Tensor numerical methods in scientific computing: Basic theory and initial applications

Boris Khoromskij

CEMRACS-2013, CIRM, Marseille-Luminy, 25.07.2013

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

Outline of the talk

Main target: tractable methods for solving d-dimensional PDEs

High-dimensional applications:

- d-dim. operators: Green's functions, Fourier, convolution and wavelet transforms.
- Molecular systems: electronic structure, quantum molecular dynamics.
- PDEs in $\mathbb{R}^d$: quantum information, stochastic PDEs, dynamical systems.

- Elliptic (parameter-dependent) BVP: Find $u \in H_0^1(\Omega)$, s.t.

$$\mathcal{H}u := -\operatorname{div}(a \operatorname{grad} u + u\mathbf{v}) + Vu = F \quad \text{in} \quad \Omega \in \mathbb{R}^d.$$

- Elliptic EVP: Find a pair $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$, s.t.

$$\mathcal{H}u = \lambda u \quad \text{in } \Omega \in \mathbb{R}^d, \quad \langle u, u \rangle = 1.$$

- Parabolic-type equations ($\sigma \in \{1, i\}$): Find $u : \mathbb{R}^d \times (0, T) \to \mathbb{R}$, s.t.

$$u(x, 0) \in H^2(\mathbb{R}^d): \quad \sigma \frac{\partial u}{\partial t} + \mathcal{H}u = 0.$$

Tensor methods adapt gainfully to main challenges:

- High spacial dimension: $\Omega = (-b, b)^d \in \mathbb{R}^d$ ($d = 2, 3, ..., 100, ...$).
- Multiparametric eq.: $a(y, x)$, $F(y, x)$, $u(y, x)$, $y \in \mathbb{R}^M$ ($M = 1, 2, ..., 100, ...$).

Tensor numerical methods in higher dimensions: focus, building blocks, benefits

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, convolution(d), FFT(d), wavelet, 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 high-dimensional PDEs.

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

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

- 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

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

- The scalar product

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

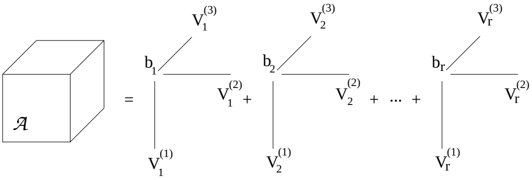
- Storage: $Stor(\mathbf{V}) = \sum_{\ell=1}^{d} n_\ell \ll \dim \mathbb{V}_{\mathbf{n}} = \prod_{\ell=1}^{d} n_\ell$.
- $O(d)$ bilinear operations: addition, Hadamard product, contraction, convolution, ...

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

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

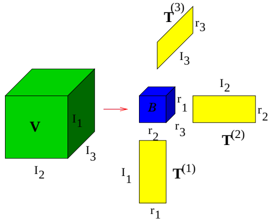
Orthogonal Tucker model

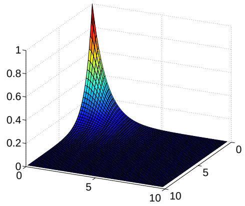
Def. **Rank $\mathbf{r} = [r_1, \ldots, r_d]$ Tucker tensors: $\mathbf{V} \in \mathcal{T}_\mathbf{r}(\mathbb{V}_\mathbf{n})$** if **[Tucker '66]**

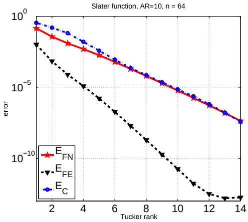
$$\mathbf{V} = \sum_{k_1,\ldots,k_d=1}^{\mathbf{r}} b_{k_1\ldots k_d} v_{k_1}^{(1)} \otimes \ldots \otimes v_{k_d}^{(d)} \in T_1 \otimes \ldots \otimes T_d, \quad T_\ell = span\{v_{k_\ell}^{(\ell)}\}_{k_\ell=1}^{r_\ell} \subset \mathbb{R}^{n_\ell}.$$

- $d = 2$: 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/information:
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 dimension splitting, $O(dr^{\log d} N)$-storage: **[BNK '06]**.
Hierarchical Tucker (HT) $\equiv$ TNS: **[Hackbusch, Kühn '09]**
Tensor train (TT) $\equiv$ MPS (open b.c.) **[Oseledets, Tyrtyshnikov '09]**.

Def. Tensor Train (MPS): 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 $\mathbb{R}^{r_{\ell-1} \times n_\ell \times r_\ell}$,

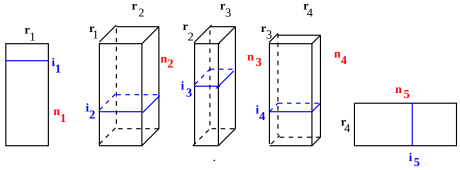
$$\mathbf{V}[i_1 ... i_d] = \sum\nolimits_{\boldsymbol{\alpha}} G^{(1)}_{\alpha_1}[i_1] G^{(2)}_{\alpha_1 \alpha_2}[i_2] \cdots G^{(d)}_{\alpha_{d-1}}[i_d]$$
$$\equiv G^{(1)}[i_1] G^{(2)}[i_2] ... G^{(d)}[i_d],$$

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

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

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

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.

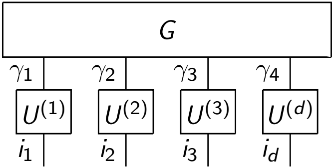
▶ **Limitations**: strong entanglements in a system, large mode-size N.
Multilinear matrix-vector algebra and DMRG iterations cost: $O(dRr^3N^2)$

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

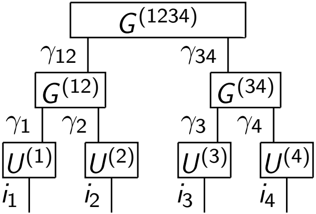
Rank bounds:

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

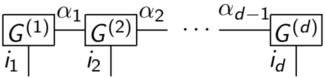
Tucker



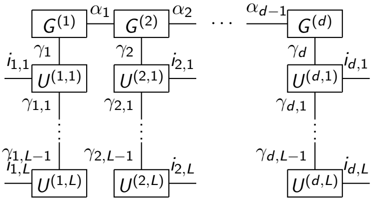
TNS (HT)



MPS (TT)



QTT-Tucker



Canonical $= \sum Tucker_{\gamma_k \in \{1\}}$

Approximation problem: Given $X \in \mathbb{V}_{\mathbf{n}}$ (in general, $X \in \mathcal{S}_0 \subset \mathbb{V}_{\mathbf{n}}$), find

$$T_{\mathbf{r}}(X) := \operatorname*{argmin}_{A \in \mathcal{S}} \|X - A\|, \quad \text{where} \quad \mathcal{S} \subset \{\mathcal{T}_{\mathbf{r}}, \mathcal{C}_{R}, \mathcal{T}\mathcal{C}_{R,\mathbf{r}}, MPS/TT[\mathbf{r}]\}.$$

Quasi-optimal (nonlinear) tensor approximation: SVD, ACA, Greedy

- SVD (or Schmidt) decomposition for matrices
- SVD-based (R)HOSVD for Tucker and canonical tensors [De Lathauwer; BNK, Khoromskaia]
- ACA interpolation [Tyrtyshnikov et al.; Grasedyck et al.].
- Greedy algorithms [Temlyakov, Maday, Cances, Lelievre, Cohen, Dahmen, Chinesta, Nouy, ...].
- SVD-based ALS/DMRG iteration in TT [Dolgov, BNK, Oseledets, Savostianov, ...]

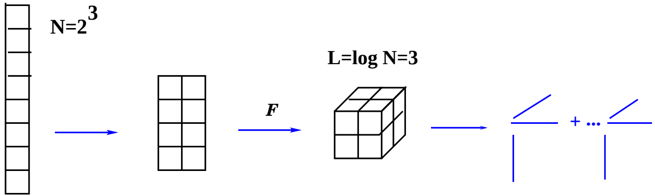
MPS/TT ranks: $TT[\mathbf{r}] := \{\mathbf{A} \in \mathbb{V}_{\mathbf{n}} : \operatorname{rank} \mathbf{A}_{[p]} \leq r_p\}$,

$$r_p = \operatorname{rank} \mathbf{A}_{[p]}(\underbrace{j_1\, j_2\, \ldots j_p}_{\text{column index}}\ ; \underbrace{j_{p+1} \ldots j_d}_{\text{row index}})$$

Canonical rank can not be presented as matrix ranks! $\Rightarrow$ unstable approximation.

Rank reduction in the canonical format:
Reduced HOSVD: Canonical $\mapsto$ Tucker $\mapsto$ canonical (ALS) [BNK, Khoromskaia, SISC '08].

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



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

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

Can/TT approximation of quantized image in $\mathbb{Q}_L \Rightarrow$ QCan/QTT 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: $rank_{QCan,\mathbb{C}}(\mathbf{X}) = rank_{QTT,\mathbb{R}}(\mathbf{X}) = 2$,

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

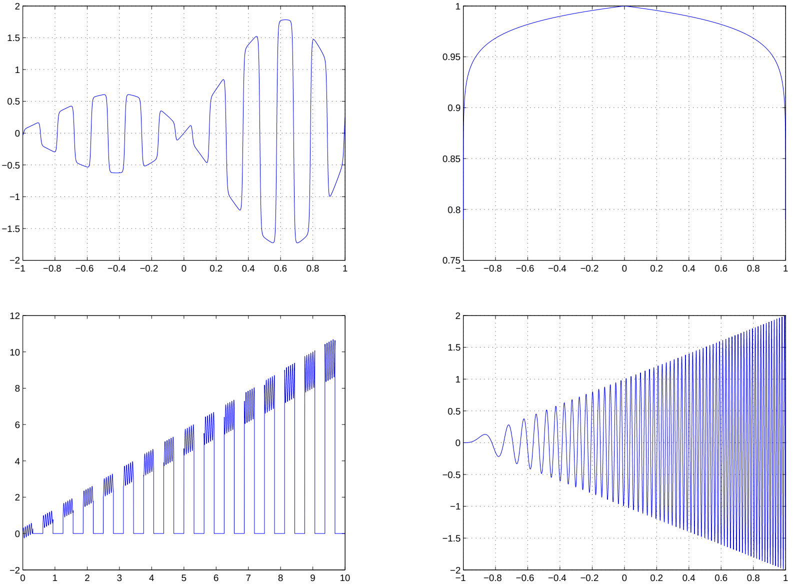
- Proof. Hint: $\sin z = \frac{e^{iz} - e^{-iz}}{2i} = Im(e^{iz})$.

- For QTT-image of polynomial of degree m, $rank_{QTT}(P_m) \leq m + 1$.
- 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

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

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)



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

TT/QTT representation of operators (large multidimensional matrices)

Example. d-dimensional discrete Laplacian.

$$\Delta_d = \Delta_1 \otimes I \otimes \ldots \otimes I + I \otimes \Delta_1 \otimes I \ldots \otimes I + \ldots + I \otimes I \ldots \otimes \Delta_1 \in \mathbb{R}^{N^{\otimes d} \times N^{\otimes d}},$$

$\Delta_1 = \mathrm{tridiag}\{-1, 2, -1\} \in \mathbb{R}^{N \times N}$, I is the $N \times N$ identity.

- Canonical/Tucker representation: $rank_{CP}(\Delta_d) = d$, $rank_{Tuck}(\Delta_d) = 2$.
- Explicit TT representation: $rank_{TT}(\Delta_d) = 2$, $rank_{QTT}(\Delta_d) \leq 4$, $\forall d$.

$$\Delta_d = \begin{bmatrix} \Delta_1 & I \end{bmatrix} \bowtie \begin{bmatrix} I & 0 \\ \Delta_1 & I \end{bmatrix}^{\bowtie (d-2)} \bowtie \begin{bmatrix} I \\ \Delta_1 \end{bmatrix}.$$

- [Kazeev, BNK '10] Explicit QTT representation: $rank_{QTT}(\Delta_1) = 3$, $rank_{QTT}(\Delta_1^{-1}) \leq 5$,

$$\Delta_1 = \begin{bmatrix} I & J' & J \end{bmatrix} \bowtie \begin{bmatrix} I & J' & J \\ & J & \\ & & J' \end{bmatrix}^{\bowtie (d-2)} \bowtie \begin{bmatrix} 2I - J - J' \\ -J \\ -J' \end{bmatrix}.$$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

"$\bowtie$" is a regular matrix product of block core matrices, blocks being multiplied by means of tensor product.

▶ **d-dimensional operators in chemical master equation:**

Thm. **[Dolgov, BNK '12].** Given matrices $E_k,\ F_k^k,\ F_k^{k+1}\ \in \mathbb{R}^{N_k\times N_k}$. The cascadic sum

$$H = F_1^1 \otimes \left(\bigotimes_{k=2}^{d} E_k\right) + \sum_{i=2}^{d}\left(\bigotimes_{k=1}^{i-2} E_k\right)\otimes F_{i-1}^i \otimes F_i^i \otimes \left(\bigotimes_{k=i+1}^{d} E_k\right)$$

possesses an explicit rank-3 TT decomposition $H = H^1(i_1,j_1)\cdots H^d(i_d,j_d)$, with

$$H^1 = \begin{bmatrix} E_1 & F_1^2 & F_1^1 \end{bmatrix}, \quad H^k = \begin{bmatrix} E_k & F_k^{k+1} & 0 \\ 0 & 0 & F_k^k \\ 0 & 0 & E_k \end{bmatrix},$$

$$H^{d-1} = \begin{bmatrix} F_{d-1}^d & 0 \\ 0 & F_{d-1}^{d-1} \\ 0 & E_{d-1} \end{bmatrix}, \quad H^d = \begin{bmatrix} F_d^d \\ E_d \end{bmatrix}.$$

CME operator: **[Kazeev, Schwab '12]**

Similar Hamiltonians arise in the spin systems models with local interactions

[Cirac, Verstraete '06; Huckle et al '12, ...]

Grid-based **Hartree-Fock calculations**

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

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:

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

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

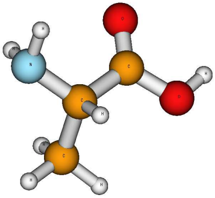
Challenges: High accuracy, 3D singular convolutions, nuclear cusps, hard scaling.

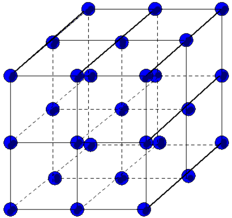
Competing grid-based tensor approach to computational quantum chemistry

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





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

Basic building blocks [Khoromskaia, BNK '08 - '13]

- **Canonical, Tucker and QTT** tensor arithmetics.
- Grid basis $\{g_\mu\}$, and QTT core Hamiltonian (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 tensor convolution** via sinc-quadrature, [BNK '08; Bertoglio, BNK '09]:

$$\frac{1}{r} = \int_0^\infty e^{-t^2 r^2} dt \approx \sum c_k e^{-t_k r^2} \quad \Rightarrow \quad \text{rank-}R_{\mathcal{N}} \quad \text{tensor} \quad \mathbf{P}_{\mathcal{N}} = [P^{(1)}, P^{(2)}, P^{(3)}].$$

- Direct or redundancy free **factorization of TEI** matrix $B = [b_{\mu\nu;\kappa\lambda}]$.
- **Cholesky decomposition** (ε-approximation) of B:

Compute columns and diagonal of B using precomputed factorization,

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

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.
- **MP2** energy correction via tensor factorizations.

Laplace transform and sinc-quadratures

[Gavrilyuk, Hackbusch, BNK '08], [Bertoglio, BNK '10]:

Green's function for Δ in $\mathbb{R}^3$, via sinc-quadrature approximation

$$\frac{1}{r} = \int_0^\infty e^{-t^2 r^2} dt \approx \sum c_k e^{-t_k^2 r^2} = \sum c_k \prod_{i=1}^3 e^{-t_k^2 x_i^2}.$$

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	$\sim$ 6000	–	–	–	$\sim$ 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

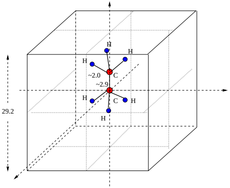
CPU time (in sec) for $V_H = \frac{1}{r} * \rho$ for H_2O.

[BNK '08; BNK, Khoromakaia '09]

Similar for Green's kernels

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

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

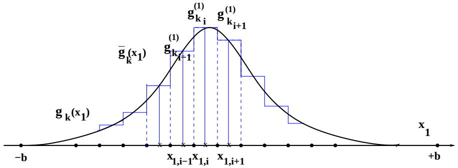


The computational box: $[-b, b]^3$,
$b \approx 15$ Å
$n \times n \times n$ 3D Cartesian grid, $n \sim 10^4$

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

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

rank-1 tensor $\mathbf{G}_k = G_k^{(1)} \otimes G_k^{(2)} \otimes G_k^{(3)}$, can. 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 $rank(\mathbf{G}_\mu) = 1 \Rightarrow \mathbf{G}_\mu = G_\mu^{(1)} \otimes G_\mu^{(2)} \otimes G_\mu^{(3)} \in \mathbb{R}^{n \times n \times n}$.
- Direct factorization: Precompute tensors $\mathbf{G}$ and full set of convolutions $\mathbf{P}_\mathcal{N} * \mathbf{G}$,

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

then

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

The unfolding matrices of the side tensor $\mathbf{G}^{(\ell)} \in \mathbb{R}^{n \times N_b \times N_b}$:

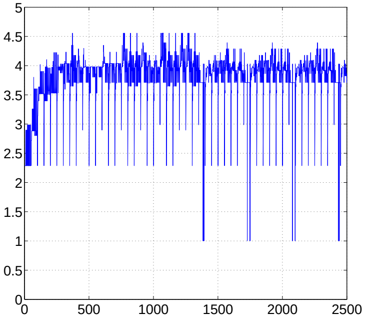
$$G^{(\ell)} = \mathrm{mat}(\mathbf{G}^{(\ell)}) \in \mathbb{R}^{n \times N_b^2}, \quad \ell = 1, 2, 3.$$

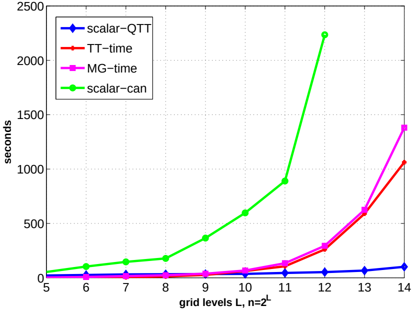
$P^{(\ell)} = [P_k^{(\ell)}] \in \mathbb{R}^{n \times R_\mathcal{N}}$ – factor matrices in the rank-$R_\mathcal{N}$ canonical tensor $\mathbf{P}_\mathcal{N} \in \mathbb{R}^{n \times n \times n}$,

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

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

Example for CH_4





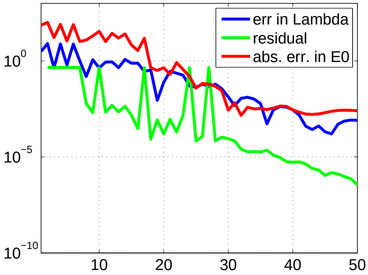
(Left) average QTT ranks of the Newton potential, $\mathbf{P}_{\mathcal{N}} * \mathbf{G}_{\mu\nu}$:
$\varepsilon = 10^{-6}$, $N_b = 55$, $n = 8192$;

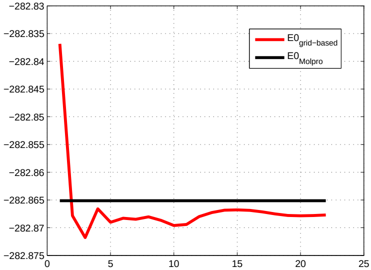
(Right) Coulomb operator:
Times vs. level L of the univariate grid size $n = 2^L$, maximal $n^3 = 2^{42} \approx 10^{12}$.

3D EVP solver for the Hartree-Fock equation in a general basis

[Khoromskaia '13]

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





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

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

Cholesky decomposed TEI matrix $B = L^T L$ 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 ≈ 1 sec.

Low rank canonical approx. Δ_d^{-1} (super-fast Poisson solver, preconditioning)

d-Laplacian: $\Delta_d U = F$ on $N \times N \times ... \times N$ – grid.

- $e^{-t\Delta_d} = \bigotimes_{\ell=1}^{d} e^{-t\Delta_1}$, $\Delta_1 = F_1^* \Lambda_1 F_1$, F_1 is the 1D sin-FFT.
- *sinc*-quadrature approximation $G_M \simeq {\Delta_d}^{-1}$ in rank-R canonical format,

$$\Delta_d^{-1} = \int_0^\infty e^{-t\Delta_d} dt \simeq \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} \exp(-t_k \Delta_1) := \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} F_1^* e^{-t_k \Lambda_1} F_1 := G_M,$$

$$t_k = e^{k\mathfrak{h}}, \quad c_k = \mathfrak{h} t_k, \quad \mathfrak{h} = \pi/\sqrt{M},$$

with the exponential convergence rate in $R = 2M + 1$, **[Gavrilyuk, Hackbusch, BNK '05]**

$$\left\|\Delta_d^{-1} - G_M\right\|_\infty \leq Ce^{-\pi\sqrt{M}}, \quad (\text{or} \ \leq Ce^{-\pi M/\log M}).$$

N	Precomp	Time for sol	Residue	Relative L_2 error
2^8	6.14	2.98	6.6e-06	7.0e-06
2^9	8.37	3.52	8.7e-06	7.0e-06
2^{10}	10.81	4.02	9.4e-06	7.0e-06

100D Poisson eq. in C-QTT, $F = 1$, $W = O(d|\log\varepsilon|^2 \log N)$ complexity.

Find $u_M \in L^2(\Gamma) \times H_0^1(D)$, s.t.

$$\begin{aligned} \mathcal{A} u_M(y,x) &= f(x) \quad \text{in } D, \quad \forall y \in \Gamma, \\ u_M(y,x) &= 0 \quad \text{on } \partial D, \quad \forall y \in \Gamma, \end{aligned}$$

$$\mathcal{A} := -\operatorname{div}\left(a_M(y,x)\operatorname{grad}\right), \quad f \in L^2(D), \quad D \in \mathbb{R}^d, \quad d = 1,2,3,$$

$a_M(y,x)$ is smooth in $x \in D$, $y = (y_1, ..., y_M) \in \Gamma := [-1,1]^M$, $M \leq \infty$.
Additive case (via the truncated Karhunen-Loéve expansion)

$$a_M(y,x) := a_0(x) + \sum_{m=1}^{M} a_m(x) y_m, \quad a_m \in L^\infty(D), \quad M \to \infty.$$

Log-additive case

$$a_M(y,x) := \exp(a_0(x) + \sum_{m=1}^{M} a_m(x) y_m) > 0.$$

- Sparse stochastic Galerkin/collocation: [Babuska, Nobile, Tempone '06-'10; Schwab el. '07-'10, Matthies '06]
- Stochastic Galerkin, $\mathcal{C}_R$ format, additive case: [BNK, Ch. Schwab '10]
- QTT, both additive and log-additive cases: [BNK, Oseledets '10]
- HT, additive: [Kressner, Tobler '10]

Stochastic collocation (additive case): discretization in x and y

A parametric linear system, N - grid size in x (Galerkin-FEM, FD in x)

$$A(y)u(y) = f, \quad f \in \mathbb{R}^N,\ u(y) \in \mathbb{R}^N, \quad y \in \Gamma, \tag{1}$$

$$A(y) = A_0 + \sum_{m=1}^{M} A_m y_m, \quad A_m \in \mathbb{R}^{N \times N}, \quad \text{parameter dependent matrix.}$$

Collocation on $n^{\otimes M}$ grid, n - grid size in y (uniform, Chebyshev, etc.)

$$\{y_m^{(k)}\} =: \Gamma_m \in [-1,1],\ k = 1,\ldots,n, \quad \Gamma_n^y = \bigotimes_{m=1}^{M} \Gamma_m$$

$\Rightarrow$ Assembled large linear system

$$\mathbb{A}\mathbf{u} = \mathbf{f}, \quad \mathbf{u}, \mathbf{f} \in \mathbb{R}^{N \times n^{\otimes M}}, \quad \mathbb{A} \in \mathbb{R}^{(N \times n^{\otimes M}) \times (N \times n^{\otimes M})},$$

$$\mathbb{A} = A_0 \times I \times \ldots \times I + A_1 \times D_1 \times I \times \ldots \times I + \ldots + A_M \times I \times \ldots \times D_M,$$

D_m, $m = 1,\ldots,M$, is $n \times n$ diagonal matrix with positions of collocation points, $\{y_m^{(k)}\} \in \Gamma_m$, on the diagonal: $rank_{CP}(\mathbb{A}) \leq M$.

$$\mathbf{f} = f \times e \times \ldots \times e, \quad e = (1,...,1)^T \in \mathbb{R}^n.$$

Stochastic/parametric PDEs: iteration on tensor manifold

- Parametric elliptic BVP on nonlinear manifold $\mathcal{S}$:

$$A(y)u(y) = f \quad \Rightarrow \quad \mathbb{A}\mathbf{u} = \mathbf{f},$$

$$\widetilde{\mathbf{u}}_{m+1} = \mathbf{u}_m - \mathbb{B}^{-1}(\mathbb{A}\mathbf{u}_m - \mathbf{f}), \quad \mathbf{u}_{m+1} := T_{\mathcal{S}}(\widetilde{\mathbf{u}}_{m+1}) \in \mathcal{S}.$$

Assumptions:

- $\mathbf{u}, \mathbf{f}$ allow the low $\mathcal{S}$-rank approximation,
- $\mathbb{A}$ and $\mathbb{B}^{-1}$ are of low matrix $\mathcal{S}$-rank,
- $\mathbb{A}$ and $\mathbb{B}$ are spectral equivalent (close).

Good candidates for $\mathbb{B}^{-1}$:

(A) Shifted FD d-Laplacian inverse $(\Delta_d + aI)^{-1}$.

$$\Delta_d = \Delta_1 \otimes I_N \otimes ... \otimes I_N + ... + I_N \otimes I_N ... \otimes \Delta_1 \in \mathbb{R}^{N^{\otimes d} \times N^{\otimes d}}.$$

(B) $\mathbb{A}(y^*)^{-1}$ with optimization in y^*.
(C) Reciprocal preconditioner $\mathbb{B}^{-1} = \Delta_d^{-1} \mathbb{A}_{[1/a(y)]} \Delta_d^{-1}$: Strongly variable coeff. $a(y,x)$.
[Dolgov, BNK, Oseledets, Tyrtyshnikov '10]

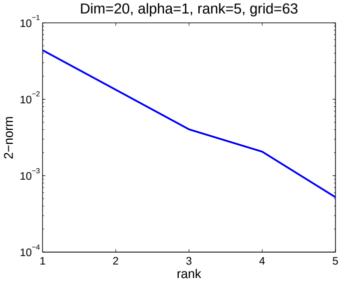
- Rank bounds: $rank_{Can}(\Delta_d^{-1}) \leq C|\log \varepsilon|^2$, $rank_{QTT}(\Delta_d^{-1}) \leq C\, rank_{Can}(\Delta_d^{-1})$.

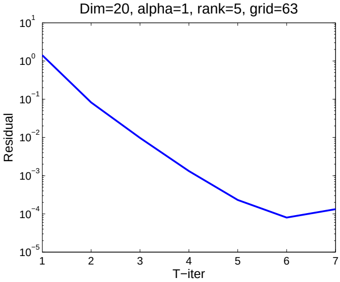
[BNK, Ch. Schwab, '11]

▶ Preconditioned $\mathcal{S}$-truncated iteration in $(d+M)$-dimensional parametric space. Rank-R canonical format, $M \leq 100$.

$N^{\otimes(M+d)}$-grid, $d=1$, $M=20$ ($\mathcal{S}=\mathcal{C}_R$, $\mathbb{B}^{-1} := \mathbb{A}(0)^{-1}$).
Variable coefficients with exponential decay ($N=63$, $R \leq 5$),

$$a_m(x) = 0.5\, e^{-\alpha m} sin(mx), \quad m = 1, 2, ..., M, \quad x \in (0, \pi).$$





Time dependent problems: dynamics on tensor manifold

- Parabolic BVP projected onto $\mathcal{S} \subset \mathbb{V}_{\mathbf{n}}$: $U \in \mathcal{S}$,

$$\sigma \frac{\partial U}{\partial t} = AU + F, \quad U(0) = T_{\mathcal{S}} U(0), \quad \sigma = 1, i.$$

Problem 1. Complex-time molecular Schrödinger eq. in QMD,

$$i \frac{\partial \psi}{\partial t} = H\psi = (-\frac{1}{2}\Delta_d + V)\psi, \quad \psi(x,0) = \psi_0(x), \quad x \in \mathbb{R}^d,$$

$V : \mathbb{R}^d \to \mathbb{R}$ is (given) approximation to the potential energy surface (PES).

Problem 2. Real-time evolution. The Fokker-Planck equation

$$\psi(0) = \psi_0, \quad \frac{d\psi}{dt} = -A\psi; \quad A\psi = -\varepsilon\Delta\psi + \mathrm{div}(\psi\mathbf{v}), \quad \psi : \mathbb{R}^d \to \mathbb{R},$$

$\mathbf{v} : \mathbb{R}^d \to \mathbb{R}^d$ is a given velocity field. $\psi(t) \to \psi_* : A\psi_* = 0$.

Problem 3. Chemical master equation. Joint probability density $\mathcal{P}(\mathbf{x}, t)$,

$$\mathcal{P}(\mathbf{x}, 0) = \mathcal{P}_0, \quad \frac{d\mathcal{P}(\mathbf{x}, t)}{dt} = \mathbf{A}\mathcal{P}(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^{n_1 \times \dots \times n_d}.$$

Main approaches

- **Time integrators**
  - Sparse grids in (x, t): **[Schwab et al.; Griebel et al.]**
  - Dirac-Frenkel projection onto Tucker/TT/QTT-manifold $\mathcal{S}$,

$$\langle \frac{dy}{dt} - Ay, \delta y \rangle = 0, \quad \delta y \in T_y \mathcal{S}.$$

    [Meyer et al. '03; Lubich '07-'12; BNK, Oseledets, Schneider '12]
  - Greedy iterations (canonical format)
    [Cancés, Le Brie, Lelievre, Maday et al; Chinesta et al; Suli et al; Binev, Cohen, Dahmen, et al]
  - Time stepping by implicit scheme + TT/QTT + ALS/DMRG local solver,
    [Dolgov, BNK, Oseledets '11; Kaseev, Schwab '12]
- **Global time-space schemes**
  - Global time-space separation by QTT-Cayley transform **[Gavrilyuk, BNK '11]**
  - QTT-Tucker + ALS-type solver on global (x, t) tensor manifold **[Dolgov, BNK '12-'13]**

Discretization on global QTT-tensor manifold in space-time

Crank-Nicolson, A, f_k are given in the TT/QTT format,

$$(I + \frac{\tau}{2}A)y_{k+1} = (I - \frac{\tau}{2}A)y_k + \frac{\tau}{2}(f_k + f_{k+1}) =: F_{k+1}.$$

(A) Time stepping by DMRG-TT iter. for

$$(I + \frac{\tau}{2}A)y_{k+1} = F_{k+1}.$$

(B) Global $O(\log N_x \log N_t)$ block solver in QTT format:

$$y_{k+1} - y_k + \frac{\tau}{2}Ay_{k+1} + \frac{\tau}{2}Ay_k = \frac{\tau}{2}(f_k + f_{k+1}).$$

Solve the huge global $N_x^d \times N_t$ system in QTT format,

$$\begin{bmatrix} I + \frac{\tau}{2}A & & & \\ -I + \frac{\tau}{2}A & I + \frac{\tau}{2}A & & \\ & \ddots & \ddots & \\ & & -I + \frac{\tau}{2}A & I + \frac{\tau}{2}A \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \left(I - \frac{\tau}{2}A\right) y^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \frac{\tau}{2} \begin{bmatrix} f_0 + f_1 \\ f_1 + f_2 \\ \vdots \\ f_{N-1} + f_N \end{bmatrix}.$$

▶ **QTT-Tucker for chemical master equation.** [Dolgov, BNK, '13]

Species $S_1, ..., S_d$ react in M channels.
The vector of concentration $\mathbf{x} = (x_1, ..., x_d)$, $x_i \in \{0, ..., N_i - 1\}$,
The stoichiometric vector $\mathbf{z^m} \in \mathbb{Z}^d$,
The propensity function $w^m(\mathbf{x})$, $m = 1, ..., M$.

$$J^{\mathbf{z}} = \begin{bmatrix} 0 & \cdots & 1 & & \\ & \ddots & & \ddots & \\ & & \ddots & & 1 \\ & & & \ddots & \vdots \\ & & & & 0 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \quad \text{row} \quad N - z, \\ \\ \leftarrow \quad \text{row} \quad N \end{matrix} \quad \text{if } z \geq 0; \qquad J^{\mathbf{z}} = (J^{-\mathbf{z}})^\top, \text{if } z < 0.$$

CME is a deterministic difference equation on the joint probability density $P(x, t)$:

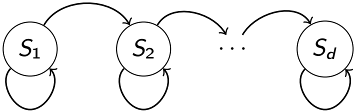
$$\frac{dP(t)}{dt} = \sum_{m=1}^{M} (\mathbf{J}^{\mathbf{z^m}} - \mathbf{J}^0) \,\mathrm{diag}(w^m) P(t), \qquad P(t) \in \mathbb{R}_+^{\prod_{i=1}^d N_i},$$

$$\mathbf{J^z} = J^{z_1} \otimes \cdots \otimes J^{z_d},$$

$$w^m = \{w^m(\mathbf{x})\} \text{ and } P(t) = \{P(\mathbf{x}, t)\}, \ \mathbf{x} \in \bigotimes_{i=1}^{d} \{0, \ldots, N_i - 1\}.$$

Long-time dynamics: CME for the signaling cascade, $d = 20$

Figure: Cascade signaling network



- $d = 20$, $M = 40$;
- for $m = 1$: $w^m(\mathbf{x}) = 0.7$, $\mathbf{z^m} = -\delta_m$: generation of the first protein;
- for $m = 2, ..., 20$: $w^m(\mathbf{x}) = \dfrac{x_{m-1}}{5 + x_{m-1}}$, $\mathbf{z^m} = -\delta_m$: succeeding creation reactions;
- for $m = 21, ..., 40$: $w^m(\mathbf{x}) = 0.07 \cdot x_{m-20}$, $\mathbf{z^m} = \delta_{m-20}$: destruction reactions.
- $N_i = 63$.

δ_m is the m-th identity vector.

Problem size 64^{20}

Convergence history

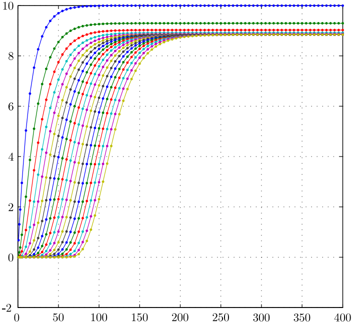


Figure: Mean concentrations $\langle x_{\boldsymbol{i}} \rangle(t)$.

Figure: Closeness to the kernel $\frac{||AP||}{||P||}(t)$

$O(\log N_t)$ CPU time

The performance of the global state-time scheme vs.

- numbers of time steps N_t in each interval $[(p-1)T_0, pT_0]$.
- the time interval widths T_0

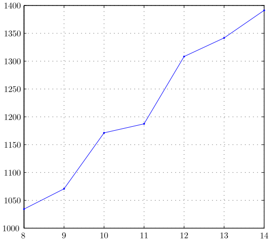


Figure: CPU time (sec.) versus $\log_2(N_t)$, $T_0 = 15$.

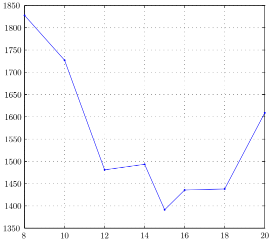


Figure: CPU time (sec.) versus T_0, $N_t = 2^{14}$.

- Logarithmic complexity in N_t.
- There is an optimal time-step T_0.

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

FFT matrix (unitary $n \times n$, $n = 2^d$, $\mathrm{FFT}_n = F_d$).

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

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

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

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

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

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

Reccurence [Cooley, Tuckey, 1965]

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

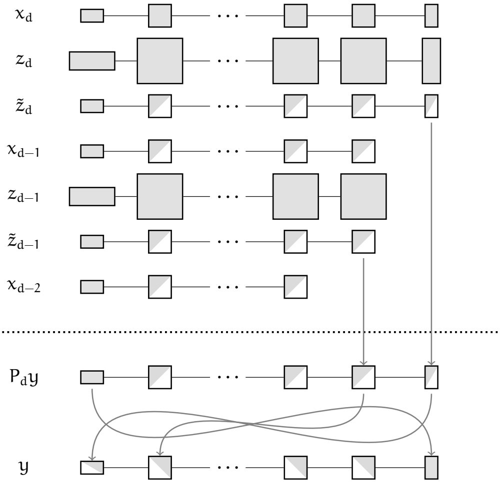
P_d is the *bit-shift* permutation, agglomerating even and odd elements of a vector.

Twiddle factors QTT rank $= 2$

$$\Omega_d = \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2^d} j\right)\right\}_{j=0}^{2^{d-1}-1} = \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2^d} j_1\right)\right\} \dots \mathrm{diag}\left\{\exp\left(-\frac{2\pi \mathrm{i}}{2} j_{d-1}\right)\right\}$$

QTT-FFT approximation scheme

Combine exact QTT-FFT with rank truncation: Algorithm (d-level approximation)



QTT-Fourier images in 1D

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

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

The Fourier integral is approximated by rectangular rule.

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

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

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

$t_k = (k + 1/2) h_t$, $\xi_j = (j + 1/2) h_\xi$, 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

Thank you for your attention !