15 Lectures on Tensor Numerical Methods for Multi-dimensional PDEs

Lect. 12. Tensor-structured calculation of 3D integrals in 1D complexity.

Primer: 3D grid-based Hartree-Fock solver

Boris Khoromskij & Venera Khoromskaia
Shanghai Jiaotong University, Institute of Natural Sciences
April 2017

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

Outline of the lecture

Electronic Schrödinger equation

The Hartree-Fock model originates from the electronic Schrödinger equation,

$$\mathcal{H}_e \Psi = E\Psi, \tag{1}$$

with the Hamiltonian

$$\mathcal{H}_e = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{x_i - x_A} + \sum_{i<j\leq N}^{N}\frac{1}{|x_i - x_j|}, \quad x_A,\ x_i,\ x_j \in \mathbb{R}^3, \tag{2}$$

describing energy of a molecule in the framework of the so-called Born-Oppenheimer approximation (fixed nuclei). M is the number of nuclei, Z_A – nuclei charges, N – number of electrons.
The electronic Schrödinger equation is practically numerically nontractable multidimensional problem in $\mathbb{R}^{3N}$ (except for Hydrogen atom).

The Hartree-Fock equation is a $3D$ problem in space variables (solvable!) obtained as a result of the minimization of the energy functional for (1).
The underlying condition for the wavefunction Ψ for electrons (fermions) Ψ is that it should be antisymmetric, therefore it is parametrized using a single Slater determinant containing the products of electronic orbitals.

$$\Psi = \frac{1}{N!}\begin{vmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \dots & \varphi_N(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \dots & \varphi_N(x_2) \\ \dots & \dots & \dots & \dots \\ \varphi_1(x_N) & \varphi_2(x_N) & \dots & \varphi_N(x_N) \end{vmatrix},$$

where $\varphi_i(x_j)$ are the one-electron wavefunctions, $i, j = 1, \dots N$.

The Hartree-Fock equation

Hartree-Fock equation is a basic model for "ab-initio" calculation of the ground state energy of molecular systems (first step in calc. of excited states).
It is a nonlinear eigenvalue problem (we consider closed shell molecular systems, $N_{orb} = N/2$)

$$\mathcal{F}\varphi_i(x) = \lambda_i\,\varphi_i(x), \qquad \int_{\mathbb{R}^3}\varphi_i\varphi_j = \delta_{ij},\ i = 1, ..., N_{orb}.$$

where the Fock operator $\mathcal{F}$ depends on the density matrix $\tau(x,y) = 2\sum_{i=1}^{N_{orb}}\varphi_i(x)\varphi_i(y)$, and electron density $\rho(y) = \tau(y,y)$,

$$\mathcal{F}\varphi := [-\frac{1}{2}\Delta - \sum_{\nu=1}^{M}\frac{Z_\nu}{\|x - a_\nu\|} + \int_{\mathbb{R}^3}\frac{\rho(y)}{\|x-y\|}\,dy]\varphi - \frac{1}{2}\int_{\mathbb{R}^3}\frac{\tau(x,y)}{\|x-y\|}\,\varphi(y)dy.$$

Numerical solution of this equation is a challenging numerical task due to nonlocal 3D and 6D integral transforms and presence of strong cusps in the electron density.
Numerical challenges: high accuracy, 3D and 6D singular integrals, strong nonlinearity.

Standard Galerkin scheme

Standart computational scheme developed in quantum chemical community is based on expansion of the molecular orbitals in Gaussian type basis function, $\{g_\mu\}_{1\le\mu\le N_b}$,

$$\varphi_i(x) = \sum_{\mu=1}^{N_b} c_{i\mu} g_\mu(x), \quad i = 1, ..., N_{orb},$$

that yields the Galerkin system of nonlinear equations for coefficients matrix $C = \{c_{i\mu}\} \in \mathbb{R}^{N_{orb}\times N_b}$, (and density matrix $D = 2CC^* \in \mathbb{R}^{N_b\times N_b}$)

$$F(C)C = SC\Lambda, \quad \Lambda = diag(\lambda_1, ..., \lambda_{N_b}), \quad C^T SC = I_{N_b},$$

where $F(C) = H + J(C) + K(C)$.
For a given basis set the core Hamiltonian $H = \{h_{\mu\nu}\}$

$$h_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla g_\mu \cdot \nabla g_\nu dx + \int_{\mathbb{R}^3} V_c(x) g_\mu g_\nu dx \quad 1 \le \mu, \nu \le N_b.$$

and two-electron integrals (TEI)

$$b_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{g_\mu(x) g_\nu(x) g_\kappa(y) g_\lambda(y)}{\|x-y\|} dxdy.$$

are precomputed analytically using erf-functions for Gaussian-type separable basis functions.

Ground state energy

With precomputed TEI and core Hamiltonian H, the Hartree-Fock EVP

$$F(C)C = SC\Lambda$$

is solved iteratively, using DIIS scheme for providing convergence **[Pulay '80]** and updating the Galerkin matrix of the Fock operator, $F(C) = H + J(C) + K(C)$, at every iteration, by recalculating the Coulomb (Hartree) potential and exchange operators

$$J(C)_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} b_{\mu\nu,\kappa\lambda} D_{\kappa\lambda}, \quad K(C) = -\frac{1}{2}\sum_{\kappa,\lambda=1}^{N_b} b_{\mu\lambda,\nu\kappa} D_{\kappa\lambda}.$$

Then the ground state energy of a molecule

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

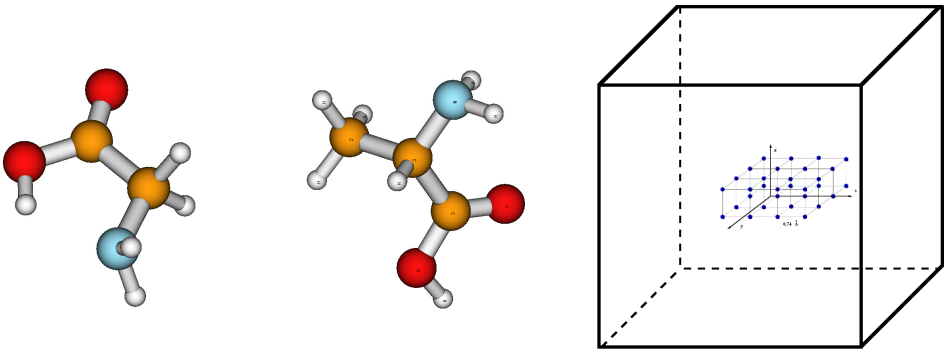
Benchmark packages: MOLPRO, GAUSSIAN, CRYSTAL, ...
All are based on analytical integration of multidimensional operators using separable GTO basis.
Require a large number of sharp Gaussians for approximating Slater-type cusps in electron density of molecules.
There are problems for larger systems and heavier atoms.

Tensor-structured HF solver

[VK & Khoromskij, 2008-2014]

Several years ago the idea to solve the Hartree-Fock equation by fully 3D grid based numerical approach seemed to be a fantasy, and the tensor-structured methods did not exist.

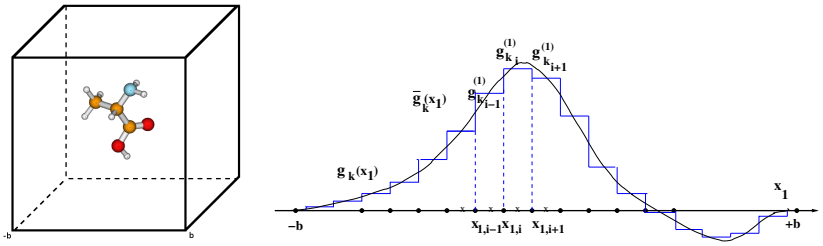
In fact, the tensor numerical methods evolved during the work on this challenging problem.



Glycine, Alanine amino acids, finite lattices in a box.

Grid-based Hartree-Fock solver

Grid-based Hartree-Fock solver **[Khoromskaia, Khoromskij '08 - '14]**
Discretization of basis functions:



Computational box $[-b, b]^3$. For single molecules, $b = 20au$, and $n \times n \times n$ 3D Cartesian grid, with $n^3 \sim 10^{15}$, and step-size $h \sim 10^{-4}$ Å.

GTO basis is presented on 3D grids, using p.w.c. interpolation,

$$g_\mu(x) = p_\mu(x_1, x_2, x_3) e^{\alpha_\mu (x_1^2 + x_2^2 + x_3^2)}$$

$$\approx \mathbf{I}_0 g_\mu := \overline{g}_k(x) = \prod_{\ell=1}^{3} \overline{g}_k^{(\ell)}(x_\ell) = \prod_{\ell=1}^{3} \sum_{i_\ell=1}^{n} g_\mu^{(\ell)}(x_{\ell,i_\ell}) \zeta_{i_\ell}^{(\ell)}(x_\ell), \quad \ell = 1, 2, 3.$$

Thus we obtain rank-1 tensors: $\mathbf{G}_\mu = \mathbf{g}_\mu^{(1)} \otimes \mathbf{g}_\mu^{(2)} \otimes \mathbf{g}_\mu^{(3)}$, with

$$\mathbf{g}_\mu^{(\ell)} = \{g_\mu^{(\ell)}(x_{\ell,i_\ell})\}_{i_\ell=1}^{n} \in \mathbb{R}^n.$$

Example: 3D integral in 1D complexity (Hartree-Fock and Kohn-Sham eqn.)

Let us consider calculation of the Hartree potential operator in the Hartree-Fock equation.

$$J_{\mu\nu} := \int_{\mathbb{R}^3} g_\mu(x) g_\nu(x) V_H(x) dx, \quad \mu, \nu = 1, \dots N_b \quad x \in \mathbb{R}^3,$$

It is a 3D convolution operator

$$V_H(x) := \int_{\mathbb{R}^3} \frac{\rho(y)}{\|x - y\|} dy$$

with the electron density (a function with multiple strong cusps),

$$\rho(x) = 2 \sum_{i=1}^{N_{orb}} (\varphi_i)^2, \quad \varphi_i(x) = \sum_{\mu=1}^{N_b} c_{i\mu} g_\mu(x), \quad i = 1, ..., N_{orb}.$$

In tensor-structured approach for calculating the Hartree potential operator we use grid-based basis functions (approximated by 1D p.w.c.finite elements),

$$g_\mu(x) \approx \mathbf{G}_\mu = \mathbf{g}_\mu^{(1)} \otimes \mathbf{g}_\mu^{(2)} \otimes \mathbf{g}_\mu^{(3)}, \quad \mathbf{G}_\mu \in \mathbb{R}^{n \times n \times n}, \ \mathbf{g}\mu \in \mathbb{R}^n$$

The tensor $n \times n \times n$ 3D grids, should be fine enough to resolve singularities in $\rho(x)$, $x \in \mathbb{R}^3$.

Calculation of the 3D Hartree potential in 1D complexity

[Khoromskij '08], **[Khoromskij, Khoromskaia '08 (SISC 2009)]**
Products of functions (orbitals squared) are substituted by Hadamard products of vectors

$$\rho \approx \Theta = \sum_{i=1}^{N_{orb}} \sum_{\mu=1}^{N_b} \sum_{\nu=1}^{N_b} C_{i,\nu} C_{i,\mu} (\mathbf{g}_\mu^{(1)} \odot \mathbf{g}_\nu^{(1)}) \otimes (\mathbf{g}_\mu^{(2)} \odot \mathbf{g}_\nu^{(2)}) \otimes (\mathbf{g}_\mu^{(3)} \odot \mathbf{g}_\nu^{(3)}).$$

Canonical-to-Tucker transform reduces the rank $N_b^2/2 \to R_\rho$ ($\sim$ from 10^4 $\downarrow$ to 10^2)

$$\Theta \to \Theta' := \sum_{t=1}^{R_\rho} \mathbf{u}_t^{(1)} \otimes \mathbf{u}_t^{(2)} \otimes \mathbf{u}_t^{(3)}.$$

Tensor product convolution (1D FFT) is calculated in $O(n \log n)$ complexity instead of required $O(n^3 \log n)$ operations for 3D FFT.

$$V_H \approx \mathbf{V}_H = \Theta' * \mathbf{P}_N = \sum_{t=1}^{R_\rho} \sum_{q=1}^{R_N} c_m b_k \left(\mathbf{u}_t^{(1)} * \mathbf{p}_q^{(1)}\right) \otimes \left(\mathbf{u}_t^{(2)} * \mathbf{p}_q^{(2)}\right) \otimes \left(\mathbf{u}_t^{(3)} * \mathbf{p}_q^{(3)}\right),$$

where Newton kernel is approximated by a canonical tensor of low rank,

$$\mathbf{P}_{\mathcal{N}} = \sum_{q=1}^{R_N} \mathbf{p}_q^{(1)} \otimes \mathbf{p}_q^{(2)} \otimes \mathbf{p}_q^{(3)}.$$

Tensor-product convolution vs. 3D FFT

Then projection of the 3D potential to 3D basis functions is computed by the scalar products

$$J_{\mu\nu} \approx \langle \mathbf{G}_\mu \odot \mathbf{G}_\nu, \mathbf{V}_H \rangle \quad \mu, \nu = 1, \dots N_b.$$

n^3	512^3	1024^3	2048^3	4096^3	8192^3	16384^3
FFT_3	582.8	~ 6000	–	–	–	~ 1 year
$C * C$	1.5	8.8	20.0	61.0	157.5	299.2
C2T	5.6	6.9	10.9	20.0	37.9	86.0

Table shows CPU time (in sec) for the computation of V_H for H_2O.
(3D FFT time for $n \geq 1024$ is obtained by extrapolation).

[Stenger], [Braess], [Gavrilyuk, Hackbusch, Khoromskij '08],
Canonical tensor for the Newton kernel: **[Bertoglio, Khoromskij '08]**
Tensor approximation of the Newton kernel using Laplace transform and and $(2M+1)$-term sinc-quadrature approximation

$$\frac{1}{\|x\|} = \int_0^\infty e^{-t^2\|x\|^2} dt \approx \sum_{k=-M}^{M} c_k e^{-t_k^2\|x\|^2} = \sum_{k=-M}^{M} c_k \prod_{\ell=1}^{3} e^{-t_k^2 x_\ell^2} \quad \mapsto \mathbf{P}_{\mathcal{N}}.$$

$\mathbf{P}_{\mathcal{N}} \in \mathbb{R}^{n\times n\times n}$, can. rank of $\mathbf{P}_{\mathcal{N}}$ $R_{\mathcal{N}} \leq 30$.

Laplacian in a Gaussian basis

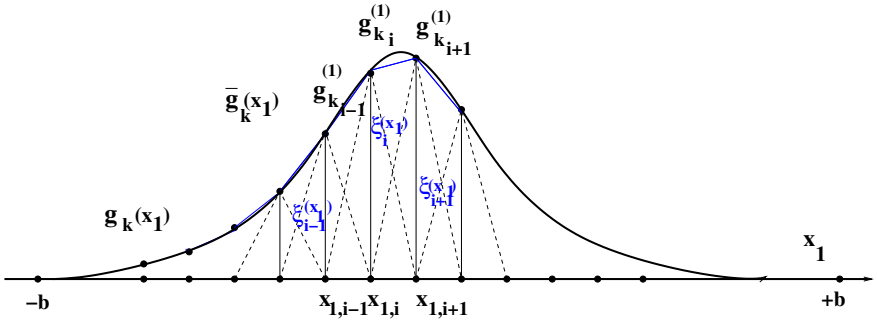
Core Hamiltonian part of the Fock operator

$$H_c = [-\frac{1}{2}\Delta - \sum_{\nu=1}^{M} \frac{Z_\nu}{\|x - a_\nu\|}$$

in a Gaussian separable basis,

$$h_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla g_\mu \cdot \nabla g_\nu dx + \int_{\mathbb{R}^3} V_c(x) g_\mu g_\nu dx \quad 1 \leq \mu, \nu \leq N_b.$$

For 3D grid-based tensor-structured calculation of the Laplace operator, we define a set of piecewise linear basis functions $\overline{g}_k := \mathbf{I}_1 g_k$, $k = 1, ..., N_b$, by linear tensor-product interpolation via the set of product hat functions, $\{\xi_{\mathbf{i}}\} = \xi_{i_1}(x_1)\xi_{i_2}(x_2)\xi_{i_3}(x_3)$, $\mathbf{i} \in \mathcal{I}$, associated with the respective grid-cells (voxels).



Tensor-based Laplacian

[Khoromskaia, Andrae, Khoromskij, CPC'12]
We approximate the exact Galerkin matrix $A_g \in \mathbb{R}^{N_b \times N_b}$,

$$A_g = \{a_{km}\} := \{\langle -\Delta_{(3)} g_k, g_m \rangle\} \equiv \{\langle \nabla_{(3)} g_k, \nabla_{(3)} g_m \rangle\}, \quad k, m = 1, \dots N_b,$$

using the piecewise linear representation of the basis functions, $\overline{g}_k(x) \in \mathbb{R}^3$ constructed on $N \times N \times N$ Cartesian grid. Here $\nabla_{(3)}$ denotes the 3D gradient operator. The approximating matrix A_G is now defined by

$$A_g \approx A_G = \{\overline{a}_{km}\} := \{\langle -\Delta_{(3)} \overline{g}_k, \overline{g}_m \rangle\} \equiv \{\langle \nabla_{(3)} \overline{g}_k, \nabla_{(3)} \overline{g}_m \rangle\}, \quad A_G \in \mathbb{R}^{N_b \times N_b}.$$

The Laplace operator applies to a separable function $\eta(x), x = (x_1, x_2, x_3) \in \mathbb{R}^3$, $\eta(x) = \eta_1(x_1)\eta_2(x_2)\eta_3(x_3)$, as follows

$$\Delta_{(3)}\eta(x) = \frac{d^2\eta_1(x_1)}{dx_1^2}\eta_2(x_2)\eta_3(x_3) + \eta_1(x_1)\frac{d^2\eta_2(x_2)}{dx_2^2}\eta_3(x_3) + \eta_1(x_1)\eta_2(x_2)\frac{d^2\eta_3(x_3)}{dx_3^2},$$

corresponding to the rank-3 tensor representation of the respective Galerkin stiffness matrix A_3 in a tensor basis $\{\xi_i(x_1)\xi_j(x_2)\xi_k(x_3)\}$, $i, j, k = 1, \dots N$,

$$A_3 := A^{(1)} \otimes S^{(2)} \otimes S^{(3)} + S^{(1)} \otimes A^{(2)} \otimes S^{(3)} + S^{(1)} \otimes S^{(2)} \otimes A^{(3)} \in \mathbb{R}^{N^{\otimes 3} \times N^{\otimes 3}}.$$

Tensor-based Laplacian

[Khoromskaia, Andrae, Khoromskij, CPC'12]
Here the 1D stiffnesss and mass matrices $A^{(\ell)}, S^{(\ell)} \in \mathbb{R}^{N \times N}$, $\ell = 1, 2, 3$, are given by $A^{(\ell)} := \{\langle \nabla_{(1)} \xi_i(x_\ell), \nabla_{(1)} \xi_j(x_\ell) \rangle\}_{i,j=1}^N = \frac{1}{h} tridiag\{-1, 2, -1\},$,
$S^{(\ell)} = \{\langle \xi_i, \xi_j \rangle\}_{i,j=1}^N = \frac{h}{6}\ tridiag\{1, 4, 1\}$, respectively, and $\nabla_{(1)} = \frac{d}{dx_\ell}$.
Since $\{\xi_i\}_{i=1}^N$ are the same for all modes $\ell = 1, 2, 3$, we further assume, $A^{(\ell)} = A_1$, and $S^{(\ell)} = S_1$.

Lemma

Assume that the basis functions $\{\overline{g}_k(x)\}$, $x \in \mathbb{R}^3$, $k = 1, \dots N_b$, are rank-1 separable, i.e., $\overline{g}_k(x) = \overline{g}_k^{(1)}(x_1)\overline{g}_k^{(2)}(x_2)\overline{g}_k^{(3)}(x_3)$. Then the matrix entries of A_G have the representation,

$$\begin{aligned} \overline{a}_{km} &= \langle A_1 \mathbf{g}_k^{(1)}, \mathbf{g}_m^{(1)} \rangle \langle S_1 \mathbf{g}_k^{(2)}, \mathbf{g}_m^{(2)} \rangle \langle S_1 \mathbf{g}_k^{(3)}, \mathbf{g}_m^{(3)} \rangle \\ &\quad + \langle S_1 \mathbf{g}_k^{(1)}, \mathbf{g}_m^{(1)} \rangle \langle A_1 \mathbf{g}_k^{(2)}, \mathbf{g}_m^{(2)} \rangle \langle S_1 \mathbf{g}_k^{(3)}, \mathbf{g}_m^{(3)} \rangle \\ &\quad + \langle S_1 \mathbf{g}_k^{(1)}, \mathbf{g}_m^{(1)} \rangle \langle S_1 \mathbf{g}_k^{(2)}, \mathbf{g}_m^{(2)} \rangle \langle A_1 \mathbf{g}_k^{(3)}, \mathbf{g}_m^{(3)} \rangle \\ &= \langle A_3 \mathbf{G}_k, \mathbf{G}_m \rangle, \end{aligned} \tag{3}$$

where $\mathbf{g}_k^{(\ell)}$, $\mathbf{g}_m^{(\ell)} \in \mathbb{R}^N$ $(k, m = 1, \dots, N_b)$, are the vectors of collocation coefficients of $\{\overline{g}_k^{(\ell)}(x_\ell)\}, \ell = 1, 2, 3$, and $\mathbf{G}_k$ are the corresponding rank-1 3-tensors $\mathbf{G}_k = \mathbf{g}_k^{(1)} \otimes \mathbf{g}_k^{(2)} \otimes \mathbf{g}_k^{(3)}$.

Tensor-based Laplacian

Representation for rank-1 tensor Laplace can be simplified by the standard lumping procedure preserving the same approximation error $O(h^2)$,

$$\overline{a}_{km} = \langle A_{3,FD}\mathbf{G}_k, \mathbf{G}_m \rangle,$$

where $A_{3,FD}$ denotes the finite difference (FD) discrete Laplacian,

$$A_{3,FD} := A^{(1)} \otimes I^{(2)} \otimes I^{(3)} + I^{(1)} \otimes A^{(2)} \otimes I^{(3)} + I^{(1)} \otimes I^{(2)} \otimes A^{(3)},$$

with $I^{(\ell)}$ being the $n \times n$ identity matrix and

$$A^{(\ell)} := \{\langle \nabla_{(1)}\xi_i(x_\ell), \nabla_{(1)}\xi_j(x_\ell)\rangle\}_{i,j=1}^{N} = \frac{1}{h} tridiag\{-1, 2, -1\},$$

is a 1D stiffnesss mass matrix, $\ell = 1,\ 2,\ 3$.

Nuclear potential operator

[Khoromskaia, '13, CMAM]

$$V_c(x) = -\sum_{\alpha=1}^{M} \frac{Z_\alpha}{\|x - a_\alpha\|}, \quad Z_\alpha > 0, \quad x, a_\alpha \in \mathbb{R}^3$$

Using the canonical tensor representation of the Newton potential

$$\widetilde{\mathbf{P}}_R = \sum_{q=1}^{R} \mathbf{p}_q^{(1)} \otimes \mathbf{p}_q^{(2)} \otimes \mathbf{p}_q^{(3)} \in \mathbb{R}^{2n\times 2n\times 2n},$$

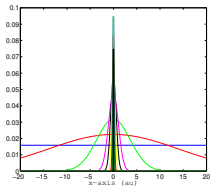
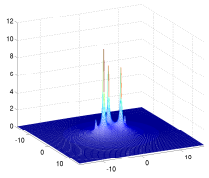
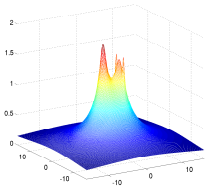
and shifting/windowing operator

$$\mathcal{W}_\alpha = \mathcal{W}_\alpha^{(1)} \otimes \mathcal{W}_\alpha^{(2)} \otimes \mathcal{W}_\alpha^{(3)}.$$

$$\mathbf{P}_c = \sum_{\alpha=1}^{M} Z_\alpha \mathcal{W}_\alpha \widetilde{\mathbf{P}}_R = \sum_{\alpha=1}^{M} Z_\alpha \sum_{q=1}^{R} \mathcal{W}_\alpha^{(1)}\mathbf{p}_q^{(1)} \otimes \mathcal{W}_\alpha^{(2)}\mathbf{p}_q^{(2)} \otimes \mathcal{W}_\alpha^{(3)}\mathbf{p}_q^{(3)} \in \mathbb{R}^{n\times n\times n},$$

$$\overline{v}_{km} = \int_{\mathbb{R}^3} V_c(x)\overline{g}_k(x)\overline{g}_m(x)dx \approx \langle \mathbf{G}_k \odot \mathbf{G}_m, \mathbf{P}_c \rangle, \quad 1 \le k, m \le N_b.$$

Tensor-structured core Hamiltonian

V_c for ethanol molecule (C_2H_5OH) at levels: $z = 0$ and $z = 0.75$ au, $n^3 = 10^{15}$, mesh-size $h \sim 2.2 \cdot 10^{-4}$ au $= 1.164 \cdot 10^{-4}$ Å $= 11.64$ fm (10^{-15} m) $\sim$ size of atomic radii.

p	13	14	15	16	17
$N^3 = 2^{3p}$	8192^3	16384^3	32768^3	65536^3	131072^3
$Er(A_G)$	0.032	0.0083	0.0021	$5.2 \cdot 10^{-4}$	$1.3 \cdot 10^{-4}$
$\lvert a_{11} - \bar{a}_{11} \rvert$	208	52	13	3.3	0.82
$Er(V_G)$	0.024	0.0083	0.0011	$3.1 \cdot 10^{-4}$	
$\lvert v_{11} - \bar{v}_{11} \rvert$	14	5.4	0.8	0.3	

Ethanol (C_2H_5OH): abs. accuracy of the Galerkin matrices corresponding to the Laplace, $Er(A_G)$, and the nuclear potential operators, $Er(V_G)$, using the discretized basis of 123 Gaussians (abs. values of a_{11} and v_{11} are of the order of 10^4 and 10^3, respectively).

Grid-based two-electron integrals (TEI)

[Khoromskaia, Khoromskij, Schneider, '12]

$$b_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{g_\mu(x)g_\nu(x)g_\kappa(y)g_\lambda(y)}{\|x-y\|}dxdy = \langle \mathbf{G}_\mu \odot \mathbf{G}_\nu, \mathbf{P}_{\mathcal{N}} * (\mathbf{G}_\kappa \odot \mathbf{G}_\lambda)\rangle_{n^{\otimes 3}},$$

$$\mathbf{G}_\mu = \mathbf{g}_\mu^{(1)} \otimes \mathbf{g}_\mu^{(2)} \otimes \mathbf{g}_\mu^{(3)} \in \mathbb{R}^{n\times n\times n}.$$

$$G^{(\ell)} = \left[\mathbf{g}_\mu^{(\ell)} \odot \mathbf{g}_\nu^{(\ell)}\right]_{1\le\mu,\nu\le N_b} \in \mathbb{R}^{n\times N_b^2} \quad \ell = 1,2,3.$$

Factorization ("1D density fitting") by Cholesky decomposition of $G^{(\ell)}G^{(\ell)\,T}$:

$$G^{(\ell)} \cong U^{(\ell)}V^{(\ell)\,T}, \quad G^{(\ell)} \in \mathbb{R}^{n\times N_b^2}, U^{(\ell)} \in \mathbb{R}^{n\times R_\ell}, \quad V^{(\ell)} \in \mathbb{R}^{N_b^2\times R_\ell},$$

$\Rightarrow$ convolutions are reduced N_b^2 to $R_\ell \sim N_b$, (131072×40000)

Grid-based two-electron integrals (TEI)

[Khoromskaia, Khoromskij, Schneider, '12]

SVD: $G^{(\ell)} \cong U^{(\ell)} V^{(\ell)\,T}$, $\ell = 1, 2, 3$, with $U^{(\ell)} \in \mathbb{R}^{n \times R_\ell}$ and $V^{(\ell)} \in \mathbb{R}^{N_b^2 \times R_\ell}$

The Newton kernel: $P^{(\ell)} \in \mathbb{R}^{n \times R_\mathcal{N}}$ are the factor matrices in the rank-$R_\mathcal{N}$ canonical tensor $\mathbf{P}_\mathcal{N} \in \mathbb{R}^{n \times n \times n}$.

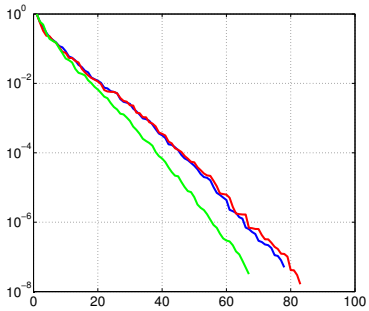
$$B \cong B_\varepsilon := \sum_{k=1}^{R_\mathcal{N}} \odot_{\ell=1}^{3} V^{(\ell)} M_k^{(\ell)} V^{(\ell)\,T},$$

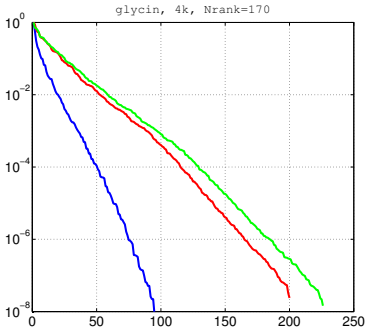
(storage: $R_\mathcal{N} \sum_{\ell=1}^{3} R_\ell^2 + N_b^2 \sum_{\ell=1}^{3} R_\ell$), with the Galerkin convolution matrix

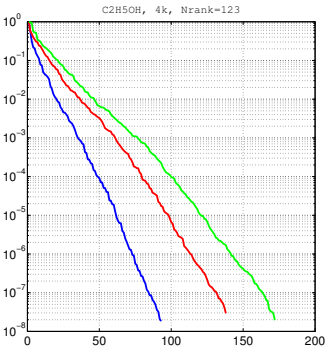
$$M_k^{(\ell)} = U^{(\ell)\,T} (P_k^{(\ell)} *_n U^{(\ell)}) \in \mathbb{R}^{R_\ell \times R_\ell}, \quad k = 1, \ldots, R_\mathcal{N}.$$

storage: $O((R_G + R_\mathcal{N})n)$, complexity: $O(R_\mathcal{N} R_G^2 n + R_G R_\mathcal{N} n \log n)$.

Factorization of two-electron integrals (TEI)







ε-ranks R_ℓ: NH3, glycin, C_2H_5OH.

Factorization of two-electron integrals (TEI)

The column and diagonal elements in the TEI matrix B :

$B(:,j^*) = \sum_{k=1}^{R_\mathcal{N}} \odot_{\ell=1}^{3} V^{(\ell)} M_k^{(\ell)} V^{(\ell)}(j^*)^T$, $B(i,i) = \sum_{k=1}^{R_\mathcal{N}} \odot_{\ell=1}^{3} V^{(\ell)}(i) M_k^{(\ell)} V^{(\ell)}(i)^T$,

with $i = vec(\mu,\nu) := (\mu-1)N_b + \nu, \quad j = vec(\kappa,\lambda), \quad i,j \in I_N := \{1,...,N\}$.

Cholesky decomposition (ε-approximation)

$$B := mat(\mathbf{B}) = [b_{\mu\nu,\kappa\lambda}] \approx LL^T.$$

▷ Coulomb matrix: given $\overline{D} = vec(D)$,

$$vec(J) = B\overline{D} \approx L(L^T\overline{D}).$$

▷ HF exchange: employing the permuted tensor $\widetilde{\mathbf{B}} = permute(\mathbf{B}, [2,3,1,4])$,

$$vec(K) = \widetilde{B}\overline{D}, \quad \widetilde{B} = mat(\widetilde{\mathbf{B}}).$$

Representation complexity of $\mathbf{B}$ using the quantized tensor format can be reduced to $O(N_b N_{orb}^2)$ (instead of $O(N_b^3)$). ($N_b \sim 10 N_{orb}$).

Fast convolution via tensor approximation of Green's kernel

Tensor approximation of the Newton kernel using Laplace transform and sinc-quadratures:
[Gavrilyuk, Hackbusch, Khoromskij '08]
[Bertoglio, Khoromskij '10]
Green's function for Δ in $\mathbb{R}^3$, via $(2M+1)$-term sinc-quadrature approximation

$$\frac{1}{\|x\|} = \int_0^\infty e^{-t^2\|x\|^2} dt \approx \sum_{k=-M}^{M} c_k e^{-t_k^2\|x\|^2} = \sum_{k=-M}^{M} c_k \prod_{\ell=1}^{3} e^{-t_k^2 x_\ell^2} \quad \mapsto \mathbf{P}_\mathcal{N}.$$

$\mathbf{P}_\mathcal{N} \in \mathbb{R}^{n\times n\times n}$, can. rank of $\mathbf{P}_\mathcal{N}$ $R_\mathcal{N} \le 30$.

Tensor-product convolution, $O(n \log n)$:
[Khoromskij '08], **[Khoromskij, Khoromskaia '09]**

$$U * P_N = \sum_{k=1}^{R_F} \sum_{m=1}^{R_N} c_k b_m (u_m^{(1)} * p_k^{(1)}) \otimes (u_m^{(2)} * p_k^{(2)}) \otimes (u_m^{(3)} * p_k^{(3)})$$

n^3	512^3	1024^3	2048^3	4096^3	8192^3	16384^3	32768^3
FFT_3	37.5	350.6	~ 3500	–	–	–	~ 1.2 years
$C_{R_F} * C_{R_\mathcal{N}}$	2.4	6.7	14.6	44	107	236	535

CPU time (in sec) for TEI: $\frac{1}{\|x\|} * g_\mu g_\nu$, $\mu,\nu = 1,...,N_b$, $\varepsilon = 10^{-7}$,

H_2O, $N_b = 41$, $\frac{N_b(N_b+1)}{2} \mapsto R_F = 71$, $R_\mathcal{N} = 27$.

Self-consistent iteration for nonlinear EVP

[Khoromskaia, CMAM'13]
[Khoromskaia, Khoromskij CPC'13]
EVP algorithm for black-box solver:

$$F(C)C = SC\Lambda, \quad F = H_0 + J(C) - K(C),$$

- Initial guess for $J = 0$, $K = 0$, $F(0) = H_0$.
- solve EVP $[H_0 + J(C) - K(C)]C = SC\Lambda$.
- Update of $J(C)$ and $K(C)$:
  ▷ Coulomb matrix: given $\overline{D} = vec(D)$,

$$vec(J) = B\overline{D} \approx L(L^T\overline{D}).$$

  ▷ HF exchange: using $D = 2CC^T$ and $B = LL^T$,

$$K(D)_{\mu\nu} = -\sum_{i=1}^{N_{orb}}\sum_{k=1}^{R_B}(\sum_{\lambda} L_{\mu\lambda k}C_{\lambda i})(\sum_{\kappa} C_{\kappa i}L_{\kappa\nu k}),$$

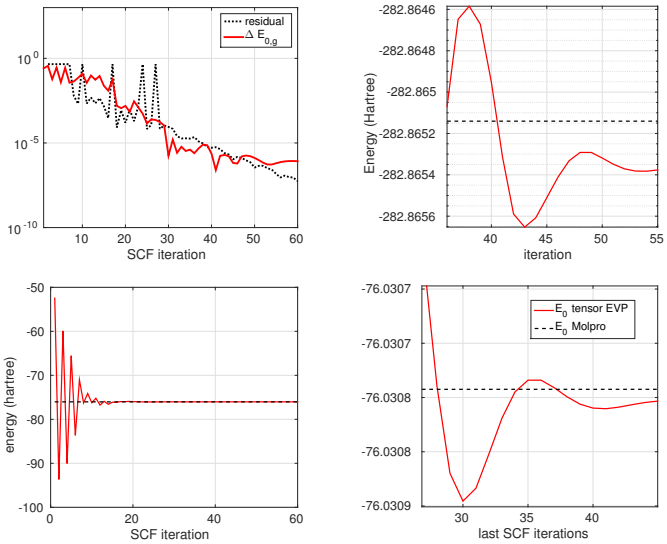
  $[L_{\mu\nu k}] = reshape(L, [N_b, N_b, R_B])$ is the $N_b \times N_b \times R_B$-folding of the Cholesky fact. L.
- DIIS for providing convergence.

Tensor-based Hartree-Fock solver

[Khoromskaia, CMAM'13], [Khoromskaia, Khoromskij PCCP'15]

Tensor-based Electronic Structure Calculations (TESC) package
[Khoromskaia, Khoromskij, 2008-2014]

SCF DIIS iteration for amino acids glycine (C_2 H_5N O_2) with TEI on $n^3 = 131072^3$, $N_b^2 = 28000$



H_2O: convergence in energy; last $k + 27$ iterations.