Tensor methods for fast multidimensional integration in the Hartree-Fock calculations

Boris Khoromskij
(joint with Venera Khoromskaia)

Dedicated to Prof. Martin Costabel on occasion of his 65th birthday

Workshop on Singularities and Integral Equations, Uni. Rennes, 26-30.08.2013

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

Outline of the talk

1. Introduction: motivations, advances, applications
2. Separation of variables: from canonical to tensor network formats
3. Quantized tensor approximation of logarithmic complexity
   - Multi-resolution folding to quantized image: $\mathbb{R}^{2^L} \mapsto \bigotimes_{\ell=1}^{L} \mathbb{R}^2$
   - Theory of quantized approximation to multidimensional functional vectors
   - Examples on fast $O(\log N)$-integration
4. Hartree-Fock eigenvalue problem for 3D integral-differential operator
   - Challenges of the grid-based Hartree-Fock calculations
   - Fast tensor convolution with 3D Newton kernel
   - Tensor Factorization of the Two-electron Integrals
   - Tensor decompositions in post-HF calculations: Møller-Plesset (MP2) energy correction
   - Towards grid-based calculations of extended and periodic systems
5. Superfast QTT-FFT, QTT-convolution and QTT-FWT in $O(\log N)$ operations
6. Conclusions

Main focus: $O(d)$ - numerical approximation to d-dimensional PDEs

Basic ingredients:

- Traditional numerical methods.
- Numerical multilinear algebra
- Low-parametric separable approximation of d-variate functions: theory/algorithms.
- Tensor representation of linear operators: Green's functions, integration, convolution(d), FFT(d), wavelet(d), multi-particle Hamiltonians, preconditioners.
- Iterative solvers to steady-state and temporal PDEs on "tensor manifolds".

"Separation" of variables beats "curse of dimensionality":

- $O(dN)$ tensor numerical methods, $N^d \to O(dN)$.

Super-compression:

- $O(d \log N)$ Quantized tensor approximation (QC, QTT), $N^d \to O(d \log N)$.

Guiding principle:

- Validation of numerical algorithms on real-life multi-dimensional PDEs.

Separable representation of (discrete) functions in a tensor-product Hilbert space

Tensor-product Hilbert space, $\mathbb{V}_{\mathbf{n}} = V_1 \otimes ... \otimes V_d$, $\mathbf{n} = (n_1, ..., n_d)$, $n_\ell = \dim V_\ell$.

- Euclidean vector space $\mathbb{V}_{\mathbf{n}} = \mathbb{R}^{n_1 \times ... \times n_d}$, $V_\ell = \mathbb{R}^{n_\ell}$ $(\ell = 1, ..., d)$,

$$\mathbf{V} = [v_{\mathbf{i}}] \in \mathbb{V}_{\mathbf{n}} : \quad \langle \mathbf{W}, \mathbf{V} \rangle = \sum\nolimits_{\mathbf{i}} w_{\mathbf{i}} v_{\mathbf{i}}, \quad \mathbf{i} = (i_1, ..., i_d) : i_\ell \in I_\ell = \{1, ..., n_\ell\}.$$

- Tensors are functions of discrete variable, $\mathbb{V}_{\mathbf{n}} \ni \mathbf{V} : I_1 \times ... \times I_d \mapsto \mathbb{R}$.

Separable representation in $\mathbb{V}_{\mathbf{n}}$: rank-1 tensors

$$\mathbf{V} = [v_{i_1 ... i_d}] = v^{(1)} \otimes \ldots \otimes v^{(d)} \in \mathbb{V}_{\mathbf{n}}, \quad v_{i_1 ... i_d} = \prod\nolimits_{\ell=1}^{d} v^{(\ell)}_{i_\ell} :$$

- The scalar product

$$\langle \mathbf{W}, \mathbf{V} \rangle = \langle w^{(1)} \otimes \ldots \otimes w^{(d)}, v^{(1)} \otimes \ldots \otimes v^{(d)} \rangle = \prod\nolimits_{\ell=1}^{d} \langle w^{(\ell)}, v^{(\ell)} \rangle_{V_\ell}.$$

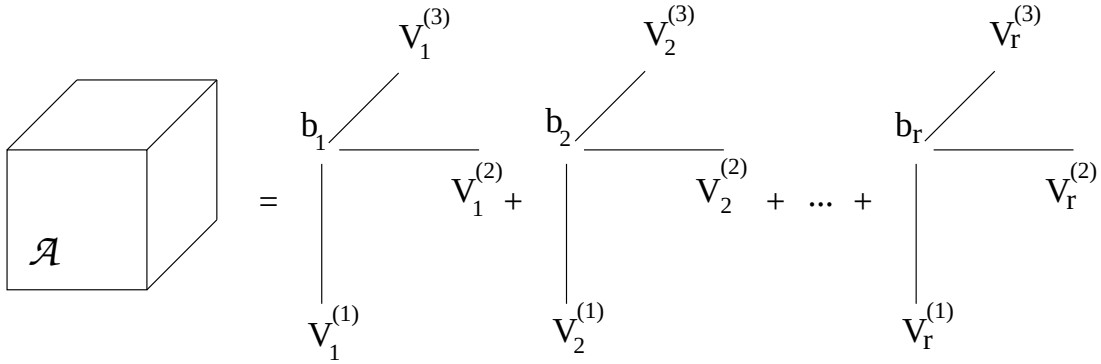
- Storage: $Stor(\mathbf{V}) = \sum_{\ell=1}^{d} n_\ell \ll \dim \mathbb{V}_{\mathbf{n}} = \prod_{\ell=1}^{d} n_\ell$.
- $O(d)$ bilinear operations: addition, Hadamard product, contraction, convolution, ...

Def. **Canonical R-term representation in $\mathbb{V}_{\mathbf{n}}$:** $\mathbf{V} \in \mathcal{C}_R(\mathbb{V}_{\mathbf{n}})$, if [Hitchcock '27, ...]

$$\mathbf{V} = \sum_{k=1}^{R} v_k^{(1)} \otimes \ldots \otimes v_k^{(d)}, \quad v_k^{(\ell)} \in V_\ell.$$

- $d = 2$: rank-R matrices, $V = \sum_{k=1}^{R} u_k v_k^T$.

Visualizing canonical model, $d = 3$.



- Advantages: Storage $= dRN$, simple multilinear algebra.
- Limitations: $\mathcal{C}_R(\mathbb{V}_{\mathbf{n}})$ is the non-closed set $\Rightarrow$ lack of stable approximation methods.

Example. $f(x) = x_1 + \ldots + x_d$. $rank_{Can}(f) = d$, but approximated by rank-2 elements

$$f(x) = \lim_{\varepsilon \to 0} \frac{\prod_{\ell=1}^{d}(1 + \varepsilon x_\ell) - 1}{\varepsilon}.$$

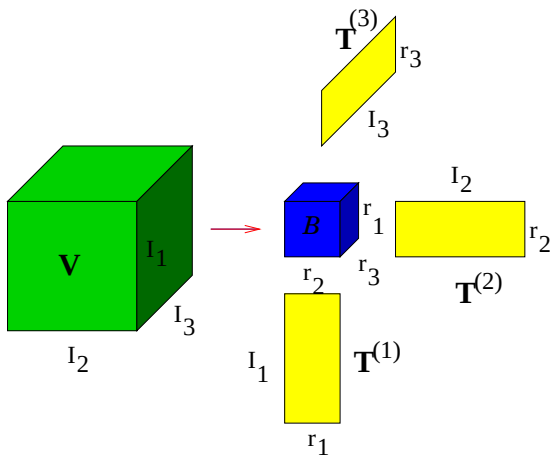
Orthogonal Tucker model

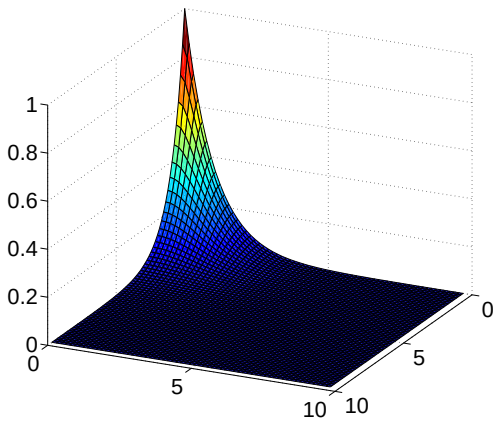
Def. **Rank $\mathbf{r} = [r_1, \ldots, r_d]$ Tucker tensors:** $\mathbf{V} \in \mathcal{T}_{\mathbf{r}}(\mathbb{V}_{\mathbf{n}})$ if [Tucker '66; De Lathawer et. al. '00, ...]

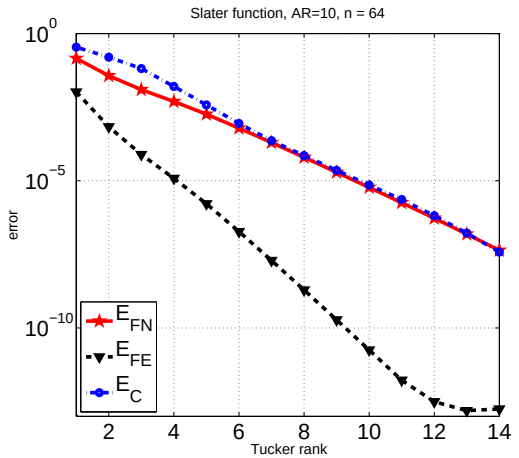
$$\mathbf{V} = \sum_{k_1, \ldots, k_d = 1}^{\mathbf{r}} b_{k_1 \ldots k_d} v_{k_1}^{(1)} \otimes \ldots \otimes v_{k_d}^{(d)} \in T_1 \otimes \ldots \otimes T_d, \quad T_\ell = span\{v_{k_\ell}^{(\ell)}\}_{k_\ell = 1}^{r_\ell} \subset \mathbb{R}^{n_\ell}.$$

- $d = 2$: SVD of a rank-r matrix, $A = UDV^T$, $U \in \mathbb{R}^{n \times r}$, $D \in \mathbb{R}^{r \times r}$, [Schmidt '1905]
- Storage: $drN + r^d$, $r = \max r_\ell \ll N$ (efficient for $d = 3$, e.g. Hartree-Fock eq.).

Beginning of tensor numerical methods: Tucker for 3D functions (e.g. $f = e^{-\mathbf{r}}, \frac{1}{\mathbf{r}}$).
[BNK, Khoromskaia '07]







Matrix Product States (MPS) factorization:

In quantum physics/chemistry:
The matrix product states (**MPS**) and tree-tensor network states (**TNS**) of slightly entangled systems, matrix product operators (**MPO**), **DMRG** optimization.
[White '92; ..., Östlund, Rommer '95; ..., Cirac, Verstraete '06, ...].

Re-invented in numerical multilinear algebra:
Hierarchical splitting of dimensions (tree structure): [BNK '06].
Hierarchical Tucker (HT) $\equiv$ TNS (without loops): [Hackbusch, Kühn '09].
Tensor train (TT) $\equiv$ MPS (open boundary cond.) [Oseledets, Tyrtyshnikov '09].

Def. MPS (TT): Given $\mathbf{r} = (r_1, ..., r_d)$, $r_d = 1$, $r_0 = 1$.
$\mathbf{V} \in TT[\mathbf{r}] \subset \mathbb{V}_{\mathbf{n}}$ is a parametrization by contracted product of tri-tensors in $\mathbb{R}^{r_{\ell-1} \times n_\ell \times r_\ell}$,

$$\begin{aligned} \mathbf{V}[i_1 ... i_d] &= \sum\nolimits_{\alpha} G^{(1)}_{\alpha_1}[i_1] G^{(2)}_{\alpha_1 \alpha_2}[i_2] \cdots G^{(d)}_{\alpha_{d-1}}[i_d] \\ &\equiv G^{(1)}[i_1] G^{(2)}[i_2] ... G^{(d)}[i_d], \end{aligned}$$

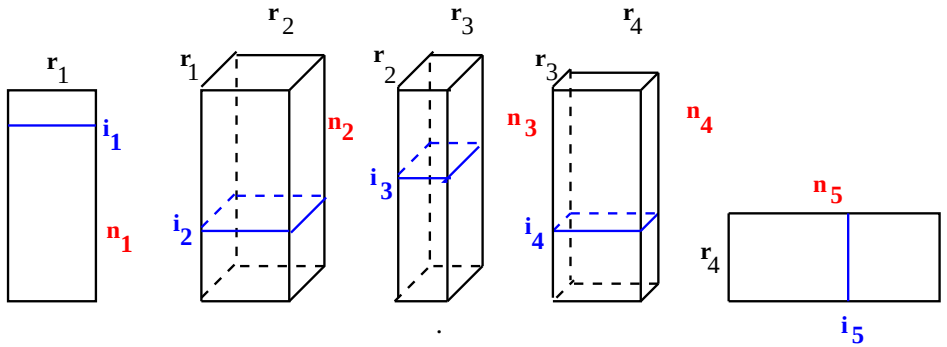
$G^{(\ell)}[i_\ell]$ is a $r_{\ell-1} \times r_\ell$ matrix, $1 \le i_\ell \le n_\ell$.

Example. $f(x) = x_1 + ... + x_d$, $rank_{TT}(f) = 2$.

$$f = \begin{bmatrix} x_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ x_2 & 1 \end{bmatrix} ... \begin{bmatrix} 1 & 0 \\ x_{d-1} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_d \end{bmatrix}.$$

Benefits and limitations of the MPS (TT) format

Example. $d = 5$.



- **Advantages**: Storage: $dr^2N \ll N^d$, $N = \max n_k$.

Efficient and robust MLA with polynomial scaling in r, linear scaling in d.
Can be implemented by stable QR/SVD algorithms.

- **Limitations**: strong entanglements in a system, large mode-size N.

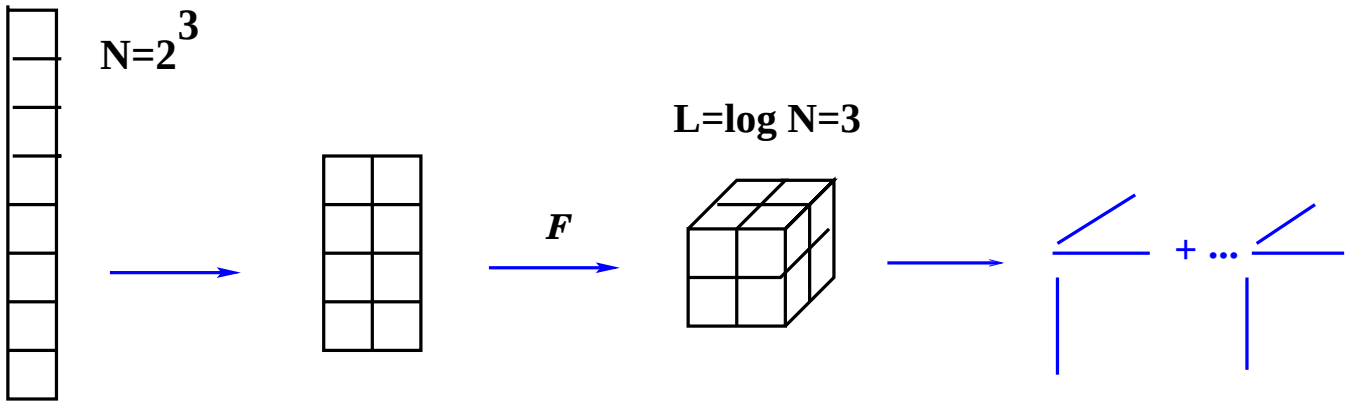
Multilinear matrix-vector algebra and DMRG iterations cost: $O(dRr^3N^2)$

$$d, R, r \sim 10^2,\ N \sim 10^3 \div 10^4 - \textbf{non-tractable local problems ?}$$

Rank bounds:

$$r_{TT} \leq R_{Can}, \quad r_{Tuck} \leq {r_{TT}}^2, \quad r_{Tuck} \leq R_{Can}.$$

- Quantized Tensor Approximation of N-vectors with $N = 2^L$. [BNK '2009]



$$(\textit{isometry}) \quad \mathcal{F}_L : [x_i]_{i=1}^N = \mathbf{X} \to \mathbf{A} = [a_{\mathbf{j}}] \in \mathbb{Q}_L := \bigotimes_{\ell=1}^{L} \mathbb{R}^2, \quad a_{\mathbf{j}} := x_i.$$

$$i - 1 = \sum\nolimits_{\nu=1}^{L} (j_\nu - 1) 2^{\nu-1}, \quad \mathbf{j} - \mathbf{1} \in \{0,1\}^{\otimes L}, \text{binary coding}$$

Canonical/TT approximation of quantized image in $\mathbb{Q}_L \Rightarrow$ QCan/QTT approx. method

- Storage in quantized tensor formats scales logarithmically in $N = 2^L$,

$$2r^2 L \ll 2^L.$$

- $N = q^L$, $q = 2, 3, \ldots$: $q_{opt} = e \approx 2,7\ldots$. Standard choice $q = 2$: binary coding.

Thm. [BNK '09]. *QTT*-approximation of functional vectors. Let $N = 2^L$.

- For quantized exponential N-vector: $rank_{QCan}(\mathbf{X}) = rank_{Can}(Q_{1,L}(\mathbf{X})) = 1$,

$$\mathbf{X} := \{z^{n-1}\}_{n=1}^{N} \in \mathbb{C}^N \mapsto \otimes_{p=1}^{L} \begin{bmatrix} 1 \\ z^{2^{p-1}} \end{bmatrix} \in \bigotimes_{p=1}^{L} \mathbb{C}^2, \quad z \in \mathbb{C}.$$

- For the quantized trigonometric N-vector: $rank_{QCan,\mathbb{C}}(\mathbf{X}) = rank_{QTT,\mathbb{R}}(\mathbf{X}) = 2$,

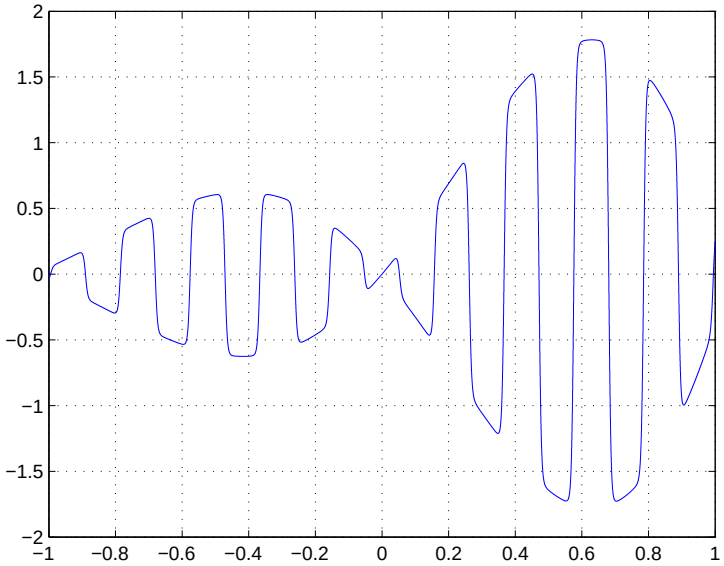
$$\mathbf{X} := \{\sin(\omega h(n-1))\}_{n=1}^{N} \in \mathbb{C}^N, \quad h = \frac{1}{N-1}, \quad \forall \omega \in \mathbb{C}.$$

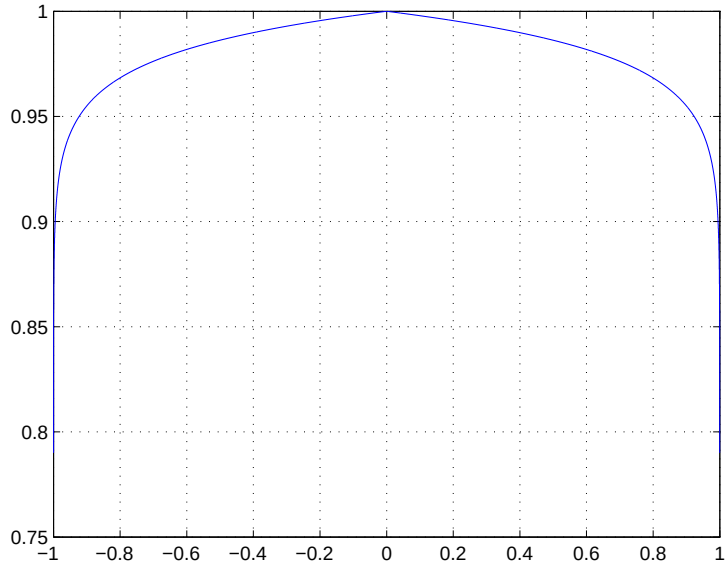
- Proof. Hint: $\sin z = \frac{e^{iz} - e^{-iz}}{2i} = Im(e^{iz})$.
- For QTT-image of polynomial of degree m, $rank_{QTT}(P_m) \le m + 1$.
- QTT-rank of the step function and Haar wavelet is 1 and 2, resp.
- Chebyshev polynomial $T_m(x) = \cos(m \arccos x)$, sampled as a vector

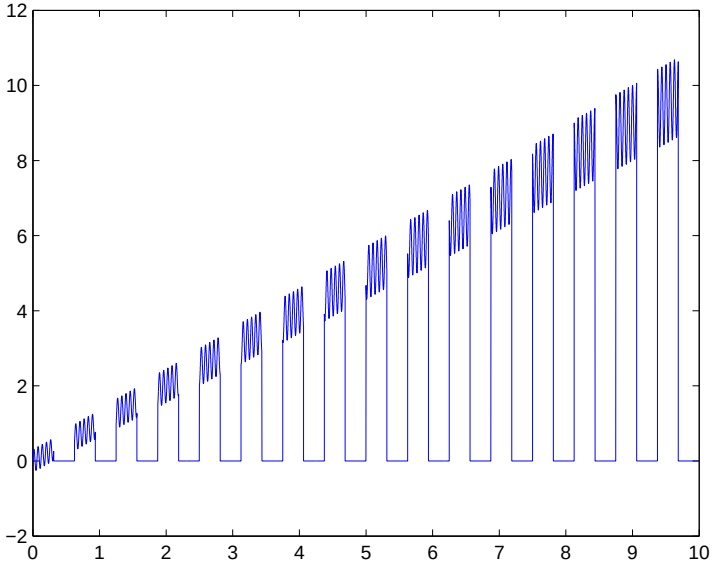
$$\mathbf{X} := \{x_n := T_m(x_n)\}_{n=0}^{N} \in \mathbb{R}^N, \quad N = 2^L - 1, \ |x_n| \le 1,$$

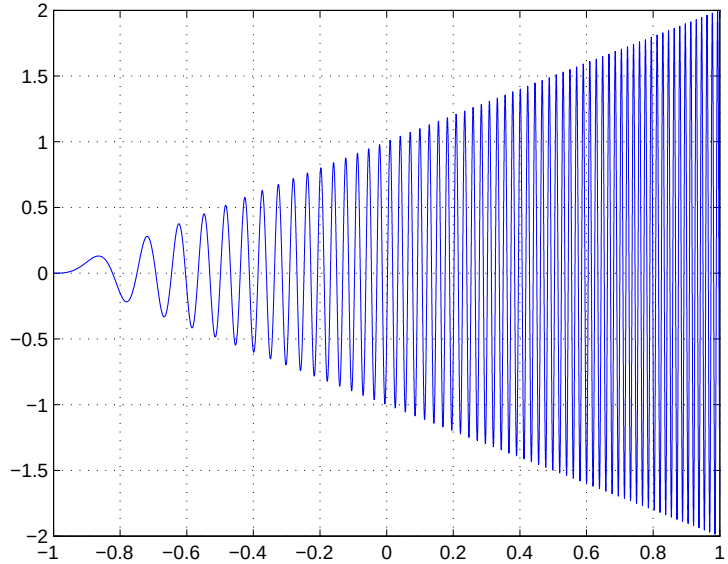
over CGL nodes $\{x_n = \cos \frac{\pi n}{N}\}$, has the explicit rank-2 QCan-image.

QTT based quadratures (cf. Chebfun2, L.-N. Trefethen, et al. '13)









QTT based quadratures of $O(\log n)$ complexity

Quantized weight function $w(x)$, integrand $f(x)$, both with moderate QTT-ranks.

The rectngular n-point quadrature, $n = 2^L$, $|I - I_n| = O(2^{-\alpha L})$, $Time = O(\log n)$.

$$\int_{-1}^{1} w(x)f(x)dx \approx I_n(f) := h\sum_{i=1}^{n} w(x_i)f(x_i) = \langle \mathbf{W}, \mathbf{F} \rangle_{QTT}, \quad \mathbf{W}, \mathbf{F} \in \otimes_{\ell=1}^{L} \mathbb{R}^2.$$

Examples. Highly oscillated and singular functions on $[-1, 1]$, $\varepsilon_{QTT} = 10^{-6}$:

$f_1(x) = e^x \sin(3x)\tanh(5\cos(30x))$, (N. Hale, L.-N. Trefethen, '12)

$f_2(x) = (1 - |x|)^q$, $q = 0.025$.

$f_3(x) =$ (homogenization example: **3 scales**).

$f_4(x) = (x+1)\sin(\omega(x+1)^2)$, $\omega = 100$ (Fresnel integral).

$n \setminus \overline{r}$	$r_{QTT}(f_1)$	$r_{QTT}(f_2)$	$r_{QTT}(f_3)$	$r_{QTT}(f_4)$
2^{14}	7.0	4.0	3.5	6.5
2^{15}	7.0	4.0	3.6	7.0
2^{16}	8.5	4.5	3.6	7.5
2^{17}	9.0	5.0	3.6	7.9

The Hartree-Fock equation

$$\mathcal{F}\varphi_i(x) \equiv (-\frac{1}{2}\Delta + V_c + V_H - \mathcal{K})\varphi_i(x) = \lambda_i \varphi_i(x), \qquad i = 1, ..., N_{orb}.$$

The Fock operator $\mathcal{F}$ depends on $\tau(x,y) = 2\sum_{i=1}^{N_{orb}} \varphi_i(x)\varphi_i(y)$,

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

The efficient Galerkin representation of the nonlinear Fock operator in low-rank basis $\{g_\mu\}$ is based on the precomputed two-electron integrals (TEI) tensor:

$$b_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3\times\mathbb{R}^3} \frac{g_\mu(x)g_\nu(x)g_\kappa(y)g_\lambda(y)}{\|x-y\|} dxdy, \quad 1 \le \mu, \nu, \kappa, \lambda \le N_b.$$

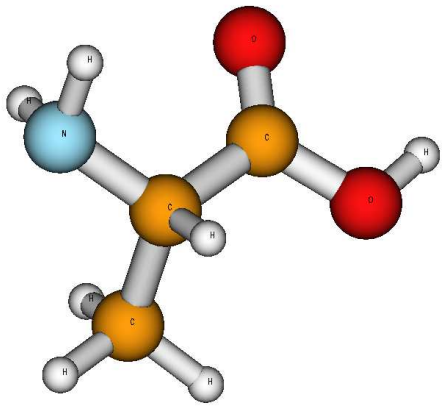
Complexity scaling $N_b^4 \times$ (3D convolution cost).

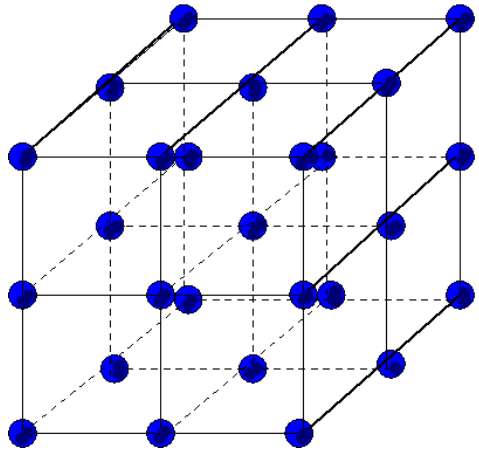
Challenges: High accuracy, 3D singular convolutions, nuclear cusps $e^{-\lambda\|x\|}$, nonlinearity.

Benchmark packages (analytic): MOLPRO [Werner et al.], GAUSSIAN, CRYSTAL, ...

Grid-based tensor methods in HF calculations: [BNK, Khoromskaia, Flad, 2009, SISC '11],

- Example of a **compact molecule** computed by tensor method: Alanin aminoacid
- 3D lattice structure





- Grid-based tensor numerical methods look promising for computation of structured **extended systems** and periodic compounds !

Building blocks in tensor HF solver [Khoromskaia, BNK '08 - '13]

- **Canonical, Tucker and QTT** tensor arithmetics.
- Grid basis $\{g_\mu\}$, and QTT core Hamiltonian in $\log n$ operations (on example of H_2O)

p	15	16	17	18	19	20
$N^3 = 2^{3p}$	32767^3	65535^3	131071^3	262143^3	524287^3	1048575^3
$Er(\Delta_G)$	0.0027	$6.8 \cdot 10^{-4}$	$1.7 \cdot 10^{-4}$	$4.2 \cdot 10^{-5}$	$1.0 \cdot 10^{-5}$	$2.6 \cdot 10^{-6}$
Richardson e.	-	$1.0 \cdot 10^{-5}$	$8.3 \cdot 10^{-8}$	$2.6 \cdot 10^{-9}$	$3.3 \cdot 10^{-10}$	0
time (sec)	12.8	17.4	25.7	42.6	77	135

- **Fast 3D tensor convolution** via sinc-quadrature, [BNK '08; Bertoglio, BNK '09]:
- Direct/redundancy free **factorizations of TEI** matrix $B = [b_{\mu\nu;\kappa\lambda}] \in \mathbb{R}^{N_b^2 \times N_b^2}$.
- **Cholesky decomposition** (ε-approximation) of B:

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

Compute columns and diagonal of B using "off-line" precomputed factorization.

- QTT compression of the Cholesky factor $L \in \mathbb{R}^{N_b^2 \times R_B}$: $N_b^2 \Rightarrow N_{orb}^2$, $N_b \approx 10 N_{orb}$.
- **DIIS** self-consistent iteration (standard in quantum chemistry).
- **MP2** energy correction via tensor factorizations.

Standard Galerkin scheme

The expansion of the molecular orbitals in $\{g_\mu\}_{1\le\mu\le N_b}$,

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

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

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

where $F(D) = H + J(C) + K(C)$.

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

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

The ground state energy

$$E_{HF} = 2\sum_{i=1}^{N_{orb}} \lambda_i - \sum_{i=1}^{N_{orb}} \left[(\varphi_i, V_H\varphi_i)_{L^2} - (\varphi_i, K\varphi_i)_{L^2}\right].$$

Fock operator via TEI

- Hartree-Fock calculations.
- ▷ Coulomb (Hartree) matrix: given $\overline{D} = vec(D)$,

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

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

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

The rank-N_{orb} decomposition of $D = 2CC^T$ reduces the cost to $O(R_B N_{orb} N_b^2)$,

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

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

- ▷ DIIS iteration to solve Hartree-Fock [Pulay].

Laplace transform and sinc-quadratures [Gavrilyuk, Hackbusch, BNK '08], [Bertoglio, BNK '10]

Green's function for Δ in $\mathbb{R}^3$, via $(2M+1)$-term sinc-quadrature approximation

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

Canonical ε-rank of $\mathbf{P}_{\mathcal{N}} \in \mathbb{R}^{n\times n\times n}$, $R_{\mathcal{N}} = rank(\mathbf{P}_{\mathcal{N}}) \le 30$.

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

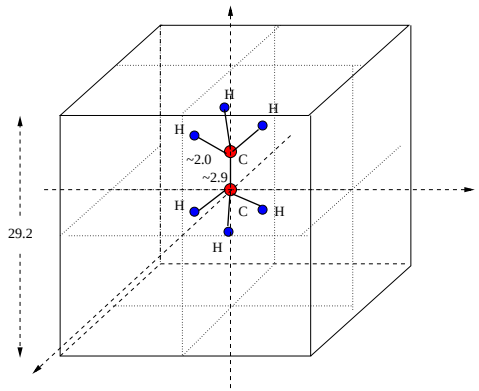
[BNK '08; BNK, Khoromakaia '09]

Applicable to Green's kernels, etc.

$$\frac{e^{-\lambda r}}{r}, \quad \frac{e^{-i\lambda r}}{r}, \quad e^{-\lambda r}.$$

Grid representation of 3D functions and operators

[BNK, Khoromskaia '08 (SISC '09)]



The computational box: $[-b, b]^3$,

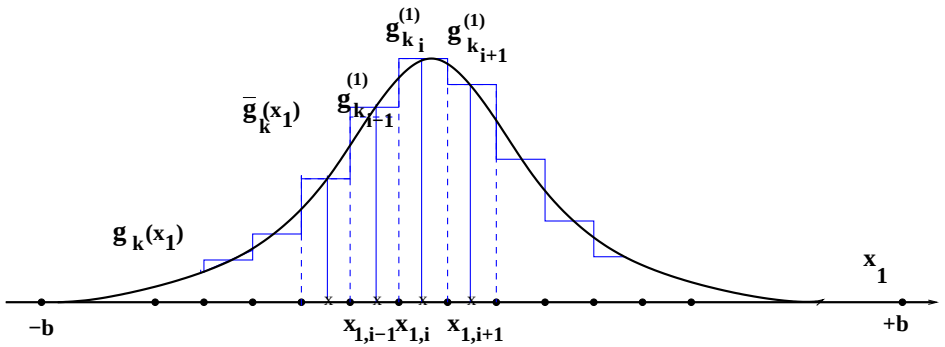
$b \approx 15$ Å

$n \times n \times n$ 3D Cartesian grid, $n \sim 10^4$

$$\mathbf{I}_0 : g_k \to \overline{g}_k := \sum_{\mathbf{i} \in \mathcal{I}} g_k(x_\mathbf{i}) \zeta_\mathbf{i}(x).$$

$$g_k(x) \approx \mathbf{I}_0 g_k := \overline{g}_k(x) = \prod_{\ell=1}^{3} \overline{g}_k^{(\ell)}(x_\ell) = \prod_{\ell=1}^{3} \sum_{i=1}^{n} g_k^{(\ell)}(x_{i_\ell}) \zeta_i^{(\ell)}(x_\ell),$$

rank-1 tensor $\mathbf{G}_k = G_k^{(1)} \otimes G_k^{(2)} \otimes G_k^{(3)}$, canonical vectors $G_k^{(\ell)} = \{g_k^{(\ell)}(x_{i_\ell})\}_{i_\ell=1}^{n} \in \mathbb{R}^{I_\ell}$.



Direct factorization of TEI matrix

Grid-based TEI: [Khoromskaia, BNK, Schneider, SISC'13]

- Separable basis $rank(\mathbf{G}_\mu) = 1 \Rightarrow \mathbf{G}_\mu = G_\mu^{(1)} \otimes G_\mu^{(2)} \otimes G_\mu^{(3)} \in \mathbb{R}^{n\times n\times n}$.
- Rank-$R_\mathcal{N}$ symmetric canonical tensor $\mathbf{P}_\mathcal{N} = \sum_{k=1}^{R_\mathcal{N}} P_k \otimes P_k \otimes P_k \in \mathbb{R}^{n\times n\times n} \approx \frac{1}{\|x\|}$.
- Direct factorization: Compute ("off-line") full set of convolutions $\mathbf{P}_\mathcal{N} * \mathbf{G}$ for tensor

$$\mathbf{G} = [\mathbf{G}_\mu \odot \mathbf{G}_\nu] \in \mathbb{R}^{n^{\otimes 3}\times N_b\times N_b}, \quad \mathbf{G}_\mu \odot \mathbf{G}_\nu \in \mathbb{R}^{n^{\otimes 3}},$$

then

$$[b_{\mu\nu\kappa\lambda}] = \langle \mathbf{G}_\mu \odot \mathbf{G}_\nu, \mathbf{P}_\mathcal{N} * (\mathbf{G}_\kappa \odot \mathbf{G}_\lambda)\rangle_{n^{\otimes 3}}.$$

Given unfolding matrices of the tensor $\mathbf{G}$,

$$G^{(\ell)} = [G_\mu^{(\ell)} \odot G_\nu^{(\ell)}] \in \mathbb{R}^{n\times N_b^2}, \quad \ell = 1, 2, 3,$$

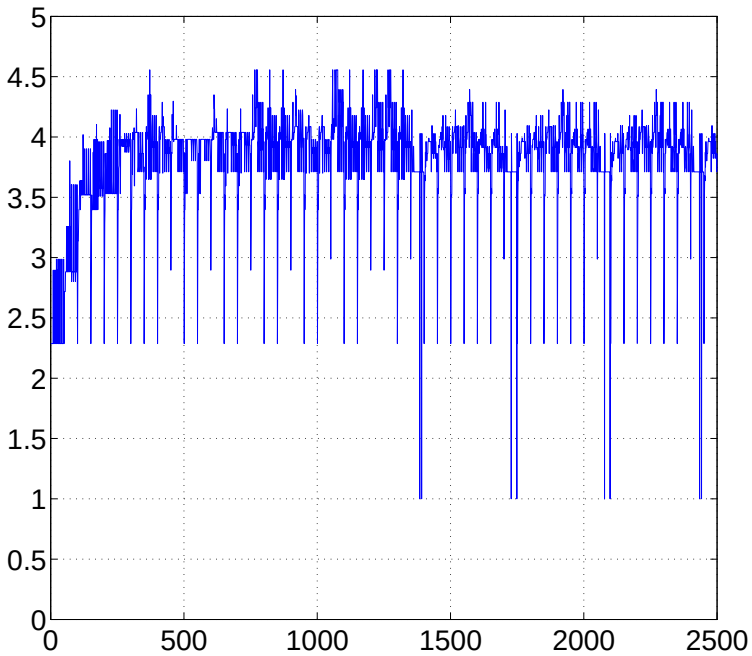
and the factor matrix $P = [P_1, ..., P_{R_\mathcal{N}}] \in \mathbb{R}^{n\times R_\mathcal{N}}$ in $\mathbf{P}_\mathcal{N}$, then

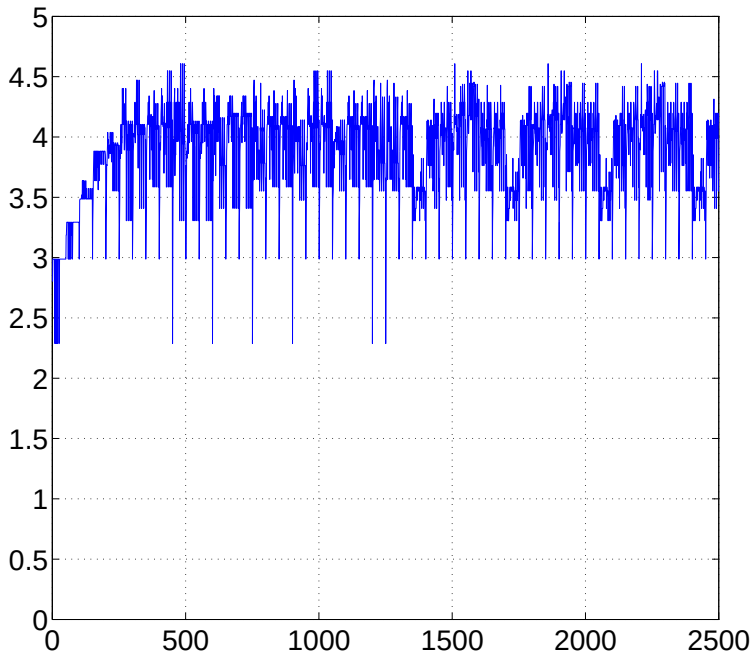
$$B = [b_{\mu\nu;\kappa\lambda}] = \sum_{k=1}^{R_\mathcal{N}} \odot_{\ell=1}^{3} G^{(\ell)^T}(P_k *_n G^{(\ell)}) \quad \Rightarrow \quad G^{(\ell)} \approx U^{(\ell)} V^{(\ell)^T}.$$

- QTT compression in n of matrices $G^{(\ell)}$ and $P_k *_n G^{(\ell)}$.

Numerics: QTT compression leading to $O(\log n)$ HF calculations

Example for CH_4





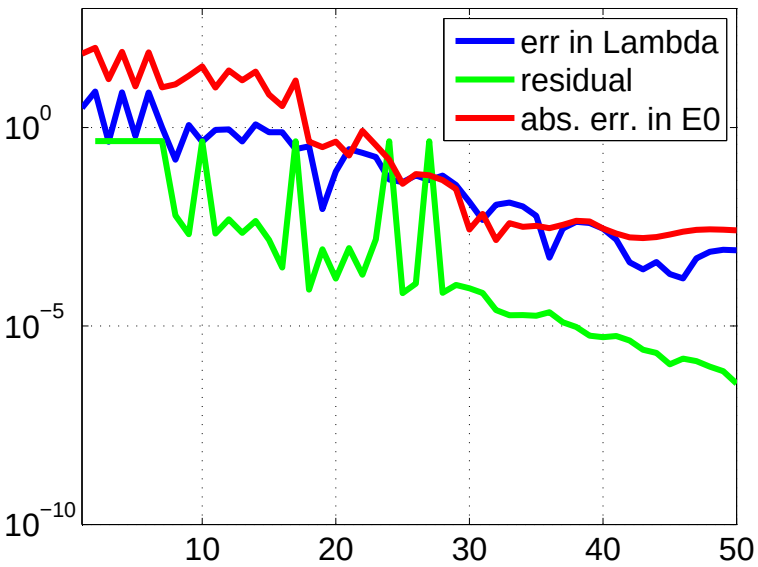
(Left) average QTT ranks of the product basis functions, $\mathbf{G}_{\mu\nu}$.

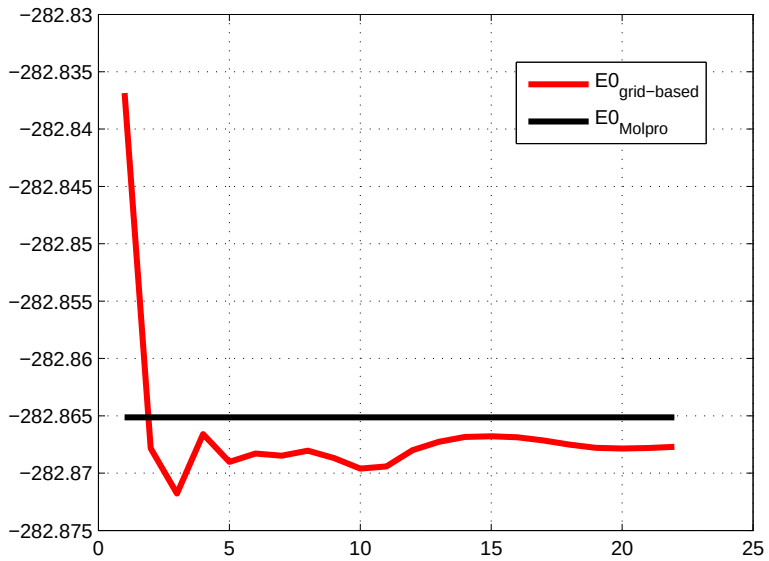
(Right) Newton potential, $\mathbf{P}_{\mathcal{N}} * \mathbf{G}_{\mu\nu}$.

$\varepsilon = 10^{-6}$, $N_b = 55$, $n = 8192$;

[Khoromskaia '13]

Left: convergence of DIIS iteration for Glycin aminoacid, $N_b = 170$ (large TEI).
Right: Computed energy (red) vs. Molpro (black) after $30 + k$ iterations.





The Laplacian $-\Delta$ is approximated on $n^{\otimes 3}$ grid, $n = 32768$.

"On-line" DIIS iterations ≤ 50, $\quad Time_{iter} \approx 2sec$.

Møller-Plesset (MP2) energy correction: factorized TEI + QTT compression

- MP2 energy correction by tensor factorization. [Khoromskaia, BNK (CPC '13)]

Cholesky decomposed TEI matrix $B = LL^T$ is transformed from the AO to the MO-basis,

$$v_{k\ell;mn} = \sum_{\mu,\nu,\kappa,\lambda=1}^{N_b} C_{\mu k} C_{\nu \ell} C_{\kappa m} C_{\lambda n} b_{\mu\nu;\kappa\lambda}, \quad k,\ell,m,n \in \{1,...,N_b\},$$

Cost $O(N_b^4) \Rightarrow O(R_B N_{orb}^3 N_{virt} N_{orb})$.

$$\text{MP2 correction:} \quad E_{MP2} = -\sum_{k,\ell \in I_{virt}} \sum_{m,n \in I_{occ}} \frac{v_{k\ell mn}(2v_{k\ell mn} - v_{knm\ell})}{\lambda_k + \lambda_\ell - \lambda_m - \lambda_n},$$

$I_{occ} := \{1,...,N_{orb}\}$, $I_{virt} := \{N_{orb}+1,...,N_b\}$, and λ_k, $k = 1,...,N_b$.

Reduced complexity $O(R_B N_{virt}^2 N_{orb}^2)$. Storage $R_B N_{virt} N_{orb}$.

MP2 calculations for H_2O. $E_{experiment} = -76.439$.

$E_{Molpro} = -76.0308$; $E_{tensor} = -76.0308$

$E_{MP2} = -76.3292$, CPU time $\approx$ 1 sec.

Main ingredients in tensor approach:

- Computing large lattice sums of the Newton kernels.
- Lattice-structured TEI computation.
- Block-structured representation of the Fock matrix.
- Fast diagonalisation of the Fock matrix.

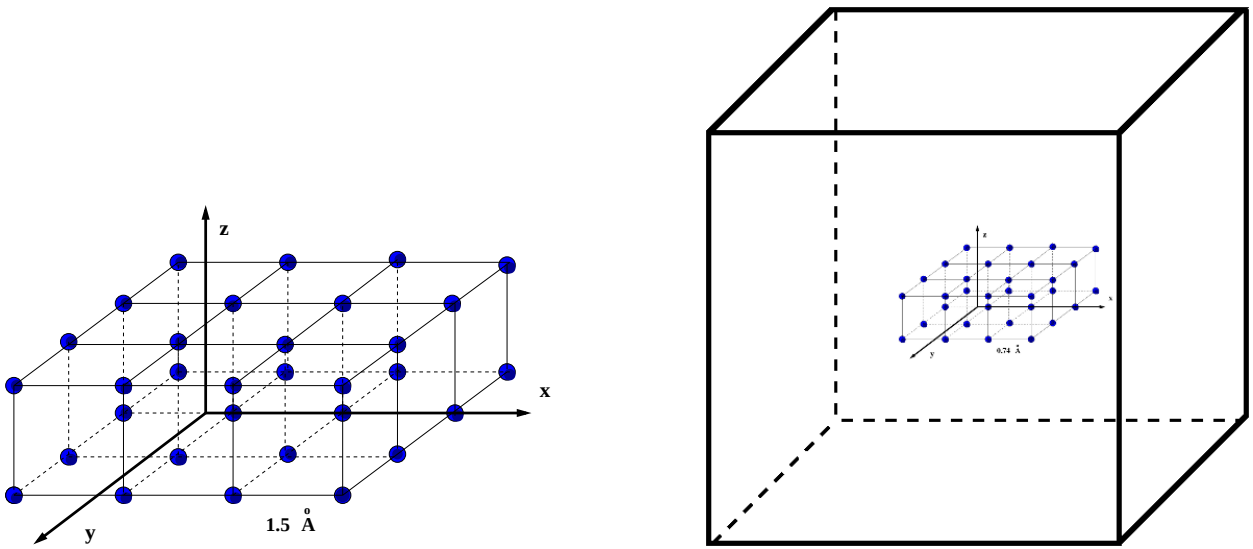


Figure: Periodic structure of size $4.5 \times 4.5 \times 1.5$ $\overset{\circ}{A}^3$ in the box $[-b, b]^3$, with $b = 16$ au (~ 8.5 $\overset{\circ}{A}$).

Examples of nuclear potential and spectrum for extended systems (H cluster)

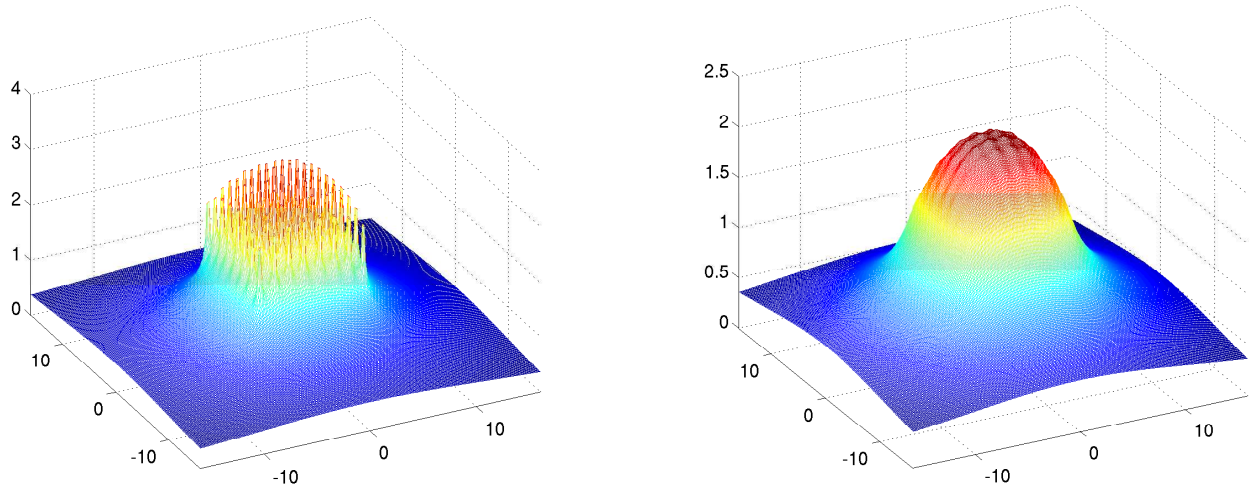


Figure: Decay of the nuclear potential for $8 \times 8 \times 1$ cluster of H_2 atoms.

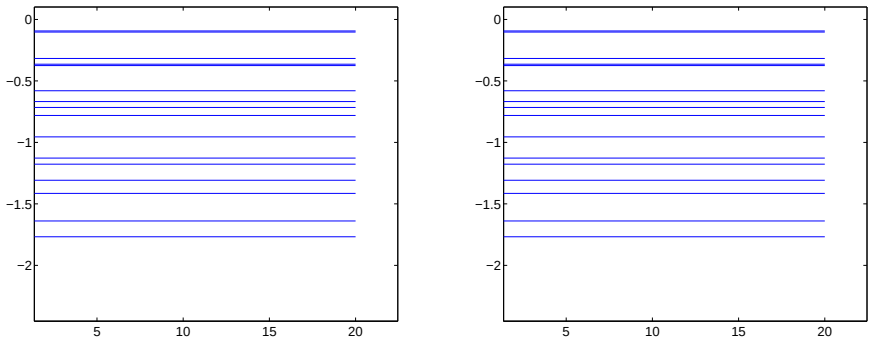


Figure: The negative part of spectra λ_i for the model systems of the size 16×2 and 8×4.

Superfast QTT-FFT transform in $O(\log^2 n)$ operations !

Fast Fourier Transform **(FFT)** – unitary $n \times n$ matrix, $n = 2^d$, $\mathrm{FFT}_n = F_d$):

$$F_d = \frac{1}{2^{d/2}} \left[\omega_d^{jk}\right]_{j,k=0}^{2^d-1}, \qquad \omega_d = \exp\left(-\frac{2\pi \mathrm{i}}{2^d}\right), \quad \mathrm{i}^2 = -1.$$

QTT format for matrix

$$a(i,j) = a(\overline{j_1 \dots j_d}, \overline{k_1 \dots k_d}) = \mathbf{A}(j_1 k_1, j_2 k_2, \dots, j_d k_d) = A^{(1)}_{j_1 k_1} A^{(2)}_{j_2 k_2} \dots A^{(d)}_{j_d k_d}$$

QTT ranks

$$r_p = \operatorname{rank} \mathbf{A}_{[p]}(\underbrace{j_1 k_1\ j_2 k_2\ \dots j_p k_p}_{\text{column index}}; \underbrace{j_{p+1} k_{p+1} \dots j_d k_d}_{\text{row index}})$$

QTT decomposition of FFT matrix has full rank

QTT-FFT matrix has full ε-rank $\Leftrightarrow$ The low-rank ε-approximation is not possible :-(

Given rank-R QTT vector x, $y = \frac{1}{\sqrt{n}} F_d x$ can be computed in QTT format!,

Cost $O(d^2 R^3)$, storage $O(dR^2)$ [Dolgov, BNK, Savostyanov '12].

Cooley-Tuckey FFT meets QTT-data: $n \log n \mapsto \log n$

In contrast to the Fourier transform:
The Hadamard (Walsh) transform has QTT–ranks one,

$$H_d = H_1^{\otimes d}, \qquad H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Fast Fourier transform $y = FFT_n(x)$ for dense vectors costs $O(n \log n)$

$$y = \frac{1}{\sqrt{n}} F_n x \quad \Leftrightarrow \quad y_k = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} x_j \exp\left(-\frac{2\pi \mathrm{i}}{n} jk\right), \quad j,k = 0,\ldots,n-1$$

Reccurence [Cooley, Tuckey, 1965]

$$P_d F_d x = \frac{1}{\sqrt{2}} \begin{bmatrix} F_{d-1} & \\ & F_{d-1} \end{bmatrix} \begin{bmatrix} I & \\ & \Omega_{d-1} \end{bmatrix} \begin{bmatrix} I & I \\ I & -I \end{bmatrix} \begin{bmatrix} x_- \\ x_+ \end{bmatrix},$$

P_d is the *bit-shift* permutation, agglomerating even and odd elements of a vector.

Twiddle factors QTT rank $= 2$

$$\Omega_{\boldsymbol{d}} = \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2^{\boldsymbol{d}}} j\right)\right\}_{\boldsymbol{j=0}}^{\boldsymbol{2^{d-1}-1}} = \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2^{\boldsymbol{d}}} j_{\boldsymbol{1}}\right)\right\} \ldots \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2} j_{\boldsymbol{d-1}}\right)\right\}$$

QTT-Fourier images in 1D

The *rectangle pulse* function, for which the Fourier transform is known,

$$\Pi(t) = \begin{cases} 0, & \text{if} \quad |t| > 1/2 \\ 1/2, & \text{if} \quad |t| = 1/2, \\ 1 & \text{if} \quad |t| < 1/2, \end{cases} \qquad \hat{\Pi}(\xi) = \mathrm{sinc}(\xi) \stackrel{\mathrm{def}}{=} \frac{\sin \pi\xi}{\pi\xi}.$$

The Fourier integral is approximated by rectangular rule.

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} f(t) \exp(-2\pi \mathrm{i} t\xi) dt.$$

$f(t) = \Pi(t)$ is real and even, we write for $k, j = 0, \ldots, n-1$, $n = 2^d$,

$$\hat{f}(\xi_j) = 2\mathrm{Re} \int_0^{+\infty} f(t) \exp(-2\pi \mathrm{i} t\xi_j) dt \approx 2\mathrm{Re} \sum_{k=0}^{n-1} f(t_k) \exp(-2\pi \mathrm{i} t_k \xi_j) h_t,$$

$t_k = (k + 1/2)h_t$, $\xi_j = (j + 1/2)h_\xi$, and use DFT for $h_t = h_\xi = \frac{1}{2^{d/2}}$ and d even.
The QTT representation of the rectangular pulse has QTT–ranks one, i.e.,

$$\Pi(t_k) = \Pi(\frac{h}{2} + \overline{k_1 \ldots k_{d/2-1}} h + \overline{k_{d/2} \ldots k_d}/2) = (1 - k_{d/2}) \ldots (1 - k_d).$$

Numerics on QTT-FFT in 1D: logarithmic complexity !

Table: Time for QTT–FFT (in milliseconds) w.r.t. size $n = 2^d$ and accuracy ε. $time_{QTT}$ is the runtime of Alg. QTT–FFT, $\mathtt{time}_{\mathrm{FFTW}}$ is the runtime of the FFT from the FFTW library, and $\mathrm{rank}\,\hat{f}$ is the effective QTT–rank of the Fourier image.

$f = \Pi(t)$		$\varepsilon = 10^{-4}$		$\varepsilon = 10^{-8}$		$\varepsilon = 10^{-12}$	
d	$\mathtt{time}_{\mathrm{FFTW}}$	$\mathrm{rank}\,\hat{f}$	$\mathtt{time}_{\mathrm{QTT}}$	$\mathrm{rank}\,\hat{f}$	$\mathtt{time}_{\mathrm{QTT}}$	$\mathrm{rank}\,\hat{f}$	$\mathtt{time}_{\mathrm{QTT}}$
16	1.7	4.66	7.9	6.85	13.8	8.85	20.0
18	8.9	4.70	9.7	6.86	16.7	8.82	23.4
20	42.5	4.75	11.3	6.85	19.8	8.86	30.6
22	180	4.77	13.1	6.83	23.3	8.89	36.4
24	810	4.74	15.0	6.72	26.3	8.94	41.7
26	4100	4.62	17.0	6.76	30.0	8.89	46.5
28	26300	4.57	18.9	6.80	33.0	8.88	51.2
30	—	4.72	20.3	6.78	36.2	8.84	57.0
40	—	4.20	29.1	6.59	53.6	8.78	83.2
50	—	3.96	39.3	6.45	70.5	8.48	109
60	—	3.69	50.0	6.25	87.6	8.32	133

Conclusions

Recent progress in fast tensor numerical methods:

- Advanced $O(d \log N)$ tensor formats: QCan, QTT, QTT-Tucker $(+)$
- QTT convolution(d) and super-fast QTT-FFT(d) in $O(d \log N)$ op. $(\pm)$.
- Low-rank preconditioning of elliptic operators on tensor grid. $(\pm)$.
- Tensor solver for the Hartree-Fock eqn. on $N \times N \times N$ grids, $N \leq 10^5$ $(+)$.
- Parametric elliptic/parabolic problems in tensor formats $(\pm)$.
- Time dependent Fokker-Planck and Master equations $(\pm)$.

Future work:

Theoretical understanding, advanced solvers in higher dimensions, real-life applications.

Lecture notes on tensor numerical methods [BNK '10]:
http://www.math.uzh.ch/fileadmin/math/preprints/06_11.pdf

More details: http://personal-homepages.mis.mpg.de/bokh

Dear Martin,
happy 65th and very best wishes !!!