Lectures on Tensor Numerical Methods for Multi-dimensional PDEs

Lect. 7-8-9. Polynomial and sinc-approximation in $\mathbb{R}^d$, TT-format, QTT approximation of functions and operators, integrating exotic oscillators, super-fast QTT-FFT/FWT. Numerical illustrations.

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

Max-Planck-Institute for Mathematics in the Sciences, Leipzig
MAX-PLANCK-GESELLSCHAFT

Polynomial and *sinc* approximation in $\mathbb{R}^d$, TT-format, QTT approximation

Outline of Lectures 7-8-9.

1. Polynomial approximation of analytic functions in $\mathbb{R}^d$.
2. Tensor product Polynomial interpolation. Example for the Helmhotz kernel.
3. *sinc*-approximation and -quadratures for analytic functions in Hardy space.
4. *sinc*-quadratures for the Laplace transform of Green's kernels: exponential convergence
5. Matrix product states (MPS) in the form of tensor train (TT) format.
6. Nonlinear approximation in tensor formats revisited. Big picture.
7. Quantized tensor approximation: Q-canonical (QCan) and QTT formats.
8. QTT approximation of functions.
9. Examples of TT/QTT representation of matrices (operators).
10. Fast QTT-based numerical quadratures for exotic oscillators. Super-fast QTT-FFT/FWT.
11. Modern tensor numerical methods: main ingredients and challenges.

Polynomial approximation. Chebyshev polynomials

▶ The Chebyshev polynomials, $T_n(w)$, $w \in \mathbb{C}$ - complex plane, are defined recursively

$$\begin{aligned} T_0(w) &= 1, \qquad T_1(w) = w, \\ T_{n+1}(w) &= 2wT_n(w) - T_{n-1}(w), \quad n = 1, 2, \ldots. \end{aligned}$$

Representation $T_n(x) = \cos(n \arccos x)$, $x \in B := [-1, 1]$, implies $T_n(1) = 1$, $T_n(-1) = (-1)^n$. There holds

$$T_n(w) = \frac{1}{2}(z^n + z^{-n}) \quad \text{with} \quad w = \frac{1}{2}(z + \frac{1}{z}). \tag{1}$$

Let $B := [-1, 1]$ be the reference interval.
Def. Denote by $\mathcal{E}_\rho = \mathcal{E}_\rho(B)$ the Bernstein's regularity ellipse

$$\mathcal{E}_\rho := \{w \in \mathbb{C} : \ |w - 1| + |w + 1| \le \rho + \rho^{-1}\}$$

with foci at $w = \pm 1$ and the sum of semi-axes equal to $\rho > 1$.

Denote by $\mathcal{P}_N(B)$ the set of polynomials of degree $\le N$ on B.

Rem. Chebyshev series provides asymptotically the same approximation error (Thm. 7.1) as for the best polynomial approximation (S. N. Bernstein, 1880-1968).

Best polynomial approximation by Chebyshev series

Thm. 7.1. (Chebyshev series). Let F be analytic and bounded by M in $\mathcal{E}_\rho$, $\rho > 1$. Then

$$F(w) = C_0 + 2\sum_{n=1}^{\infty} C_n T_n(w), \tag{2}$$

holds for all $w \in \mathcal{E}_\rho$, with $C_n = \frac{1}{\pi}\int_{-1}^{1} \frac{F(w)T_n(w)}{\sqrt{1-w^2}} dw$.
Moreover, $|C_n| \le M/\rho^n$. For $w \in B$, and for $m = 1, 2, 3, \ldots,$

$$|F(w) - C_0 - 2\sum_{n=1}^{m} C_n T_n(w)| \le \frac{2M}{\rho - 1}\rho^{-m}, \quad w \in B. \tag{3}$$

▶ Given the set $\{\xi_j\}_{j=0}^N$ of interpolation points on B, the Lagrangian interpolant $\mathcal{I}_N$ of $F \in C[B]$ has the form

$$\mathcal{I}_N F := \sum\nolimits_{j=0}^{N} F(\xi_j) l_j(x) \in \mathcal{P}_N(B) \tag{4}$$

with $l_j(x)$ being the set of interpolation polynomials

$$l_j := \prod_{k=0, j\neq k}^{N} \frac{x - \xi_k}{\xi_j - \xi_k} \in \mathcal{P}_N(B), \quad j = 0, \ldots, N.$$

Clearly, $\mathcal{I}_N(\xi_j) = F(\xi_j)$, since $l_j(\xi_j) = 1$ and $l_j(\xi_k) = 0 \ \forall k \neq j$.

Lagrangian polynomial interpolation

▶ The inf-norm of the interpolant $\mathcal{I}_N$ is bounded by the Lebesque constant $\Lambda_N \in \mathbb{R}_{>1}$,

$$\|\mathcal{I}_N u\|_{\infty,B} \leq \Lambda_N \|u\|_{\infty,B} \qquad \forall u \in C(B). \tag{5}$$

Let $[\mathcal{I}_N F](x) \in \mathcal{P}_N(B)$ define the interpolation polynomial of F w.r.t. the Chebyshev-Gauss-Lobatto (CGL) nodes

$$\xi_j = \cos\frac{\pi j}{N} \in B, \quad j = 0, 1, \ldots, N, \qquad \text{with } \xi_0 = 1,\ \xi_N = -1,$$

where ξ_j are zeros of the polynomials $(1 - x^2)T_N'(x)$, $x \in B$.

▶ In the case of Chebyshev interpolation Λ_N grows at most logarithmically in N,

$$\Lambda_N \leq \frac{2}{\pi} \log N + 1.$$

The interpolation points which produce the smallest value Λ_N^* of all Λ_N are not known, but Bernstein (1854) proves that

$$\Lambda_N^* = \frac{2}{\pi} \log N + O(1).$$

▶ The interpolation operator $\mathcal{I}_N$ is a projection, that is, for all $v \in \mathcal{P}_N$ we have $\mathcal{I}_N v = v$.

Optimal error bound for polynomial interpolation. Multivariate case

Thm. 7.2. Let $u \in C^\infty[-1,1]$ have an analytic extension to $\mathcal{E}_\rho$ bounded by $M > 0$ in $\mathcal{E}_\rho$ (with $\rho > 1$). Then we have

$$\|u - \mathcal{I}_N u\|_{\infty,I} \leq (1 + \Lambda_N)\frac{2M}{\rho - 1}\rho^{-N}, \quad N \in \mathbb{N}_0. \tag{6}$$

Proof. Due to (3) one obtains for the best polynomial approximations to u on $[-1,1]$,

$$\min_{v \in \mathcal{P}_N} \|u - v\|_{\infty,B} \leq \frac{2M}{\rho - 1}\rho^{-N}.$$

The interpolation operator $\mathcal{I}_N$ is a projection. Now apply the triangle inequality,

$$\|u - \mathcal{I}_N u\|_{\infty,B} = \|u - v - \mathcal{I}_N(u - v)\|_{\infty,B} \leq (1 + \Lambda_N)\|u - v\|_{\infty,B}.$$

▶ Given a set of interpolating functions $\{\varphi_j(x)\}$, $x \in B$, and sampling points $\xi_i \in B$ $(i, j = 0, 1, ..., N)$, s.t. $\varphi_j(\xi_i) = \delta_{ij}$. For $f \in C[B^d]$, the *tensor-product interpolant* $\mathbf{I}_N$ in d variables reads

$$\mathbf{I}_N f = \mathcal{I}_N^1 \times \mathcal{I}_N^2 \times \ldots \times \mathcal{I}_N^d f := \sum_{\mathbf{j}=\mathbf{0}}^{\mathbf{N}} f(\xi_{j_1}, ..., \xi_{j_d})\varphi_{j_1}^{(1)}(x_1)...\varphi_{j_d}^{(d)}(x_d).$$

Tensor product polynomial interpolation

To derive an multidimensional analogue of Thm. 7.2, introduce the product domain

$$\mathcal{E}_\rho^{(j)} := B_1 \times \ldots \times B_{j-1} \times \mathcal{E}_\rho(I_j) \times B_{j+1} \times \ldots \times B_d,$$

and denote by X_{-j} the $(d-1)$-dimensional (single-hole) subset of variables

$$\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d\} \quad \text{with} \quad x_j \in B_j, \quad j = 1, ..., d.$$

Assump. 7.1. Given $f \in C^\infty(B^d)$, assume there is $\rho > 1$ s.t. for all $j = 1, \ldots, d$, and each fixed $\xi \in X_{-j}$, there exists an analytic extension of $f(x_j, \xi)$ to $\mathcal{E}_\rho(B_j) \subset \mathbb{C}$ w.r.t. x_j, $\hat{f}_j(x_j, \xi)$, bounded in $\mathcal{E}_\rho(B_j)$ by certain $M_j > 0$, independent on ξ.

Thm. 7.3. For $f \in C^\infty(B^d)$, let Assump. 7.1 be satisfied. Then the interpolation error can be estimated by

$$\|f - \mathbf{I}_N f\|_{\infty, B^d} \le \Lambda_N^d \frac{2M_\rho(f)}{\rho - 1} \rho^{-N}, \tag{7}$$

where Λ_N is the maximal Lebesque const. of the 1D interpolants $\mathcal{I}_N^k$, $k = 1, \ldots, d$, and

$$M_\rho(f) := \max_{1 \le j \le d} \{ \max_{x \in \mathcal{E}_\rho^{(j)}} |\hat{f}_j(x, \xi)| \}.$$

Proof of Thm. 7.3

Proof. Multiple use of (5), (6) and the triangle inequality lead to

$$\begin{aligned}
|f - \mathbf{I}_N f| &\le |f - \mathcal{I}_N^1 f| + |\mathcal{I}_N^1 (f - \mathcal{I}_N^2 \times \ldots \times \mathcal{I}_N^d f)| \\
&\le |f - \mathcal{I}_N^1 f| + |\mathcal{I}_N^1 (f - \mathcal{I}_N^2 f)| + \\
&+ |\mathcal{I}_N^1 \mathcal{I}_N^2 (f - \mathcal{I}_N^3 f)| + \ldots + |\mathcal{I}_N^1 \times \ldots \times \mathcal{I}_N^{d-1} (f - \mathcal{I}_N^d f)| \\
&\le [(1 + \Lambda_N) \max_{x \in \mathcal{E}_\rho^{(1)}} |\hat{f}_1(x, \xi)| + \Lambda_N (1 + \Lambda_N) \max_{x \in \mathcal{E}_\rho^{(2)}} |\hat{f}_2(x, \xi)| \\
&+ \ldots + \Lambda_N^{d-1} (1 + \Lambda_N) \max_{x \in \mathcal{E}_\rho^{(d)}} |\hat{f}_d(x, \xi)|] \frac{2}{\rho - 1} \rho^{-N} \\
&\le \frac{(1 + \Lambda_N)(\Lambda_N^d - 1)}{\Lambda_N - 1} \frac{2M_\rho}{\rho - 1} \rho^{-N}.
\end{aligned}$$

Hence (7) follows since for $x \ge 1$ we have

$$\frac{(1 + x)(x^n - 1)}{x - 1} \le x^n,$$

which complete the proof.

Application to the Helmholtz kernel: overview of the main results

Are the Tucker/canonical models robust to the frequency κ ?

Goal: Separable approximation of the Newton kernel ($\kappa = 0$), [Hackbush, Khoromskij '07]

$$f(x) = \frac{1}{\|x\|}, \quad x \in \mathbb{R}^3,$$

and the oscillatory potentials (polynomials in $\|x\|^2$), [Khoromskij, Constr. Approx. '09]

$$f_{1,\kappa}(\|x\|) := \frac{sin(\kappa\|x\|)}{\|x\|}; \quad f_{2,\kappa}(\|x\|) := \frac{2sin^2(\frac{\kappa}{2}\|x\|)}{\|x\|} = \frac{1}{\|x\|} - \frac{cos(\kappa\|x\|)}{\|x\|}, \quad x \in \mathbb{R}^d.$$

▶ Construct exponentially convergent tensor decompositions of the classical Helmholtz kernel in $\mathbb{R}^3$, $\frac{e^{i\kappa\|x-y\|}}{\|x-y\|}$, $\kappa \in \mathbb{R}$, s.t. its real and imaginary parts are treated separately, [Khoromskij '09],

$$\frac{\cos(\kappa\|x-y\|)}{\|x-y\|} \quad \text{and} \quad \frac{\sin(\kappa\|x-y\|)}{\|x-y\|}, \quad x, y \in \mathbb{R}^3.$$

Main result 1: The ε-rank for both Tucker and canonical approx. to $\frac{1}{\|x\|}$, is bounded by

$$r_T \le R_{CP} \le Cd(\log^2 \varepsilon).$$

Main result 2: The Tucker and canonical approximations to $f_{1,\kappa}$, $f_{2,\kappa}$, allow the ε-rank bound

$$r_T(f_{1,\kappa}) \le R_{CP} \le Cd(|\log\varepsilon| + \kappa), \quad r_T(f_{2,\kappa}) \le R_{CP} \le Cd^2|\log\varepsilon|(|\log\varepsilon| + \kappa).$$

Approximation via *sinc* interpolation and quadratures

▶ The Tucker/CP models apply to analytic functions with point singularities (say, $f = f(\|x\|)$).

I. Approximating by exponential sums (canonical model)

- Sinc quadratures (simple direct method)

▶ The canonical format applies well to functions depending on a sum of single variables. Assume a function of $\rho = \sum_{i=1}^d x_i$ be given by the integral

$$f(\rho) = \int_\Omega G(t)e^{\rho F(t)}\mathrm{d}t, \quad \Omega \in \{\mathbb{R}, \mathbb{R}_+, (a,b)\}.$$

Apply the Sinc-quadrature to the Laplace-type transform ⇒ separable approximation

$$f(\rho) = f(x_1 + \ldots + x_d) \approx \sum_{\nu=1}^R \omega_\nu G(t_\nu)e^{\rho F(t_\nu)} = \sum_{\nu=1}^R c_\nu \prod_{i=1}^d e^{x_i F(t_\nu)}, \quad c_\nu = \omega_\nu G(t_\nu).$$

Examples of $f(\rho)$: Green's kernels and classical potentials,

$$f(x) = \frac{1}{x_1 + \ldots + x_d}, \quad x_i \ge 0, \qquad \frac{1}{\rho} = \int_0^\infty e^{-\rho t}dt, \ \rho > 0.$$

$$f(x) = 1/\|x\|, \quad x \in \mathbb{R}^d, \qquad \frac{1}{\rho} = \frac{2}{\pi}\int_0^\infty e^{-\rho^2 t^2}dt, \quad \rho = \|x\|.$$

II. Separation by tensor-product interpolation (Tucker model)

- Tensor-product polynomial interpolation
- Tensor-product Sinc interpolation

Approximation by *sinc* interpolation (band-limited signals)

How to discretise analog signals ?

▶ The class of functions $f(t)$, $t \in \mathbb{R}$ can be discretized by recording their sample values $\{f(n\mathfrak{h})\}_{n\in\mathbb{Z}}$ at intervals $\mathfrak{h} > 0$.

Def. The sinc function (also called Cardinal function) is given by

$$\text{sinc}(x) := \frac{\sin(\pi x)}{\pi x} \quad \text{with convention } \text{sinc}(0) = 1.$$

V.A. Kotelnikov (1933) and J. Whittaker (1935) proved a celebrated theorem:

▶ *Band-limited signals* can be exactly reconstructed via their sampling values.

$$\widehat{f}(\omega) := \int_{\mathbb{R}} f(t) e^{-i\omega t} dt \quad \text{(continuous Fourier transform)}.$$

Thm. 7.3. (Sampling Theorem, Kotelnikov, Shannon, Whittaker)
If the support of $\widehat{f}$ *is included in* $[-\pi/\mathfrak{h}, \pi/\mathfrak{h}]$ *then for* $t \in \mathbb{R}$,

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\mathfrak{h}) S_{n,\mathfrak{h}}(t), \quad \text{with} \quad S_{n,\mathfrak{h}}(t) = \text{sinc}(t/\mathfrak{h} - n).$$

Proof. **Exer. 7.1.** Use properties of the Fourier transform (FT). [Khoromskij, Zurich-Lectures '2010].

Generalizing Sampling Theorem

Exer. 7.2. Let $\chi_{[-T,T]}(t) = 1$ if $t \in [-T, T]$ and 0 otherwise (characteristic, indicator, step function). Prove $\frac{1}{2T}\widehat{\chi} = \frac{\sin(T\omega)}{T\omega}$.

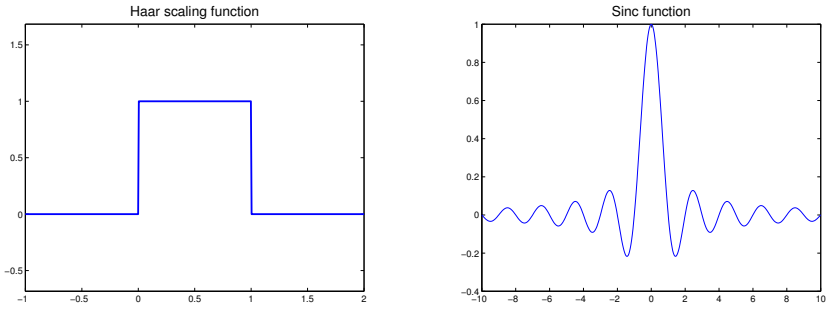


Figure: Haar (cf. $\widehat{f}$ of $f = \text{sinc}(x)$) and Sinc scaling functions.

Sampling theorem plays an important role in tele/radio communications, signal processing, stochastic models etc.

Def. The space $\mathbf{U}_{\mathfrak{h}}$ is a set of functions whose FTs have a support included in $[-\pi/\mathfrak{h}, \pi/\mathfrak{h}]$.

Lem. 7.4. [Stenger] A set of functions $\{S_{n,\mathfrak{h}}(t)\}_{n\in\mathbb{Z}}$ is an orthogonal basis of the space $\mathbf{U}_{\mathfrak{h}}$.

$$\text{For} \quad f \in \mathbf{U}_{\mathfrak{h}} : \quad f(n\mathfrak{h}) = \frac{1}{\mathfrak{h}} \left\langle f(t), S_{n,\mathfrak{h}}(t) \right\rangle.$$

sinc interpolation on Hardy space of analytic functions

Cor. 7.5. The *sinc*-interpolation formula of Thm. 7.1 can be interpreted as a decomposition of $f \in \mathbf{U}_\mathfrak{h}$ in an orthogonal basis of $\mathbf{U}_\mathfrak{h}$:

$$f(t) = \frac{1}{\mathfrak{h}} \sum_{n=-\infty}^{\infty} \langle f(\cdot), S_{n,\mathfrak{h}}(\cdot) \rangle \, S_{n,\mathfrak{h}}(t).$$

- If $f \notin \mathbf{U}_\mathfrak{h}$, one finds the orthogonal projection of f in $\mathbf{U}_\mathfrak{h}$.

When the Sinc-interpolant represents a function exactly?

$$C(f,\mathfrak{h})(x) = \sum_{k=-\infty}^{\infty} f(k\mathfrak{h}) S_{k,\mathfrak{h}}(x), \quad x \in \mathbb{R}.$$

Interpolant $C(f,\mathfrak{h})$ provides **an incredibly accurate approximation** on $\mathbb{R}$ for functions which are analytic and uniformly bounded on the strip

$$D_\delta := \{z \in \mathbb{C} : |\mathrm{Im}\ z| \le \delta\}, \quad 0 < \delta < \frac{\pi}{2}.$$

Def. Define the Hardy space $H^1(D_\delta)$ of functions which are analytic in D_δ and

$$N(f, D_\delta) := \int_{\mathbb{R}} (|f(x + i\delta)| + |f(x - i\delta)|)\, dx < \infty.$$

Approximation by *sinc*-interpolation and quadrattures

For $f \in H^1(D_\delta)$ we have exponential convergence in $1/\mathfrak{h}$ (Stenger)

$$\sup_{x\in\mathbb{R}} |f(x) - C(f,\mathfrak{h})(x)| = O(e^{-\pi\delta/\mathfrak{h}}), \quad \mathfrak{h} \to 0. \tag{8}$$

Likewise, if $f \in H^1(D_\delta)$, the integral

$$I(f) = \int_\Omega f(x)dx \quad (\Omega = \mathbb{R} \text{ or } \Omega = \mathbb{R}_+)$$

can be approximated by the Sinc-quadrature (trapezoidal rule)

$$T(f,\mathfrak{h}) := \mathfrak{h} \sum_{k=-\infty}^{\infty} f(k\mathfrak{h}) \quad \left(= \int_{\mathbb{R}} C(f,\mathfrak{h})(x)dx \approx I(f) \right),$$

$$|I(f) - T(f,\mathfrak{h})| = O(e^{-\pi\delta/\mathfrak{h}}), \quad \mathfrak{h} \to 0. \tag{9}$$

Analogues estimates hold for (computable) **trucated sums** (exponentially convergent)

$$C_M(f,\mathfrak{h}) := \sum_{k=-M}^{M} f(k\mathfrak{h}) S_{k,\mathfrak{h}}(x),$$

$$T_M(f,\mathfrak{h}) := \mathfrak{h} \sum_{k=-M}^{M} f(k\mathfrak{h}).$$

Exponential conergence rate for the truncated *sinc* interpolation/quadratures

Thm. 7.6. [Stenger] If $f \in H^1(D_\delta)$ and $|f(x)| \le C\exp(-b|x|)$ for all $x \in \mathbb{R}$ $b, C > 0$, then

$$\|f - C_M(f,\mathfrak{h})\|_\infty \le C\left[\frac{e^{-\pi\delta/\mathfrak{h}}}{2\pi\delta}N(f,D_\delta) + \frac{1}{b\mathfrak{h}}e^{-b\mathfrak{h}M}\right], \tag{10}$$

$$|I(f) - T_M(f,\mathfrak{h})| \le C\left[\frac{e^{-2\pi\delta/\mathfrak{h}}}{1-e^{-2\pi\delta/\mathfrak{h}}}N(f,D_\delta) + \frac{1}{b}e^{-b\mathfrak{h}M}\right]. \tag{11}$$

- For interpolation error (10), the choice

$$\mathfrak{h} = \sqrt{\pi\delta/bM}$$

implies the exponential convergence rate (usually we choose $\delta = \pi/2$)

$$\|f - C_M(f,\mathfrak{h})\|_\infty \le CM^{1/2}e^{-\sqrt{\pi\delta bM}}. \tag{12}$$

In fact, for the chosen $\mathfrak{h}$, the first term in the r.h.s. in (10) dominates, hence (12) follows.

- For the quadrature error (11), the "optimal" choice

$$\mathfrak{h} = \sqrt{2\pi\delta/bM}$$

yields

$$|I(f) - T_M(f,\mathfrak{h})| \le Ce^{-\sqrt{2\pi\delta bM}}. \tag{13}$$

Examples related to basic applications

- Low rank separable approximation of the multi-variate functions in $\mathbb{R}^d$

$$(a)\ \frac{1}{x_1^2+\ldots+x_d^2},\quad (b)\ \frac{1}{\sqrt{x_1^2+\ldots+x_d^2}},\quad (c)\ \frac{e^{-\lambda\|x\|}}{\|x\|},\quad \|x\| = \sqrt{x_1^2+\ldots+x_d^2}.$$

Example 7.3. In case (a), the Sinc method applies to the Laplace integral transform

$$\frac{1}{\rho} = \int_{\mathbb{R}_+} \mathrm{e}^{-\rho t}\mathrm{d}t \qquad (\rho = \|x\|^2 \in [1,R],\ R > 1)\,. \tag{14}$$

Exer. 7.3. Compute low-rank approximations to the Hilbert matrix (tensor).

Examlpe 7.4. In case (b), $\rho = \|x\|$, apply the Gauss integral ($1/\|x\|$ is the Newton kernel in $\mathbb{R}^3$)

$$\frac{1}{\rho} = \frac{2}{\sqrt{\pi}}\int_{\mathbb{R}_+} \mathrm{e}^{-\rho^2 t^2}\mathrm{d}t \qquad (\rho \in [1,R])\,. \tag{15}$$

- To maintain robustness in ρ, rewrite the Gauss integral (15) using substitutions $t = \log(1+\mathrm{e}^u)$ with $u = \sinh(w)$,

$$\frac{1}{\rho} = \int_{\mathbb{R}} f(w)\mathrm{d}w \quad \text{with} \quad f(w) := \cosh(w)F(\sinh(w)), \tag{16}$$

$$F(u) := \frac{2}{\sqrt{\pi}}\frac{\mathrm{e}^{-\rho^2\log^2(1+\mathrm{e}^u)}}{1+\mathrm{e}^{-u}}, \quad w, u \in (-\infty,\infty).$$

Low rank CP approximation to classical Green's functions, [Khoromskij, Bertoglio '08-'09]

▶ Elliptic Green's function via sinc-quadrature approximation (CP rank $R = 2M+1$):

$$\rho = \|x\| \geq h > 0: \quad \frac{1}{\rho} = \int_0^\infty e^{-t^2\rho^2} dt \approx \sum_{k=-M}^{M} c_k e^{-t_k^2(x_1^2+\ldots+x_d^2)} =: G_M.$$

The choice $t_k = e^{k\hbar}$, $c_k = \hbar t_k$, $\hbar = \pi/\sqrt{M}$, implies exponential convergence rate in M,

$$\left| \frac{1}{\|x\|} - G_M \right| \leq Ce^{-\pi\sqrt{M}}, \quad (\text{or } \leq Ce^{-\pi M/\log M},\ t_k = k\hbar,\ c_k = \hbar,\ \hbar = \frac{C_0 \log M}{M}).$$

▶ Slater function $e^{-\lambda\|x\|}$, $x \in \mathbb{R}^3$, represents typical singularity in quantum chemistry. For any $M = 1, 2, \ldots$, there is a sequence c_k, t_k (see above) s.t.

$$e^{-2\lambda\|x\|} = \frac{\lambda}{\sqrt{\pi}} \int_0^\infty t^{-3/2} e^{-\lambda^2/t} e^{-t\|x\|^2} dt \approx \sum_{k=-M}^{M} c_k e^{-t_k\|x\|^2} =: G_M,$$

$$\left| e^{-2\lambda\|x\|} - G_M \right| \leq Ce^{-\pi\sqrt{M}}, \quad (\text{or } \leq Ce^{-\pi M/\log M}).$$

▶ Similar low-rank approximations can be derived for $\frac{e^{-\lambda\|x\|}}{\|x\|}$, $\frac{e^{-i\lambda\|x\|}}{\|x\|}$. [Khoromskij '09]

Rank-**r** Tucker approximation to $1/\|x\|$, $d = 3$, $\|x\| \leq 10$.

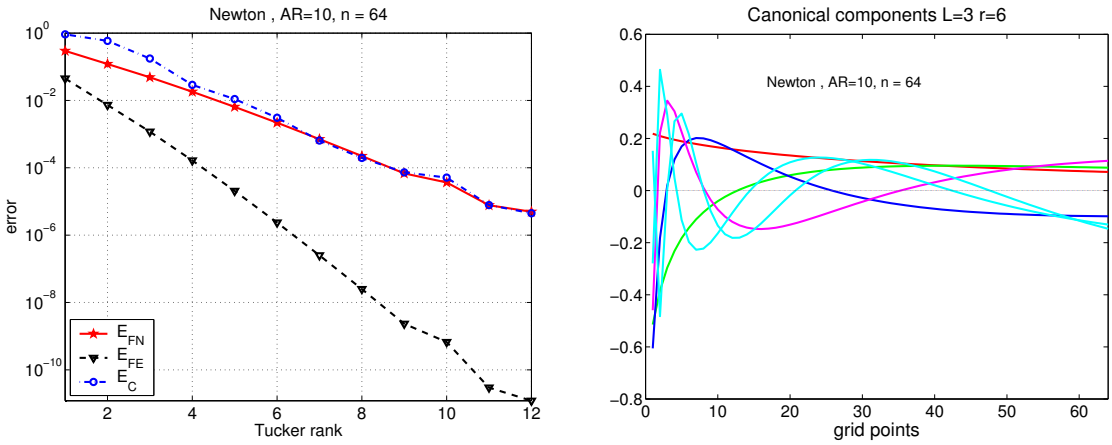


Figure: Convergence history for the Newton potential on $n \times n \times n$ grid.

Numerics to approximation of analytic functions with singularities

Rank-**r** Tucker approximation to $\exp(-\|x\|^\gamma)$, $d = 3$, $\|x\| \le 10$.

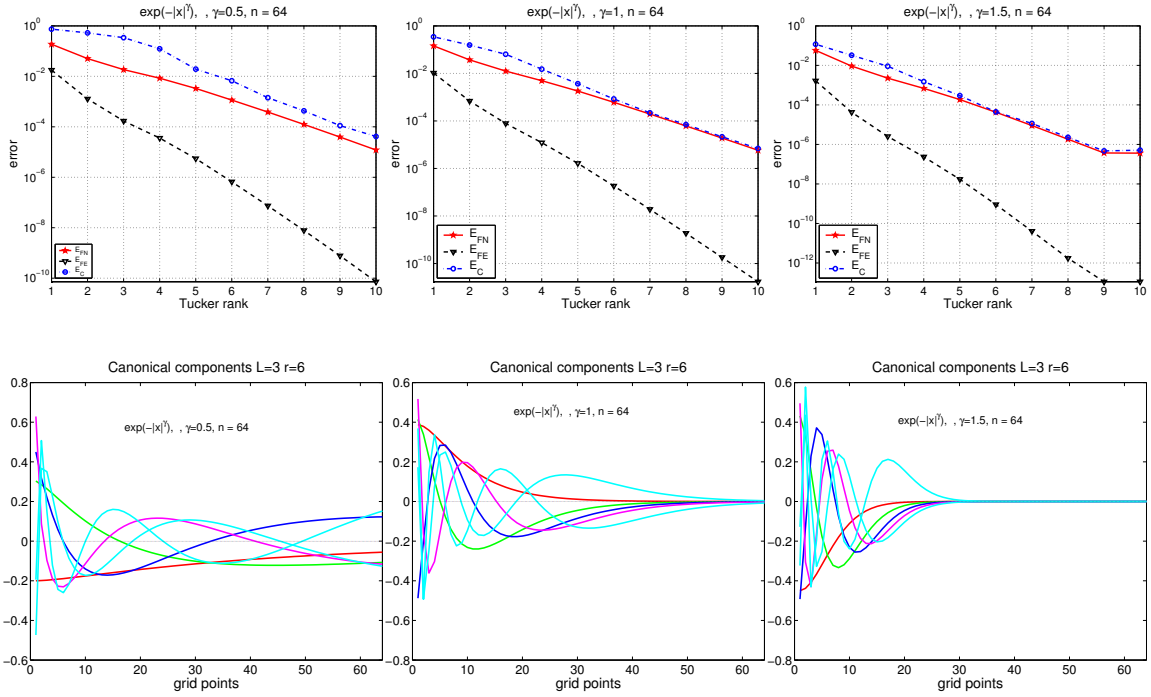


Figure: Orthogonal Tucker vectors for the γ-Slater potential on $n \times n \times n$ grid.

Matrix Product States (MPS) factorization:

In quantum physics, spin systems:

The matrix product states (**MPS**), (**MPO**) and tree-tensor network states (**TNS**)
[White '92; Fannes, Nachtergaele '92, Östlund, Rommer '95; ..., Cirac, Verstraete '06, ...].

Re-invented in numerical MLA:

Hierarchical dimension splitting, $O(dr^{\log d} N)$-storage: [Khoromskij '06].
Hierarchical Tucker (HT) $\equiv$ TNS: [Hackbusch, Kühn '09]
Tensor train (TT) $\equiv$ MPS (for open boundary conditions) [Oseledets, Tyrtyshnikov '09].

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

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

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

▶ $\mathbf{V} \in TT[\mathbf{r}]$ is represented by a product of matrices (matrix product states), each depending on a single "physical" mode: cf. Tucker with localized connectivity constraints.

▶ $d = 2$: TT is a skeleton factorization of a rank-r matrix: $A = UV^T$.

MPS-type dimension splitting: benefits of the MPS/TT format

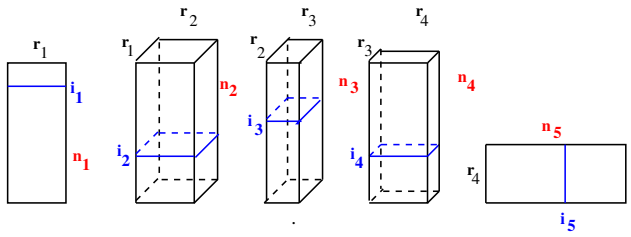
Rem. TT factorization can be derived from CP-format by RHOSVD [Khoromskij, Khoromskaia '09].

▶ For fixed $\mathbf{r} = [r_1, \dots, r_d]$ both Tucker and TT parametric representations in T[**r**] and TT[**r**] define a manifold ⇒ Dirac-Frenkel dynamics on low-parametric manifolds.

▶ Existence of best rank-**r** approximation: ALS/DMRG iteration.

▶ Stable quasi-optimal approximation by ℓ-mode SVD (Schmidt decomposition). Practically applicable only to the canonical or TT input tensor!

Visualizing MPS (TT) for $d = 5$: Contracted product of tri-tensors over $J_1 \times \dots \times J_5$.



Example 7.1. $f(x) = x_1 + \dots + x_d$. Explicit TT representation: $rank_{TT}(f) = 2$.

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

Main properties of the MPS (TT) representations

Def. $V_{[\ell]} := [V(i_1, \dots, i_\ell; i_{\ell+1}, \dots, i_d)]$ is the ℓ-mode TT unfolding matrix.

Thm. 7.7. (TT-tensors: Storage, rank bound, concatenation, quasioptimality).

(A) Storage: $\sum_{\ell=1}^{d} r_{\ell-1} r_\ell \, N \le d r^2 N$ with $r = \max_\ell r_\ell$.

(B) Rank bound: $r_\ell \le rank_{[\ell]}(\mathbf{V}) := rank(V_{[\ell]}) \le rank_{Can}(\mathbf{V})$, $r_1 = r_{1,Tuck}, r_{d-1} = r_{d,Tuck}$.

(C) Canonical embeddings:

$$\mathcal{C}_{R,\mathbf{n}} \subset TT[\mathbf{r}, \mathbf{n}, d] \quad \text{with} \quad \mathbf{r} = (R, \dots, R), \quad TT[\mathbf{r}] \subset TC[\mathbf{r}].$$

(D) Concatenation to higher dimension: $\mathbf{V}[d_1] \otimes \mathbf{V}[d_2] \to D = d_1 + d_2$ (look how).

(E) Quasi-optimal TT[**r**]-approximation $\mathbf{T}_*$ of $\mathbf{V} \in \mathbb{V}_\mathbf{n}$ exists and it satisfies

$$\min_{\mathbf{T} \in TT[\mathbf{r}]} \|\mathbf{V} - \mathbf{T}\|_F \le \Big(\sum_{\ell=1}^{d-1} \varepsilon_\ell^2\Big)^{1/2}, \quad \varepsilon_\ell = \min_{rank\, B \le r_\ell} \|V_{[\ell]} - B\|_F,$$

and $\mathbf{T}_*$ can be computed by QR/SVD (DMRG) algorithm.

(F) **Summary on rank bounds**

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

SVD-based approximation in tensor formats revisited

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 via matrix SVD:

- SVD or Schmidt decomposition: for matrices
- SVD-based (R)HOSVD: for Tucker and canonical tensors
- SVD-based ALS/DMRG iteration: for MPS/TT tensors
- ACA interpolation: heuristic approach for matrices and tensors.

Tucker ranks: $\mathcal{T}_{\mathbf{r}} := \{\mathbf{A} \in \mathbb{V}_{\mathbf{n}} : \operatorname{rank}\mathbf{A}_{(p)} \le r_p\}$,

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

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

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

Canonical (CP) rank can't be presented as the matrix rank! $\Rightarrow$ unstable approximation

Rank reduction in the canonical format: Reduced HOSVD: CP $\mapsto$ Tucker $\mapsto$ CP (ALS)

Example: TT decomposition of the function $\sin(\sum_{j=1}^d x_j)$

Example. 7.2. $f(x) := \sin(\sum_{j=1}^d x_j)$, $x \in \mathbb{R}^d$, has the explicit rank-2 TT factorization

$$f(x) = \begin{pmatrix} \sin x_1 & \cos x_1 \end{pmatrix} \begin{pmatrix} \cos x_2 & -\sin x_2 \\ \sin x_2 & \cos x_2 \end{pmatrix} \cdots \begin{pmatrix} \cos x_{d-1} & -\sin x_{d-1} \\ \sin x_{d-1} & \cos x_{d-1} \end{pmatrix} \begin{pmatrix} \cos x_d \\ \sin x_d \end{pmatrix}.$$

Proof. Induction, cf. Example 3.1,

$$\begin{aligned} f(x) &= \sin x_1 \cos(x_2 + \ldots + x_d) + \cos x_1 \sin(x_2 + \ldots + x_d) \\ &= \begin{pmatrix} \sin x_1 & \cos x_1 \end{pmatrix} \begin{pmatrix} \cos(x_2 + \ldots + x_d) \\ \sin(x_2 + \ldots + x_d) \end{pmatrix} \\ &= \begin{pmatrix} \sin x_1 & \cos x_1 \end{pmatrix} \begin{pmatrix} \cos x_2 & -\sin x_2 \\ \sin x_2 & \cos x_2 \end{pmatrix} \begin{pmatrix} \cos(x_3 + \ldots + x_d) \\ \sin(x_3 + \ldots + x_d) \end{pmatrix} \ldots \end{aligned}$$

Lem. 7.8. For any $d \ge 3$, $\varepsilon > 0$, we have for the high-frequency Helmholtz kernels,

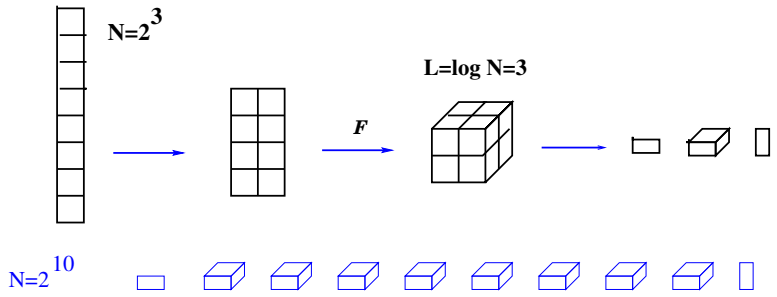
$$rank_{TT,\varepsilon}(f_{1,\kappa}) \le C(|\log\varepsilon| + \kappa), \quad f_{1,\kappa}(\|x\|) := sin(\kappa\|x\|)/\|x\|,$$

$$rank_{TT,\varepsilon}(f_{2,\kappa}(\|x\|)) \le Crank_{Can}(\|x\|)|\log\varepsilon|(|\log\varepsilon|+\kappa)\log n, \quad f_{2,\kappa}(\|x\|) := \frac{2sin^2(\frac{\kappa}{2}\|x\|)}{\|x\|}.$$

Hint: Follows from rank bounds in Thm. 7.7, (F). [Khoromskij '09].

The quantized image of functional vectors: Tensor approx. in higher virtual dimensions

- Quantized TT (QTT) approximation of functional N-vectors ($N = 2^L$). [Khoromskij '2009]



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

$$i \mapsto \mathbf{i} \in \{1,2\}^{\otimes L} : \quad i - 1 = \sum\nolimits_{\ell=1}^{L} (i_\ell - 1) 2^{\ell-1}.$$

Canonical/TT approximation of quantized L-dimensional image in $\mathbb{Q}_L$
$\Rightarrow$ QCan/QTT method

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

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

- Numerical observation: $2^L \times 2^L$ Laplacian reshapes to a low TT-rank. [Oseledets '09]

Reshaping to high-dimensional Q-image via q-adic coding

- $N = q^L$, $q = 2, 3, 5, \ldots$. Standard choice $q = 2$: binary coding. $q_{opt} = e \approx 2,7\ldots$.

Def. [Khoromskij '09] $N = q^L$, $q = 2, 3, 5, \ldots$. QTT as the q-adic folding of degree $L = \log_q N$.

- $d = 1$: a vector $\mathbf{X} = [x_i]_{i=1}^N \in \mathbb{R}^N$, is reshaped to L-dimensional tensor,

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

Fixed i, the $\mathbb{Q}$-multiindex $\mathbf{j} - \mathbf{1} \in \{0, 1, ..., q\}^{\otimes L}$ is defined via q-adic coding of $i - 1$,

$$i - 1 = \sum\nolimits_{\ell=1}^{L} (j_\ell - 1) q^{\ell-1}, \quad j_\ell - 1 \in \{0, 1\}.$$

- $d \geq 2$: multivariate reshaping.
- **Generalization**: Decomposition into smallest nontrivial prime factors $N = q_1 q_2 \ldots q_L$. The corresponding index factorization, say, $N = 30 = 2 \cdot 3 \cdot 5$, allows the QTT format.

Quantization (folding) of **a vector/tensor** to higher dimension leads to super-compressed representation of functions and operators,

$$N^d \to O(d \log_q N).$$

Numerical methods in **QTT format** lead to super-fast PDEs solvers at log-cost.

Approximation power by the QTT method for functional vectors: basic results

Thm. 7.9. [Khoromskij '09]. *QTT*-approximation of functional vectors, $N = 2^L$.

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

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

- For the quantized sin N-vector $\mathbf{X}$ (same for cos): $rank_{QTT}(\mathbf{X}) = 2$,

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

- Proof. **Hint:** $\sin z = \frac{e^{iz} - e^{-iz}}{2i} = Im(e^{iz})$.
- For QTT-image of polynomial of degree m we have $rank_{QTT}(P_m) \le m + 1$.
- 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

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

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

- Gaussian on quadratic grid $\mathbf{G} = \{e^{-pt_n^2}\}$, $t_n = \sqrt{h(n-1)}$: $rank_{Can}(\mathbf{G}) = 1$.

Function of form $f(x) = \sum_{\ell=1}^{d} f^{(\ell)}(x_\ell)$

- For Gaussian $g(x) := e^{-x^2/2p^2}$, $x \in [-a, a]$,

$$rank_{QTT}(\mathbf{G}) \le c\frac{a}{p}\sqrt{\log(\varepsilon^{-1}\frac{p}{1+a})}.$$

Proof. The Fourier transf. of $g(x)$ + $\text{rank}_{QTT}(\cos) = 2$: $\int_{\mathbb{R}} g(x)\cos(\omega x)dx = pe^{-\omega^2 p^2/2}$.

- Rank decomposition of $f(x) = f_1(x_1) + f_2(x_2) + \ldots + f_d(x_d)$,

$$f(x) = \begin{pmatrix} f_1(x_1) & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ f_2(x_2) & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ f_{d-1}(x_{d-1}) & 1 \end{pmatrix} \begin{pmatrix} 1 \\ f_d(x_d) \end{pmatrix}.$$

$\text{Rank}_{Can}(f) = d$,

$\text{Rank}_{Tuck}(f) = \text{Rank}_{TT}(f) = 2$,

ℓ-mode QTT-rank: $\text{Rank}_{\ell,QTT}(f) \le 1 + \text{Rank}_{QTT}(f_\ell)$, $\ell = 1, \ldots, d$.

- Harmonic potential: QTT-ranks are bounded by 4,

$$V(\mathbf{q}) = \sum_{k=1}^{d} w_k q_k^2, \quad rank_{TT}(V) \le 2, \ \ rank_{QTT}(V) \le 4$$

Numerics: QTT approximation of functional tensors

$$\text{Average QTT-rank:} \quad \overline{r}^2 = \frac{1}{L}\sum_{\ell=1}^{L} r_{\ell-1} r_\ell,$$

$$\text{Storage} \le 2L\overline{r}^2 \log N.$$

Function-related N-vector: $\mathbf{F} = \{f(a + (i - \frac{1}{2})h)\}_{i=1}^{N}$, $h = \frac{b-a}{N}$, $\varepsilon = 10^{-6}$

$N \setminus \overline{r}$	$e^{-\alpha x^2}$, $\alpha = 0.1 \div 10^2$	$\frac{\sin(\alpha x)}{x}$, $\alpha = 1 \div 10^2$	$1/x$	e^{-x}/x	$x, x^{10}, \sqrt[10]{x}$
2^{12}	3.1/2.9/2.9/2.6	3.8/4.8/5.6	4.2	3.8	1.9/2.6/3.9
2^{14}	2.9/2.8/2.8/2.8	3.6/4.7/5.5	4.2	3.8	1.9/2.5/3.9
2^{16}	2.8/2.7/2.8/2.8	3.6/4.5/5.4	4.2	5.3	1.9/2.4/3.9

$N \setminus \overline{r}$	$1/(x_1 + x_2)$	$e^{-\|x\|}$	$e^{-\|x\|^2}$	$\Delta_2^{-1} 1$, $\varepsilon = 10^{-6}, 10^{-7}, 10^{-8}$
2^9	5.0	9.4	7.8	3.6/3.6/3.6
2^{10}	5.1	9.4	7.7	3.6/3.6/3.6
2^{11}	5.2	9.3	7.5	3.7/3.7/3.7

Super-compression in high dimension?

Exer. 7.1. Linear-log-log scaling via quantics in auxiliary dimension:

dth order Hilbert N-d tensor $\mathbf{A}$ of dimension $N^{\otimes d}$, $N = 2^L$,

$$a(i_1, ..., i_d) = \frac{1}{i_1 + i_2 + ... + i_d} \approx \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} e^{-t_k i_\ell},$$

$i_1, ..., i_d = 1, ..., N$, can be approximated by a rank-$|\log\varepsilon|$ canonical tensor of order $D = d \log_2 N$ and size $2^{\otimes D}$, requiring only

$$Q = d|\log\varepsilon| \log N \ll N^d \quad \text{reals to store it.}$$

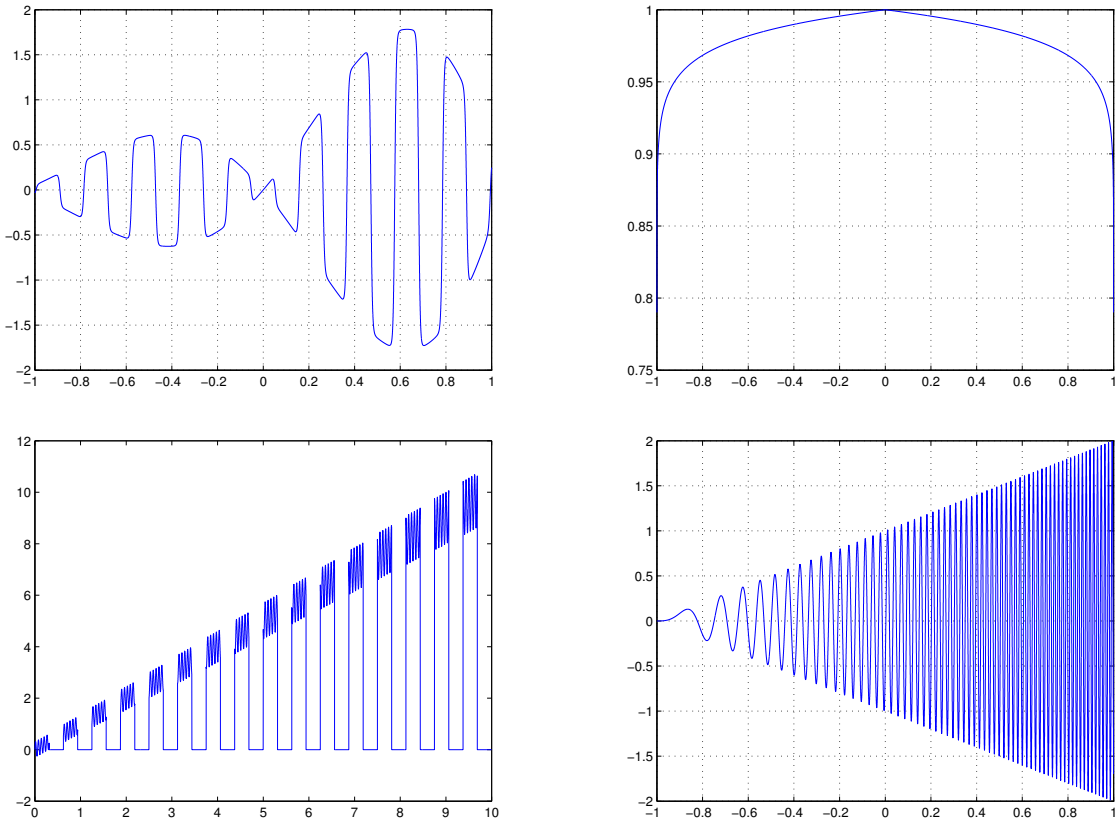
Using our canonical decomposition, compute its QTT approximation applying C-to-QTT.

Computational gain:

Matrix case: $d = 2$, $N = 2^{20} \Rightarrow Q = 40|\log\varepsilon| \ll 2^{40}$.

High dimension: $d = 2^{10}$, $N = 2^{20} \Rightarrow Q = 20 \cdot 2^{10}|\log\varepsilon| \ll 2^{2\cdot 10^4}$.

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



QTT based quadratures for highly oscillating and singular functions

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

The rectangular 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 oscillating 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).

$n \setminus r$	$r_{QTT}(f_1)$	$r_{QTT}(f_2)$	$r_{QTT}(f_3)$	$r_{QTT}(f_4)$
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

Summary: Canonical/Tucker/TT/QTT approximation of functional tensors

Piece of theory.

- Exponential, polynomials, wavelets, sum/product of them: $O(\log N)$ complexity.
- Gaussian type and highly oscillating functions: $O(|\log \varepsilon| \log N)$ complexity.
- $f(x+y)$ separable with low rank $\Rightarrow$ low QTT rank.
- Multivariate functions in the form $f(x_1 + \ldots + x_d)$ inherit QTT ranks of $f(t)$.

Recent applications.

- Tucker/TT/QTT: Hartree-Fock $\Rightarrow$ Green functions, two-electron integrals (TEI), Hartree and exchange potentials, electron density, molecular orbitals. Many particles electrostatic potentials in range-separated formats. High-dim. integration.
- QTT: sPDEs, QMD (PES), FCI electronic structure, chemical master eq.
- QTT representation to highly oscillating functions (geometric homogenization).
- The Bethe-Salpeter equation (BSE) for excitation energies, density of states.

Limitations.

- "Curse of ranks", dominance of rank-truncation (hope on QTT-Tucker)
- Schrödinger, Hartree-Fock, Fokker-Planck hamiltonians are not (naively) separable :-(

TT/QTT representation of operators (MPO)

Def. Matrix product operators (MPO): A multi-way TT/QTT-matrix is defined by

$$\mathbf{A} : \mathbb{X} := \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_d} \mapsto \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_d} =: \mathbb{Y}$$

$$\begin{aligned} \boldsymbol{A}(i_1, j_1, \ldots, i_d, j_d) &= \sum_{\alpha_1=1}^{r_1} \ldots \sum_{\alpha_{d-1}=1}^{r_{d-1}} U_1(i_1, j_1, \alpha_1)\, U_2(\alpha_1, i_2, j_2, \alpha_2) \cdot \ldots \cdot \\ &\cdot \; U_{D-1}(\alpha_{d-2}, i_{d-1}, j_{d-1}, \alpha_{d-1})\, U_D(\alpha_{d-1}, i_d, j_d), \end{aligned}$$

where $U_k(i_k, j_k)$ is a $r_{k-1} \times r_k$ matrix.

- Two approaches to define the tensor rank of a multi-way matrix (operator).

Def. For $\mathbf{X} \in \mathbb{X}$ denote by $r_1 \ldots r_{d-1}$ the TT ranks of the matr.-by-vect. prod. $\mathbf{AX} \in \mathbb{Y}$.

▷ The *operator TT rank* of $\boldsymbol{A}$ is defined by

$$\max_{\substack{k=1,\ldots,d-1, \\ \mathbf{X} \text{ is of vector TT rank } 1\ldots1}} r_k(\mathbf{AX}).$$

▷ *k-th vector TT rank* of $\boldsymbol{A}$ is the rank of its TT unfolding $A_{[k]}$ with entries

$$A_{[k]}(i_1 j_1 \ldots i_k j_k \;;\; i_{k+1} j_{k+1} \ldots i_d j_d) = \boldsymbol{A}(i_1 j_1 \ldots i_d j_d), \quad k = 1, .., d-1.$$

TT/QTT representation (approximation) of the elliptic operators

Example. d-dimensional discrete FDM 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: $\mathrm{rank}_{CP}(\Delta_d) = d$, $\mathrm{rank}_{Tuck}(\Delta_d) = 2$.
- Explicit TT representation: $\mathrm{rank}_{TT}(\Delta_d) = 2$, $\mathrm{rank}_{QTT}(\Delta_d) \le 4$, $\forall d$.

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

- Explicit QTT representation: $\mathrm{rank}_{QTT}(\Delta_1) = 3$, $\mathrm{rank}_{QTT}(\Delta_1^{-1}) \le 5$,

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

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

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

Collection of CP/TT/QTT-rank estimates for Δ_d-related matrices

Lem. 7.10. TT/QTT rank estimates hold:

- Explicit representations hold true

$$\mathrm{rank}_{QTT}(\Delta_1) = 3, \quad \mathrm{rank}_{QTT}(\Delta_1^{-1}) \le 5.$$

$$\mathrm{rank}_{TT}(\Delta_d) = 2, \quad \mathrm{rank}_{QTT}(\Delta_d) = 4.$$

- Matrix valued exponential function:

$$\mathrm{rank}_{CP}(e^{-\Delta_d}) = \mathrm{rank}(e^{-\Delta_1} \otimes e^{-\Delta_2} \otimes \cdots \otimes e^{-\Delta_d}) = 1.$$

- ε-rank:

$$\mathrm{rank}_{TT}(\Delta_d^{-1}) \le \mathrm{rank}_{CP}(\Delta_d^{-1}) \le C |\log \varepsilon| \log N.$$

- ε-rank:

$$\mathrm{rank}_{QTT}(\Delta_d^{-1}) \le C |\log \varepsilon|^2 \log N.$$

- Variable coefficients: 1D FEM elliptic operator (stiffness matrix of div $a(x)$ grad)

$$\mathrm{rank}_{QTT}\left(\nabla^T \mathrm{diag}\{a\} \nabla\right) \le 7 \, \mathrm{rank}_{QTT}(a).$$

Fast Fourier Transform (FFT)

Let S_N be the space of sequences $\{f[n]\}_{0\le n<N}$ of period N (in $\mathbb{R}^N$ or $\mathbb{C}^N$).

S_N is an Euclidean space, $\langle f, g\rangle = \sum_{n=0}^{N-1} f[n]g^*[n]$.

Def. 1.3. The *discrete Fourier transform* (DFT) of f is

$$\widehat{f}[k] := \langle f, e_k\rangle = \sum_{n=0}^{N-1} f[n]\exp\left(\frac{-2i\pi kn}{N}\right), \quad (N^2 \text{ complex multiplications}).$$

The DFT matrix $F_N = \{f_{k,n}\}_{k,n=1}^{N}$ is given by

$$f_{k,n} := \exp(\frac{-2i\pi kn}{N}) = W^{-nk}, \quad W = e^{2i\pi/N}.$$

- Fast Fourier Transform (FFT) in $C_F N\log_2 N$ operations, $C_F \approx 4$.

The FFT traces back (1805) to Gauss (1777 - 1855).

The first computer program Coolly/Tukey (1965).

- Fast wavelet transform (FWT) in $O(N)$ operations.
- QTT-tensor based Super-fast FFT and FWT in $O(\log^2 N)$ operations !

Superfast QTT-FFT (another direction: Super-fast wavelet transform (FWT)

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

$$a(i,j) = a(\overline{j_1\ldots j_d}, \overline{k_1\ldots k_d}) = \mathbf{A}(j_1k_1, j_2k_2, \ldots, j_dk_d) = A^{(1)}_{j_1k_1}A^{(2)}_{j_2k_2}\ldots A^{(d)}_{j_dk_d}$$

QTT ranks

$$r_p = \mathrm{rank}\,\mathbf{A}_{[p]}(\underbrace{j_1k_1\ j_2k_2\ \ldots j_pk_p}_{\text{column index}}; \underbrace{j_{p+1}k_{p+1}\ldots j_dk_d}_{\text{row index}})$$

QTT decomposition of FFT matrix has full rank :(
QTT-FFT matrix has full ε-rank $\Leftrightarrow$ The low-rank ε-approximation is not possible :(
[Dolgov, Khoromskij, Savostianov, J. Fourier Anal. Appl., 2012]

Example. The Hadamard (Walsh) transform has QTT–ranks one,

$$H_d = H_1^{\otimes d} \equiv \underbrace{H_1\otimes H_1\ldots\otimes H_1}_{\text{d times}}, \qquad H_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}.$$

Cooley-Tuckey FFT in QTT format

Fourier transform $y = FFT_n(x)$, $n = 2^d$.

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

FFT for dense vectors costs $O(n \log n)$

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

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

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.,

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

Fourier images in 1D: QTT-FFT vs. FFTW

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

Algebra of circulant matrices

Def. A one-level block circulant matrix $A \in \mathcal{BC}(L, m_0)$ is defined by

$$A = \mathrm{bcirc}\{A_0, A_1, ..., A_{L-1}\} = \begin{bmatrix} A_0 & A_{L-1} & \cdots & A_2 & A_1 \\ A_1 & A_0 & \cdots & \vdots & A_2 \\ \vdots & \vdots & \ddots & A_0 & \vdots \\ A_{L-1} & A_{L-2} & \cdots & A_1 & A_0 \end{bmatrix} \in \mathbb{R}^{Lm_0 \times Lm_0}, \qquad (17)$$

where $A_k \in \mathbb{R}^{m_0 \times m_0}$ for $k = 0, 1, \ldots, L-1$, are matrices of general structure. The equivalent Kronecker product representation is defined by the associated matrix polynomial,

$$A = \sum_{k=0}^{L-1} \pi^k \otimes A_k =: p_A(\pi), \qquad (18)$$

where $\pi = \pi_L \in \mathbb{R}^{L \times L}$ is the periodic downward shift (cycling permutation) matrix,

$$\pi_L := \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \qquad (19)$$

Diagonalizing circulant matrix

In the case $m_0 = 1$ a matrix $A \in \mathcal{BC}(L, 1)$ defines a circulant matrix generated by its first column vector $\mathbf{a} = (\mathbf{a_0}, ..., \mathbf{a_{L-1}})^{\mathsf{T}}$. The associated scalar polynomial then reads

$$p_A(z) := a_0 + a_1 z + ... + a_{L-1} z^{L-1},$$

so that (18) simplifies to

$$A = p_A(\pi_L).$$

Let $\omega = \omega_L = \exp(-\frac{2\pi i}{L})$, we denote the unitary matrix of Fourier transform by

$$F_L = \{f_{k\ell}\} \in \mathbb{R}^{L\times L}, \quad \text{with} \quad f_{k\ell} = \frac{1}{\sqrt{L}} \omega_L^{(k-1)(\ell-1)}, \quad k, l = 1, ..., L.$$

Since the shift matrix π_L is diagonalizable in the Fourier basis,

$$\pi_L = F_L^* D_L F_L, \quad D_L = \mathrm{diag}\{1, \omega, ..., \omega^{L-1}\}, \tag{20}$$

the same holds for any circulant matrix,

$$A = p_A(\pi_L) = F_L^* p_A(D_L) F_L, \tag{21}$$

where

$$p_A(D_L) = \mathrm{diag}\{p_A(1), p_A(\omega), ..., p_A(\omega^{L-1})\} = \mathrm{diag}\{F_L \mathbf{a}\}.$$

Matrix-vector product in $O(L \log L)$ operations

$$A\mathbf{x} = F_L^* p_A(D_L) F_L \mathbf{x} = F_L^* \left(\mathrm{diag}\{F_L \mathbf{a}\}(F_L \mathbf{x})\right).$$

Discrete circulant/Toeplitz convolution

Def. g is the discrete convolution of signals f, h supported by the indices $0 \le n \le M-1$,

$$g_n = (f * h)_n = \sum_{k=-\infty}^{\infty} f_k h_{n-k}.$$

The naive implementation requires $M(M+1)$ operations.

It can be represented as a matrix-by-vector product (MVP) with the Toeplitz matrix

$$g = f * h = Tf: \qquad T = \{h_{n-k}\}_{0\le n,k<M} \in \mathbb{R}^{M\times M}.$$

Extending f and h with over M samples by

$$\tilde{h}_M = 0, \quad \tilde{h}_{2M-i} = h_i, \quad i = 1, ..., M-1,$$

$$\tilde{f}_n = 0, \quad n = M, ..., 2M-1,$$

we reduce the problem to the MVP with a circulant matrix $\mathcal{C} \in \mathbb{R}^{2M\times 2M}$ specified by the first row $\tilde{h} \in \mathbb{R}^{2M}$.

The latter can be multiplied with a vector by FFT algorithm (diagonalization by DFT).

- Toeplitz/circulant type matrices apply to quasi-periodic systems, in homogenization.

Summary: Tensor representation of operators

Constructive results.

- *sinc*-quadrature representation of A^{-1}, e^A, Green's functions.
- Explicit Tucker, TT, QTT representation of Δ_d related operators.
- PES (Henon-Heiles potential), spin Hamiltonians.
- d-dimensional convolution: canonical/Tucker, explicit QTT in $O(d \log N)$ complexity.
- d-dimensional QTT-FFT, QTT-FWT, $O(d \log N)$ complexity.

Recent applications.

- Operators in (post) Hartree-Fock eqn., lattice-structured systems, master eqn., QMD, sPDEs, geometric homogenization, many-particle interaction potentials, the Bethe-Salpeter eqn., density of states, the Poisson-Boltzmann eqn. for proteins, ...

Limitations.

- "Curse" of ranks, high cost of rank reduction, tensor representation of the Schrödinger and Fokker-Planck Hamiltonians, log-additive case of sPDEs, non-rectangular geometries (IGA), stochastic homogenization, tensor algebra in the new formats.

Tensor numerical methods: algebraic ingredients and main targets

1. Discretization in tensor-product Hilbert space of N-d tensors,

$$\mathbf{V} = [V(i_1, ..., i_d)] \in \mathbb{V}_{\mathbf{n}} = \mathbb{R}^{n_1 \times \cdots \times n_d}, \quad n_k = N.$$

2. MLA in rank-$\mathbf{r}$ tensor formats $\mathcal{S} \subset \mathbb{V}_{\mathbf{n}}$:

$$\mathcal{S} \subset \{\mathcal{C}_R,\ \mathcal{T}_{\mathbf{r}},\ \mathcal{T}_{\mathcal{C}_R,\mathbf{r}},\ TT/TC[\mathbf{r}],\ QTT[\mathbf{r}]\}, \quad \mathbf{r} = [r_1, \ldots, r_d].$$

- Tensor truncation (projection), $T_{\mathcal{S}} : \mathcal{S}_0 \to \mathcal{S} \subset \mathcal{S}_0 \subset \mathbb{V}_{\mathbf{n}}$, based on

$$\text{SVD} + \text{(R)HOSVD} + \text{ALS/DMRG} + \text{multigrid}.$$

- Scalar/Hadamard/contracted/convolution products on $\mathcal{S}$.

3. $\mathcal{S}$-tensor approximation of functions and operators.

4. Tensor-truncated solvers on low-parametric manifold $\mathcal{S}$:
- Multigrid $\mathcal{S}$-truncated preconditioned iteration.
- Direct minimization on $\mathcal{S}$:

  ALS/DMRG, in CP, Tucker, MPS (TT), QTT formats and their generalizations.

- Direct $\mathcal{S}$-tensor solution operators via A^{-1}, $\exp(tA)$, Green's functions.