 \div 10^4$

$$\rho(x) = 2\sum_{i=1}^{N_{orb}} (\varphi_i)^2, \quad \varphi_i = \sum_{k=1}^{R_0} c_{i,k} g_k(x), \quad x = (x_1, x_2, x_3) \in \mathbb{R}^3$$

$g_k(x) = g_{k,1}(x_1)\, g_{k,2}(x_2)\, g_{k,3}(x_3)$,
$g_{k,\ell} = (x_\ell - A_{k,\ell})^{\beta_{k,\ell}} e^{-\lambda_k (x_\ell - A_{k,\ell})^2}$,

$g_{k,\ell} \to u_k^{(\ell)}; x_\ell \in [-b, b]$,
$b = 6 \div 10 (\text{in } \mathring{A})$.



$R_0(CH_4) = 55, \quad R_0(C_2H_6) = 96, \quad R_0(H_2O) = 41, \quad R_\rho = \frac{R_0(R_0+1)}{2}$.

$$\rho \approx F_\rho = \sum_{k=1}^{R_\rho} c_k u_k^{(1)} \otimes u_k^{(2)} \otimes u_k^{(3)}, \quad c_k \in \mathbb{R}, \quad u_k^{(\ell)} \in \mathbb{R}^n.$$

$\Rightarrow \rho$ is presented in the computation box $[-b, b]^3$ with $R_\rho \sim 10^3 \div 10^4$.

$O(n)$ computation of the Hartree potential on huge 3D grids

$R_\rho(CH_4) = 1540, \quad R_\rho(C_2H_6) = 4656, \quad R_\rho(H_2O) = 861$

Next : MGA C2T + T2C transformations to reduce the rank:

$$R_{\rho_r} \sim 10^2.$$

Canonical tensor G for the projected Newton kernel $G = \left[\langle \frac{1}{\|x\|}, \phi_{\mathbf{i}} \rangle\right]$, where $\phi_{\mathbf{i}}$ are p.w.c. basis functions. (Obtained using *sinc*-quadratures) [C. Bertoglio, B. Khoromskij '08]).
Rank of G, $R_N \sim 20 \div 30$.

The rank-R_{ρ_r} canonical tensor F for ρ:

$$V_H = F * G = \sum_{m=1}^{R_{\rho_r}} \sum_{k=1}^{R_N} c_k b_m \left(u_1^k * v_1^m\right) \otimes \left(u_2^k * v_2^m\right) \otimes \left(u_3^k * v_3^m\right)$$

Coulomb matrix J: tensor inner products in GTO basis $\{g_k\}_{k=1}^{R_0}$

$$J_{km} := \int_{\mathbb{R}^3} g_k(x) g_m(x) V_H(x) dx, \quad k, m = 1, \ldots R_0.$$

Tensor-product convolution vs. 3D FFT

The cost of computation of V_H by tensor product convolution (1D FFT):

$$\mathcal{N}_{C*C} = O(R_{\rho_r} R_N n \log n)$$

instead of $O(n^3 \log n)$ for 3D FFT.

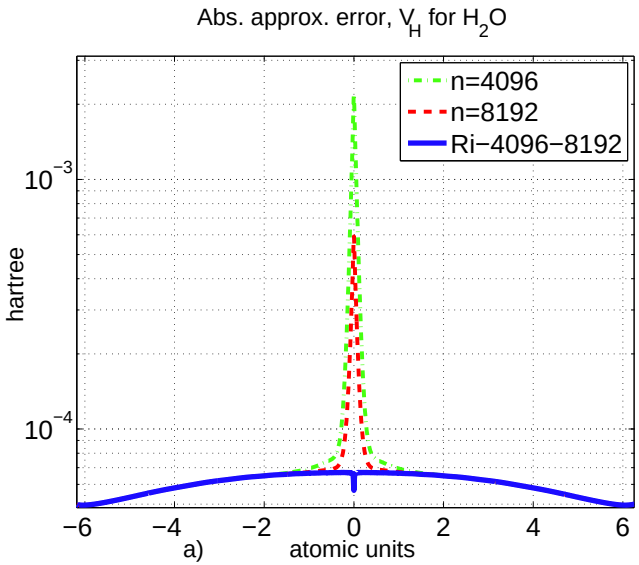
n^3	128^3	256^3	512^3	1024^3	2048^3	4096^3	8192^3	16384^3
FFT_3	4.3	55.4	582.8	~ 6000	–	–	–	~ 2 years
$C * C$	0.2	0.9	1.5	8.8	20.0	61.0	157.5	299.2
C2T	4.2	4.7	5.6	6.9	10.9	20.0	37.9	86.0

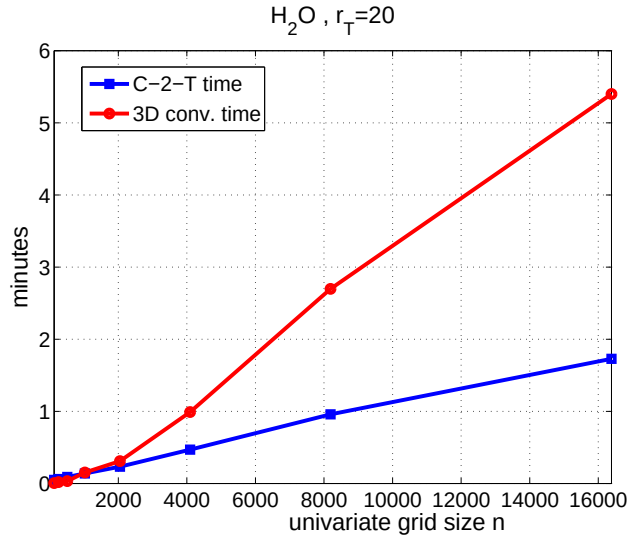
CPU time (in sec) for the computation of V_H for H_2O.

(3D FFT time for $n \geq 1024$ is obtained by extrapolation).

High accuracy computations of V_H

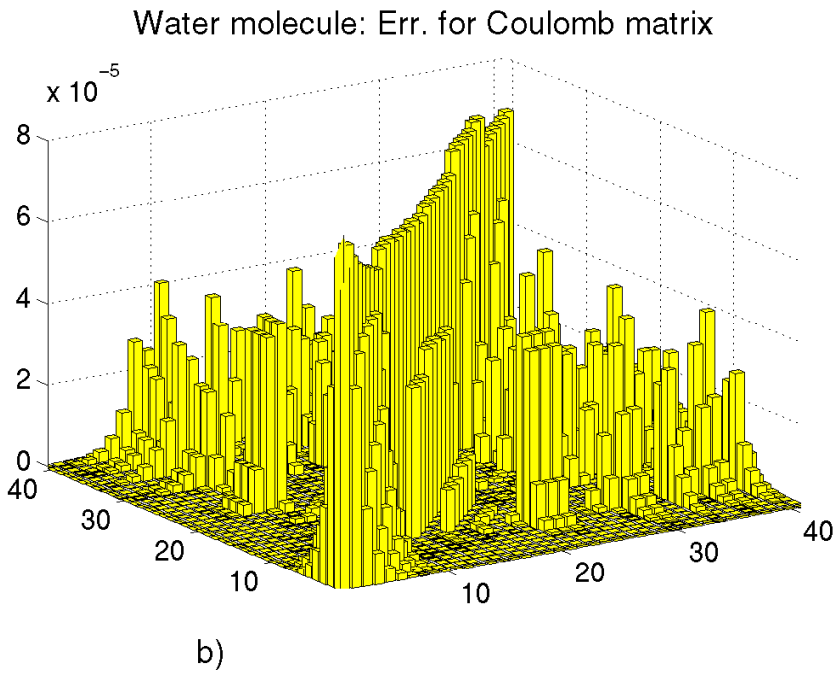
(Calculation over the number of entries $N \approx 4.3 \cdot 10^{12}$).

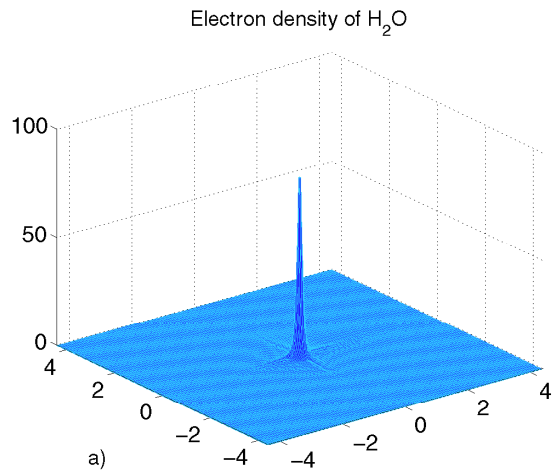




High accuracy computations of the Coulomb matrix for H_2O

(for calculation over the number of entries $N \approx 4.3 \cdot 10^{12}$). Approx. error $O(h^3)$, $h = 2b/n$ ($h \approx 0.0008$ $\mathring{A}$).

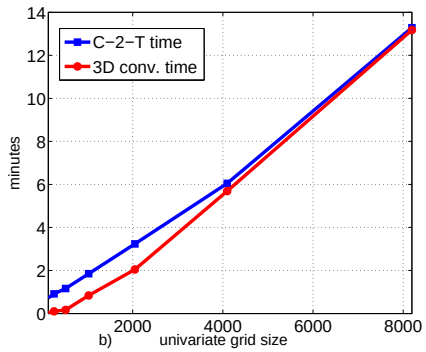


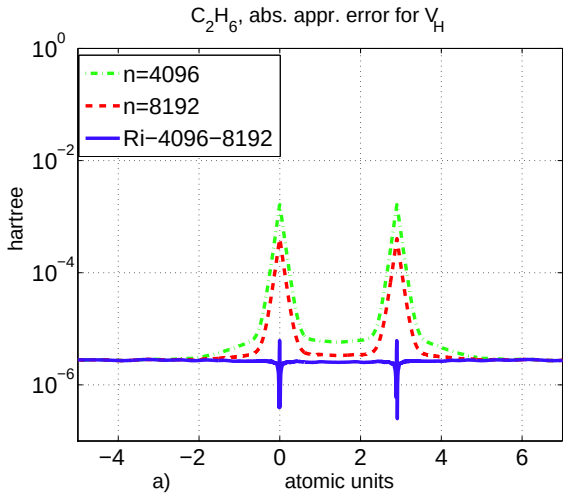


Richardson extrapolation: $J^{(n)}_{Rich} = (4 \cdot J^{(2n)} - J^{(n)})/3.$

High accuracy computations of the Coulomb matrix for C_2H_6

a) CPU times for V_H : $O(n \log n)$ on $n \times n \times n$ grids with max.: $n = 8192$.
b) Absolute approximation error of the tensor-product computation of the Coulomb matrix for C_2H_6 molecule.





TS computation of the exchange matrix

[V.Kh., '09]

$$K_{k,m} := \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} g_k(x)\frac{\tau(x,y)}{|x-y|}g_m(y)dxdy, \quad k,m = 1,\dots R_0$$

1. Compute the discrete tensor product convolution

$$W_{a,m}(x) = \int_{\mathbb{R}^3} \frac{\varphi_a(y)g_m(y)}{|x-y|}dy, \quad \text{cost: } O(R_N R_0 n \log n) \text{ per element } (\times R_0 N)$$

2. Entries of the Exchange matrix for every orbital,

$$V_{km,a} = \int_{\mathbb{R}^3} \varphi_a(x)g_k(x)W_{a,m}(x)dx$$

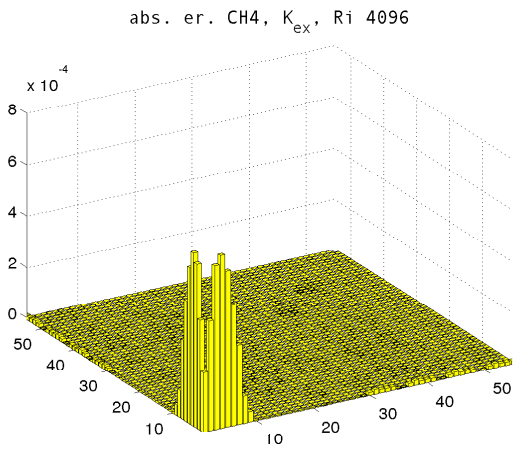
with the cost $O(R_N R_0^4 n_f)$.

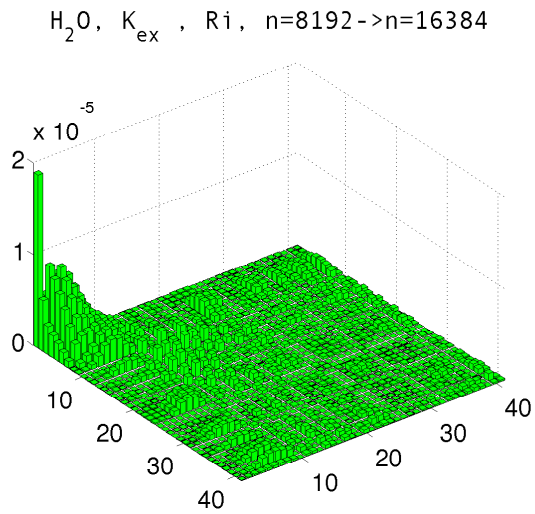
3. Entries of the Exchange matrix

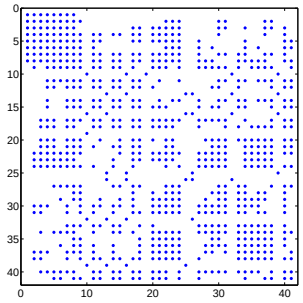
$$K_{k,m} = \sum_{a=1}^{N_{orb}} V_{km,a}$$

TS computation of the exchange matrix on huge 3D grids

Abs. error for the Exchange matrices K_{ex} for CH_4 (Richardson extrapolation ($4096 \times 4096 \times 4096$, for $n = 4096$)
and H_2O ($N_{vol} = 16384^3$), $h \approx 0.0008$ $\mathring{A}$, $O(h^3)$.







Sparsity of K_{ex} computations on large grids.

Solution of the Hartree-Fock on a sequence of 3D grids

[Flad, Khoromskij, V.Kh., (SISC, 2010, accepted)]

Discretized GTO basis, $\varphi_i(x) \approx \sum_{k=1}^{R_0} c_{ki}\overline{g}_k(x)$, $x \in \mathbb{R}^3$, $i = 1, ..., N_{orb}$

$$\mathcal{F}\varphi_i(x) = \lambda_i \varphi_i(x), \quad \Rightarrow \quad \mathbf{F}C_i = \lambda_i \mathbf{S}C_i$$

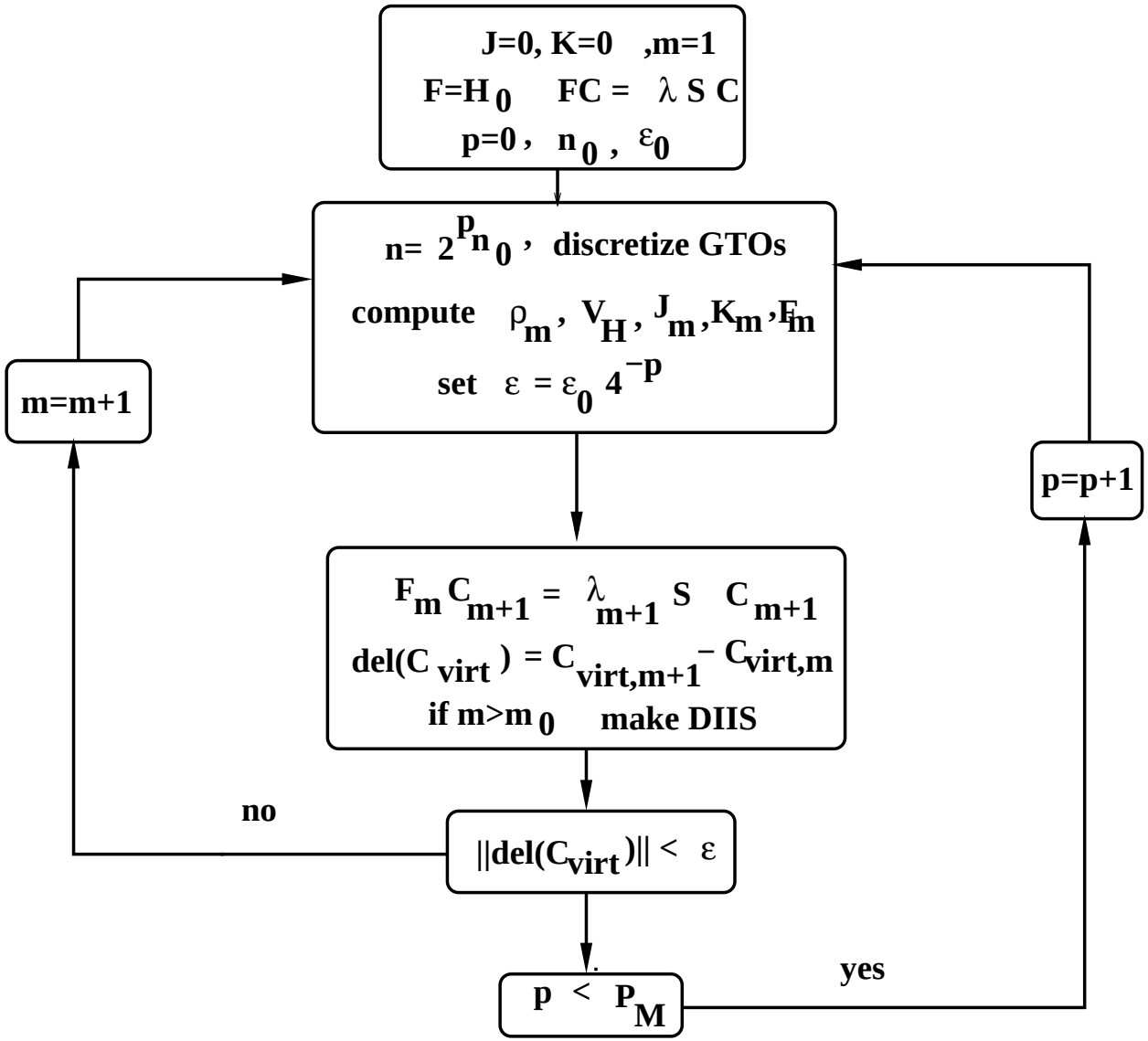
$$\mathbf{F} = H_0 + J(C) - K(C), \quad C = \{c_{ki}\} = [C_1 C_2 \dots C_{N_{orb}}] \in \mathbb{R}^{R_0 \times N_{orb}}$$

H_0 is the stiffness matrix of the core Hamiltonian $-\frac{1}{2}\Delta + V_c$

Our iterative scheme

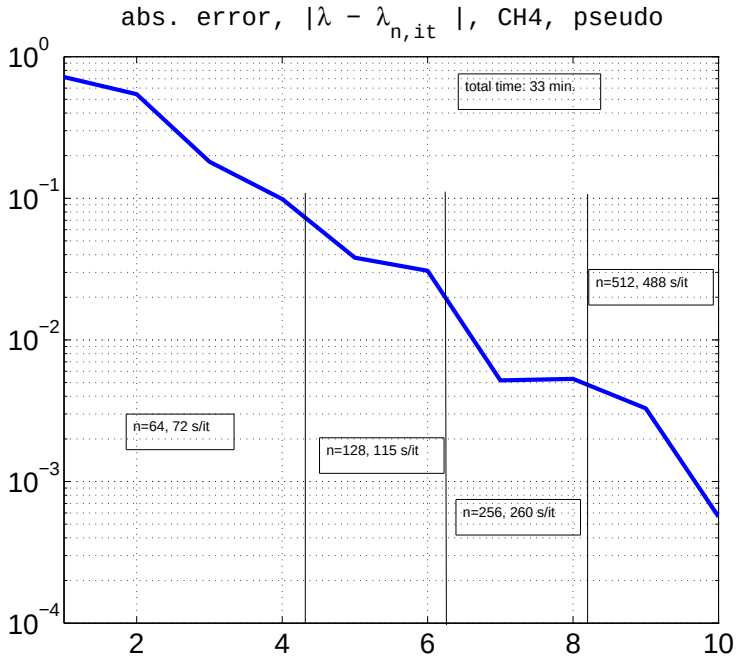
- Initial guess for $m = 0$: $J = 0$, $K = 0$ $\Rightarrow$ $\mathbf{F}_0 = H_0$.
- On a sequence of refined grids $n = n_0, 2n_0, \dots, 2^p n_0$, solve EVP $[H_0 + J_m(C) - K_m(C)]U = \lambda SU$.
- Fast update of $J_m(C)$ and $K_m(C)$.
- Grid dependent termination criteria $\varepsilon_p = \varepsilon_0 \cdot 4^{-p}$.
- Use DIIS **[Pulay '80]** to provide fast convergence.

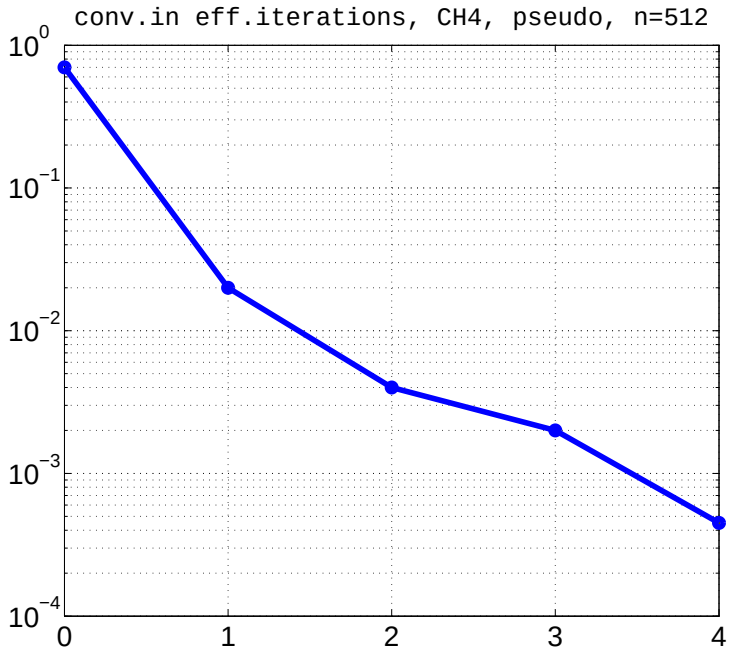
Flow-chart of the HF solver



HF solver for the pseudopotential case of CH_4

Left: multilevel convergence of iterations for the pseudopotential case of CH_4.
Right: convergence in effective iterations for finest grid.

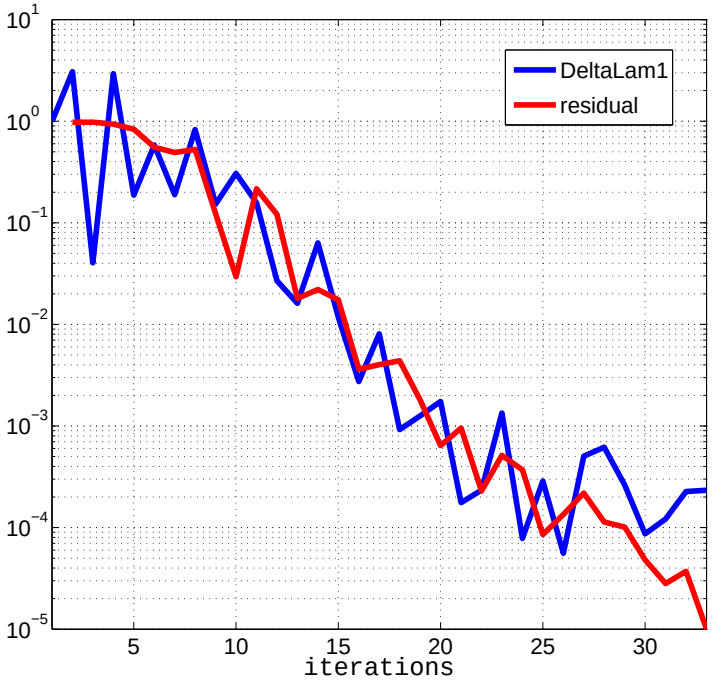


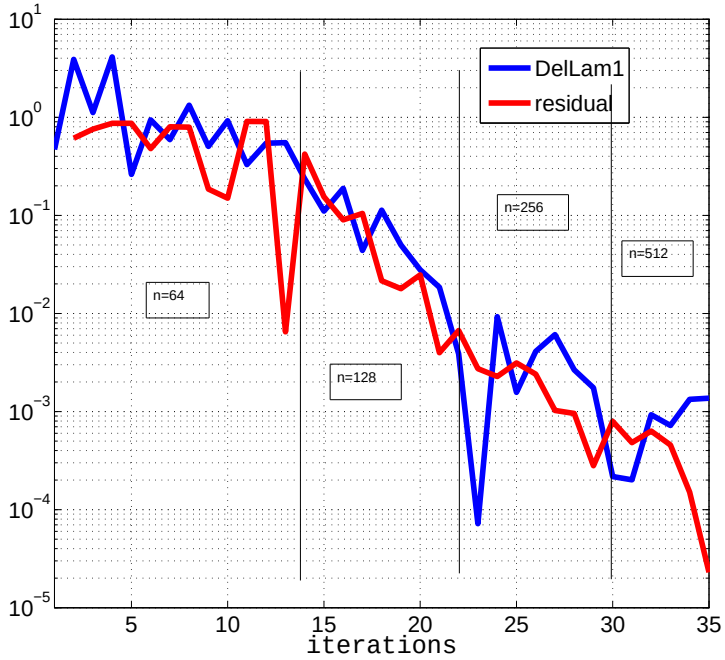


Cost per iteration $\sim O(n_p)$.

Iterations history of the nonlinear 3D EVP solver

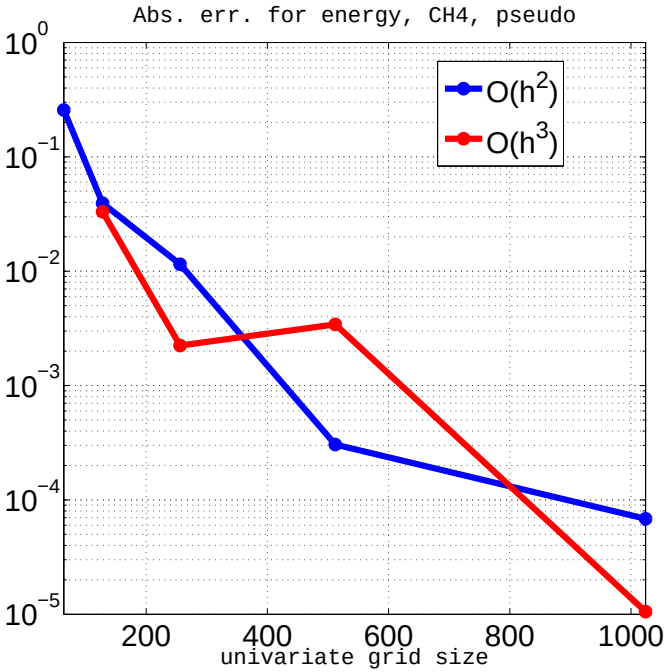
Pseudopotential case of CH_3OH and C_2H_5OH. $\varepsilon = 2^{-2p}\varepsilon_0$.

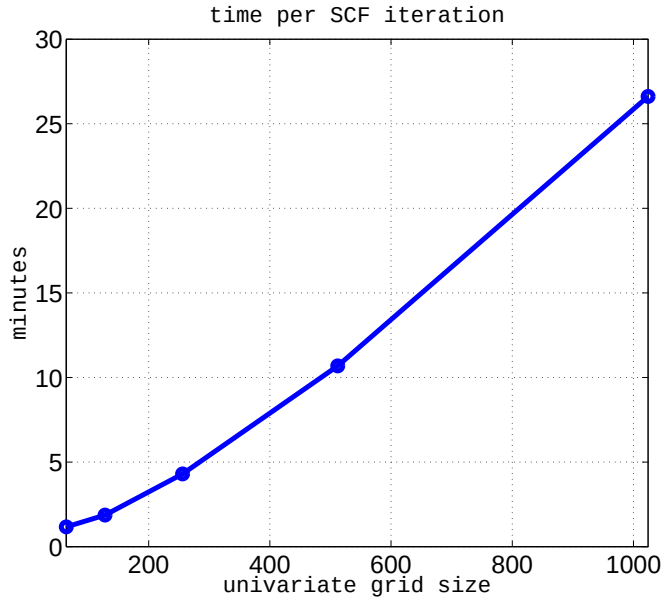




Convergence in HF energy

Left: convergence in HF energy.
Right: linear scaling of the iteration time in the univariate grid size.





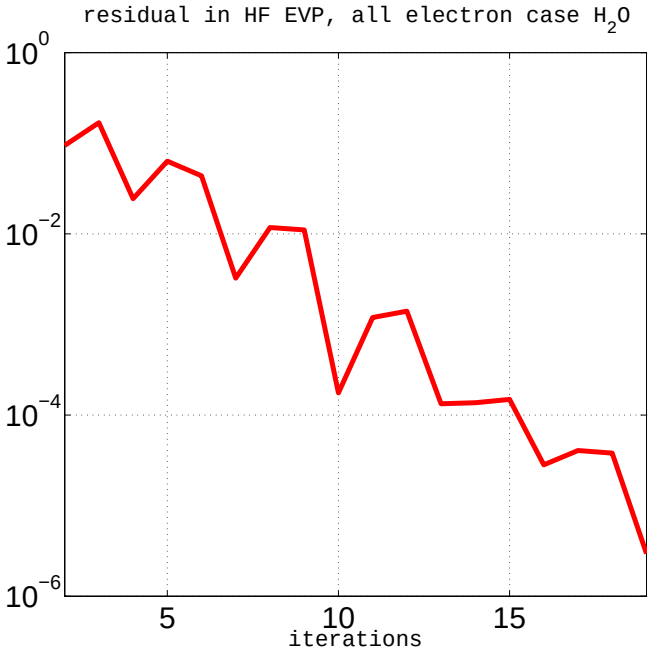
The total energy is calculated by

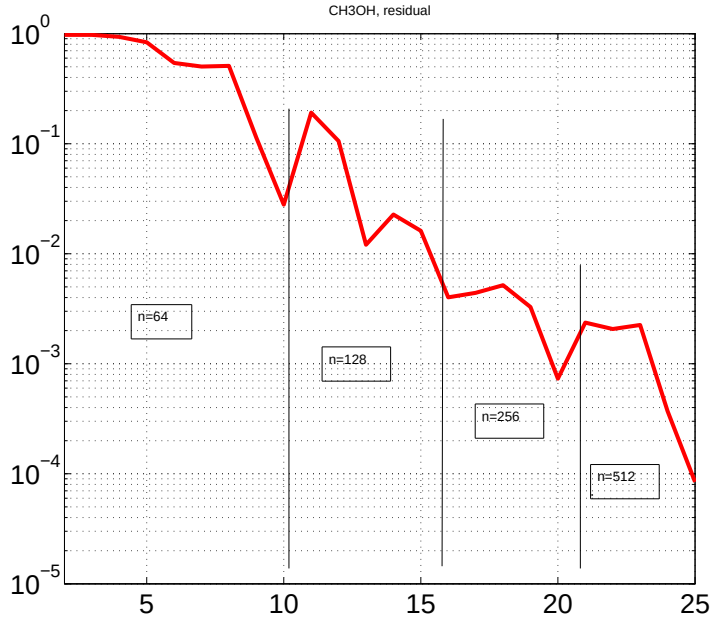
$$E_{HF} = 2\sum_{i=1}^{N_{orb}} \lambda_i - \sum_{i=1}^{N_{orb}} \left(\widetilde{J}_i - \widetilde{K}_i\right)$$

with $\widetilde{J}_i = (\varphi_i, V_H\varphi_i)_{L^2}$ and $\widetilde{K}_i = (\varphi_i, \mathcal{V}_{ex}\varphi_i)_{L^2}$ $(i = 1, \ldots, N_{orb})$ the Coulomb and exchange integrals, computed with respect to the orbitals φ_i.

Iterations history of the nonlinear 3D EVP solver

Left: all electron case of H_2O.
Right: the pseudopotential case of CH_3OH, constant ε.





Resume:

- Tensor-structured numerical methods for a class of multidimensional problems based on separable representation of functions and operators are first developed.
- The complete numerical scheme using tensor-structured methods for the solution of the Hartree-Fock equation is presented.
- Grid-based evaluation of 3D and 6D integral operators scales linearly with respect to 1D grid size.
- Comprehensive numerical tests for moderate size molecules confirm efficiency of algorithms.

Publications and presentations (2006-2010):

- The results are published in SIAM Journal of Scientific Computing (2), Journal of Computational physics (1), Computational Methods in Applied mathematics (1), Central European Journal of Mathematics (1).
- Talks (9) and posters (2) at the international conferences and seminars.
- The tensor-structured HF solver package (Matlab) contains more than 100 functional routines and about 200 test programs.

Thank you for attention