Multi-scale tensorization - the blessing of dimensions

R. Schneider (TUB Matheon)

Announcement

International Focus Workshop on **Entanglement Based Approaches in Quantum Chemistry**

to be held from 04 - 06 September 2012 in Dresden.
Max-Planck-Institut für Physik komplexer Systeme

http://events.mipks-dresden.mpg.de/register/EBAQC12/registration

Information: http://www.mipks-dresden.mpg.de/ebaqc12
Sparsity versus rank sparsity
So far, concepts of **adaptivity** are based on **sparsity**.

1. Sparsity depends on the underlying bases (or frame, dictionary).
2. But, how to find and construct appropriate bases (hp FE, Fourier, Gaussians, splines, wavelets, *dreamlets* ...)
3. Adaptive concepts (best $N$-term approx.) are well understood, but often difficult to implement.

**Rank sparsity** (in $2D$) is provided low rank approximation, e.g. SVD (singular value decomposition)

$$
A(x_1, x_2) \approx \sum_{k=1}^{r} \sigma_k (u_k(x_1) \otimes v_k(x_2))
$$

In this case, the **non-zero entries** $\sigma_k$ together with the bases $u_k \otimes v_k$ are **adaptively** computed (*dreamlets*).

Examples: matrix completion, reduced basis functions, etc.

In the present talk, we pursue the second kind of adaptivity.
I.

Best one term approximation!

Does it make sense?

yes-

each function has its own multi-scale transform!
by adaptive multi-wavelet transformation
(subdivision scheme bored from Alperts multi-wavelets)
I. Best one term approximation! Does it make sense?

yes-
each function has its own multi-scale transform!
by adaptive multi-wavelet transformation
(subdivision scheme bored from Alpert's multi-wavelets)
Polynomial Multi-Wavelets - revisited

Two scale relations, subdivision scheme for discont. multiwavelets (Alpert et al.), orth. polynomials

\[ f^0_{k_0}(x) = \sum_{\mu_0=0,1}^{r_1} \sum_{k_1=1}^{r_1} U_0(k_0, \mu_0, k_1)f^1_{k_1}(2x + \mu_0), \mu_0 = 0, 1. \]

\[ f^1_{k_1} = \sum_{\mu_1=0,1}^{r_2} \sum_{k_2}^{r_2} U_1(k_1, \mu_1, k_2)f^2_{k_2}(2x + \mu_1), \mu_1 = 0, 1. \]

\[ f^i_{k_i} = \sum_{\mu_i=0,1}^{r_i} \sum_{k_i}^{r_i} U_i(k_i, \mu_i, k_{i+1})f^{i+1}_{k_{i+1}}(2x + \mu_i), \mu_i = 0, 1. \]

\[ \ldots \]

\[ f^{d-1}_{k_{d-1}} = \sum_{\mu_{d-1}=0,1}^{r_{d-1}} \sum_{k_{d-1}}^{r_{d-1}} U_{d-1}(k_{d-1}, \mu_{d-1}, k_d)f^d_{k_d}(2x + \mu_{d-1}) \]

Filter coefficients \( U_i(k_i, \mu_i, k_{i+1}), i = 0, \ldots, d - 1 \) defines the scaling functions \( f^0_{k_0}(x) \)
i.e. the function \( f \) will be encoded by its filter coefficients

\( f \) in a non-stationary subdivision scheme.
Subdivision - Discrete version

\[ f_{k_0}(x) = \sum_{n=0}^{2^d-1} F(k_0, n) f^d_{k_d}(2^d x - n), \]

\( f^d \) ON scaling function. binary representation \( n = 0 \ldots, 2^d - 1 \),

\[ n = \sum_{j=0}^{d-1} \mu_j 2^{d-j-1}, \quad \mu_j = 0, 1, \]

\[ F(k_0, n) = F(k_0, \sum_{j=0}^{d-1} \mu_j 2^{d-j-1}) =: U(k_0, \mu_0, \ldots, \mu_{d-1}). \]

\((k_0, \mu) \mapsto U(k_0, \mu_0, \ldots, \mu_{d-1}) = \]

\[ = \sum_{k_0=1}^{r_0} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} U_0(k_0, \mu_0, k_1) U_1(k_1, \mu_1, k_2) \cdots U_{d-1}(k_{d-1}, \mu_{d-1}, k_d) \]

Discrete version \( f^d_{k_d}(x) \approx \delta_{0,n} \).
Two-scale decomposition-similarity rank

Figure: Function $f(x) = x^2 \sin \frac{8x}{2\pi}$ and functions $f_k^1$, $k = 0, 1$

first half interval $f_0^1$

second half interval $f_1^1$
Similarity rank and Multi-scale decomposition

Figure: $f_k^2$, $k = 0, 1, 2, 3$

POD (proper orthogonal decomposition (SVD)) ⇒
Similarity rank and Multi-scale decomposition

Figure: $f_k^2, k = 0, 1, 2, 3$, $(f_3^2 = 0)$

Recursively performed SVD from level to level $\Rightarrow$
Two-scale decomposition-similarity rank

Figure: Function \( f(x) = x^2 \sin \frac{8x}{2\pi} \) and translated functions \( f_k^3 \)

Figure: functions \( f_k^3, k = 0, \ldots 7 \)
Similarity rank and Multi-scale decomposition

Figure: POD (SVD) functions $f_k^3$, $k = 0, \ldots, 2, (\ldots, 7)$

Recursively performed SVD from level to level $\Rightarrow$ multiscale decomposition, i.e. a TT representation of $U$. 
Higher order SVD (HOSVD)

- Vidal (2003), Oseledets (2009), Grasedyck (2009), Kühn (2012)

1. Matriziation - unfolding

\( (\mu_0, k_0 \text{ i row index, } \mu_1, \ldots, \mu_{d-1} \text{ - column index}) \)

\[
U(k_0, \mu_0, \ldots, \mu_{d-1}) \approx U_{k_0, \mu_0}^{\mu_1, \ldots, \mu_{d-1}}.
\]

decompose by SVD, with orthogonal matrices \( (u_{k_0, \mu_0}^{k_1}) \),

\[
u_{k_0, \mu_0}^{k_1, \ldots, k_{d-1}} = \sum_{k_1=0}^{r_0} u_{k_0, \mu_0}^{k_1} v_{k_1, \mu_1}^{k_1, \ldots, k_{d-1}},
\]

\[
U_0(k_0, \mu_0, k_1) := u_{k_0, \mu_0}^{k_1}, \quad k_1 = 1, \ldots, r_0 \text{ (} \leq 2 \text{)}.
\]

2. Decompose \( V(k_1, \mu_1, \ldots, \mu_{d-1}) \) via matricisation

\[
v_{k_1, \mu_1}^{k_2, \ldots, k_{d-1}} = \sum_{k_2=1}^{r_1} u_{k_1, \mu_1}^{k_2} v_{k_1, \mu_1}^{k_2, \ldots, k_{d-1}},
\]

\[
U_1(k_1, \mu_1, k_2) := u_{k_1, \mu_1}^{k_2}.
\]

3. repeat with \( V(k_{d-1}, \mu_{d-2}, \mu_{d-1}) \) until one ends with

\[
\begin{bmatrix} v_{k_{d-1}, \mu_{d-1}}^{k_d} \end{bmatrix} \mapsto U_{d-1}(k_{d-1}, \mu_{d-1}, k_d).
\]
What do we gain?

For \( n := \#\{0, 1\} = 2 \), \( r := \max\{r_i : 0 \leq i \leq d - 2\} \), storage complexity is bounded by

\[
\text{DOFs: } \leq 2dr^2, \text{ instead of } 2^d
\]

Compression: Whenever \( r = O(1) \) then complexity \( N = 2^d \) is reduced to \( O(\log N) \)!

The optimal ranks \( r_i, r = (r_1, \ldots, r_{d-1}) \) (multi-rank), of the above decomposition are equal to the

\[
i\text{-th separation rank } r_i \text{ of } U,
\]

i.e. the ranks of the following matrices, \( i = 1, \ldots, d - 1 \)

\[
\text{rank of } A_i = U_{\mu_0, \ldots, \mu_i}^{\mu_{i+1}, \ldots, \mu_{d-1}} = r_{i+1}.
\]
I. Vector tensorization and tensor product approximation

Can we work with individual filter coefficients?

\[ \begin{array}{c}
U_1 & r_1 & U_2 & r_2 & U_3 & r_3 & U_4 & r_4 & U_5 \\
\bullet & \bullet & \bullet & \bullet & \bullet
\end{array} \]

\[ \begin{array}{c}
n_1 & n_2 & n_3 & n_4 & n_5
\end{array} \]

yes-
in novel tensor formats
I. Vector tensorization and tensor product approximation

Can we work with individual filter coefficients?

$U_1 \quad r_1 \quad U_2 \quad r_2 \quad U_3 \quad r_3 \quad U_4 \quad r_4 \quad U_5$

$n_1 \quad n_2 \quad n_3 \quad n_4 \quad n_5$
I.

Vector tensorization and tensor product approximation

Can we work with individual filter coefficients?

\[ U_1 \quad r_1 \quad U_2 \quad r_2 \quad U_3 \quad r_3 \quad U_4 \quad r_4 \quad U_5 \]

\[ n_1 \quad n_2 \quad n_3 \quad n_4 \quad n_5 \]

yes-
in novel tensor formats
Setting - Tensors

\[ V_\nu := \mathbb{R}^n, \quad \mathcal{H}_d = \mathcal{H} := \bigotimes_{\nu=1}^d V_\nu \quad d\text{-fold tensor product Hilbert-s.}, \]

\[ \mathcal{H} \simeq \{(x_1, \ldots, x_d) \mapsto U(x_1, \ldots, x_d) \in \mathbb{R} : x_i = 1, \ldots, n_i\}. \]

The function \( U \in \mathcal{H} \) will be called an order \( d \)-tensor. For notational simplicity, we often consider \( n_i = n \). Here \( x_1, \ldots, x_d \in \{1, \ldots, n\} \) will be called variables or indices.

\[ k \mapsto U(k) = U(k_1, \ldots, k_d) = (U_{k_1, \ldots, k_d}) \quad k_i = 1 \ldots, n_i. \]

Or in index (vectorial) notation

\[ U = \left(U_{k_1, \ldots, k_d}\right)_{k_i=0,1 \leq i \leq d}^{n_i} \]

\[ \dim \mathcal{H} = n^d \quad \text{curse of dimensions!!!} \]

E.g. wave function \( \psi(r_1, s_1, \ldots, r_N, s_N) \)
Vector-Tensorization - e.g. Binary coding

1D example: vector, e.g. signal

\[ k \rightarrow f(k) , \quad \left( \text{or } g\left(\frac{k}{2^d}\right) \right), \quad k = 0, \ldots, 2^d - 1. \]

Labeling of indices \( k \simeq \mu \in \mathcal{I} \) by an binary string of length \( d \),

\[ \mu = \mu(k) = (0, 0, 1, 1, 0, \ldots) \simeq \sum_{j=0}^{d-1} \mu_i 2^j = k(\mu), \quad \mu_i = 0, 1. \]

Tensorization

\[ \mu \mapsto U(\mu) := f(k(\mu)) \in \bigotimes_{j=0}^{d-1} \mathbb{R}^2, \quad \text{or} \quad \bigotimes_{j=0}^{d-1} \mathbb{C}^2. \]

This provides an isomorphism \( T : \mathbb{R}^{2^d} \leftrightarrow \bigotimes_{j=0}^{d-1} \mathbb{R}^2 \) by \( Tf := U \).

So far no information is lost, \( N = 2^d \) or \( d = \log_2 N \).
Hierarchical Tucker (HT) format
Hackbusch & Kühn (2009), Grasedyck (2010)


Noteable special case of HT: TT format , Oseledets & Tyryshnikov
TT- or matrix product representation of $U$

$$U(x) = U_1(x_1) \cdots U_i(x_i) \cdots U_d(x_d)$$

$$= \sum_{k_1=1}^{r_1} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} U_1(x_1, k_1) U_2(k_1, x_2, k_2) \cdots U_{d-1}(k_{d-2}x_{d-1}, k_{d-1}) U_d(k_{d-1}, x_d, k_d)$$

- component tensors $U_i(k_{j-1}, x_j, k_j) \in \mathbb{R}^{r_{i-1} \times n_i \times r_i}$, if $r := \max\{r_1, \ldots, r_{d-1}\}$, here $n = 2$, ( or 4, \ldots small!)

storage complexity DOFs: $\leq ndr^2$, 
Binary coding - signal compression - 1 D functions

Quantized TT - Oseledets (2009), Khoromskij (2009):

TT approximation of $U$

- Storage complexity $N$ is reduced to $2r^2 \log_2 N!$ (linear in $d = \log_2 N$)

- Allow extreme fine grid size $h = o(\epsilon) = 2^{-d} = \frac{1}{N}$.
  Example: $d = 50$, then $h \leq 10^{-15}$

Basic question: when is $r$ small or moderate?

(Grasedyck (2010), Hackbusch (2010), Oseledets (2010))

Examples:

1. For Kronecker $\delta_{i,j}$ (Dirac function) is $r = 1$.

2. For plane wave (fixed $k = \sum_{j=1}^{d} \nu_j 2^{j-1}$)

$$e^{2\pi ik} = e^{2\pi i \sum_{j=1}^{d} \nu_j 2^{j-1}} = \prod_{j=1}^{d} e^{2\pi i \nu_j 2^{j-1}}, \quad \nu_j = 0, 1,$$

again (complex) $r = 1$, or (real $r = 2$).
TT representation of tensorized $f$

The TT ranks $r_i$, $r = (r_1, \ldots, r_{d-1})$, of the TT decomposition of $U = T(f)$ are the ranks of the following matrices, $i = 1, \ldots, d-1$

$$r_i = \text{rank of } A_i = U_{\mu_1, \ldots, \mu_i}^{\mu_{i+1}, \ldots, \mu_d} = U_{(1,0,\ldots)}^{(1,0,\ldots)}.$$

or

$$A_i = U_p^q, \ p = p(\mu_1, \ldots, \mu_i) \text{ and } q = q(\mu_{i+1}, \ldots, \mu_d)$$

consider $g : [0, 1] \rightarrow \mathbb{R}, \ x = \frac{k}{2^d}, \ k = 0, \ldots, 2^d - 1, \ f(k) = g\left(\frac{k}{2^d}\right)$, then $k_1$ is the number of the subinterval

$$I_{k_1}^i := \left[\frac{k_1 - 1}{2^i}, \frac{k_1}{2^i}\right]$$

and $\left\{ \frac{k_1 - 1}{2^i} + \frac{1}{2^d} [0, \ldots, k_2, \ldots, 2^d - i - 1] : k_2 = 0, \ldots, 2d - i - 1 \right\}$

the corresponding grid.
Two-scale decomposition-similarity rank

Figure: Function $g(x) = x^2 \sin \frac{8x}{2\pi}$ and translated functions $g_k^3$
Similarity rank and Multi-scale decomposition

Figure: PCA principal component analysis (SVD) functions $g_k^3$, $k = 1, \ldots, 3$

Recursively performed SVD from level to level (Vidal-decomposition) $\Rightarrow$ multiscale decomposition, i.e. a TT representation of $U$. 
Examples for admitting low rank representations

Low ranks

1. (Piecewise) polynomials of degree $p + 1$.
2. $e^{i\langle k,x \rangle}$, i.e. Fourier polynomials
3. Splines, wavelets etc.
4. adaptive - best n-term approximation
   For homogenization:
5. Let $\phi : \mathbb{R} \to \mathbb{K}$ be $2^{-j}$-periodic, then $\phi|_{\Omega}$ is if rank $i\phi = 1$.
6. Let $f : [0, 1] \to \mathbb{K}$ has sim rank $f = p + 1$, and $\phi$ 1-periodic, then the modulated function $x \mapsto g(x) := f(x)\phi(\frac{x}{2^{-j}})$ has ranks rank $i\phi \leq p + 1$.

see e.g. appendix.
Low TT rank approximation

Theorem

1. \( f \in H^s(\Omega^j_k), k \in \Delta_j, (\Omega \subset \mathbb{R}^D) \). There exists \( f_\epsilon \) with rank \( r_j \) on level \( j \) satisfying

\[
\| f - f_\epsilon \|^2 \lesssim r_j^{-s/D} \sum_{k \in \Delta_j} |f|_{H^s}^2 .
\]  

(1)

2. \( f \) analytic on the domains \( \Omega^j_k, k \in \Delta_j \)

\[
\| f - f_\epsilon \| \lesssim e^{-\alpha r_j} , \text{ for some } \alpha > 0 .
\]  

(2)

3. Let \( f \) be a piecewise analytic function on \( \Omega \setminus \{x_0\} \) satisfying

\[
|\partial^\alpha f(x)| \leq c_\alpha |x - x_0|^{-\alpha} |\alpha|! 
\]

then there exists \( \alpha > 0 \) s.t.

\[
\| f - f_\epsilon \| \lesssim e^{-\alpha r_j} .
\]  

(3)
Examples for admitting low rank representations

- From this perspective, the approach can be easily extended to 2D and 3D Finite Elements with uniform refinement.
- **Black-box algorithm**: The multi-level scheme corresponds to a (multi-) wavelet packet decomposition.
- The **two-scale relations** are not fixed, but optimized in each component (such that no wavelets appears).
- Ranks $r_j \leq cN$ for best N-term approximation (e.g. Fourier, wavelets etc.), even with $c \ll 1$ but scaling is $\mathcal{O}(r^2)$!
Two-scale decomposition - matricisation

Reference domain $\Omega_0$ and $\Omega = \bigcup_{k \in \Delta_j} \Omega_k^j$, above example

$\Omega = \Omega_0$, $\Omega_k^j = [k2^{-j}, (k + 1)2^{-j}]$ together with isomorphisms (renormalization group)

$$T_k^j : \Omega_k^j \to \Omega_0 \ , \Omega = \bigcup_{k \in \Delta_j} \Omega_k^j .$$

$$f_l := \sum_{k \in \Delta_j} f_l|_{\Omega_k^j} : \Omega \to \mathbb{K}, \ l = 1, \ldots, m \text{ simultaneously},$$

$$f_{k,l}^j : \Omega_k^j \to \mathbb{K} , \ x \mapsto f_{k,l}^j(x) := f_l((T_k^j)^{-1}x) = f_l \circ (T_k^j)^{-1}(x) , \ x \in \Omega_0 .$$

The similarity rank of $f$ at level $j$

$$\text{sim rank}_jf := \dim \text{span}\{f_{k,l}^j : k \in \Delta_j \ l = 1, \ldots, n\} = r_j .$$

is equal to the TT Rank $r_j$ of $U$. 
Multi-scale decomposition by recursion

$$\Omega_0^j = \bigcup_{\mu_j=0}^{n_j-1} \Omega_{\mu_j}^{j+1} , \quad T_{\mu_j}^j : \Omega_{\mu_j}^{j+1} \to \Omega_0^{j+1} , \quad \mu_j = 0, \ldots, n_j$$

$k \in \Delta_{j+1}$ can be encoded by a multi-index

$$\mu = \mu(k) = (\mu_0, \ldots, \mu_j) \in I_0 \times \cdots \times I_j , \quad I_j = 0, \ldots, n_j - 1,$$

We define multi-scale scaling functions (see multi-wavelets)

$$x \mapsto \varphi_{\alpha_{j-1}}^j(x) , \quad x \in \Omega_0 .$$

Two-scale relation

$$\varphi_{\alpha_{j-1}}^j(x) = \sum_{\mu_j=1}^{n_j} \sum_{\alpha_j=1}^{r_d} U_j(\alpha_{j-1}, \mu_j, \alpha_j) \varphi_{\alpha_j}^{j+1}(T_{\mu_j}^j x) , \quad x \in \Omega_0 .$$
Multi-scale decomposition by recursion

Iterating

\[ T_{k}^{j+1} = T_{k(\mu)}^{j+1} = T_{\mu_0}^{0} \circ \cdots \circ T_{\mu_j}^{j} . \]

yields nonlinear (multi wavelet (packets)-) subdivision scheme

reconstructing a vector, resp. function \( f \) corresponds to a

tensor \( U \)

\[
x \mapsto f(x), \ x \in \Omega_0, \ (\mu) \mapsto U(\mu) = U(\mu_0, \ldots, \mu_d) ,
\]

\[
\Rightarrow f(x) = \sum_{\mu : k(\mu) \in \Delta_d} U(\mu) \varphi^d(T_{k(\mu)}^d x) , \ x \in \Omega_0 .
\]

\[
U(\mu, \alpha_d) = \sum_{\alpha_0=1}^{r_0} \cdots \sum_{\alpha_{d-1}}^{r_d-1} U_i(\alpha_{i-1}, \mu_i, \alpha_i) .
\]

\( U_i \) computed e.g. sequentially by SVDs (opt. - black box)
Binary coding - linear operators and matrices

\[ A = \begin{bmatrix} a(k_1, k_2) \end{bmatrix} : \otimes_{j=1}^d K^2 \to \otimes_{j=1}^d K^2 \cong K^{2^d} \]. coding

\[ \theta = (\theta_1, \ldots, \theta_{2d}) := ((\mu_1, \nu_1), \ldots, (\mu_d, \nu_d)) . \]

\[ ((\mu_1, \nu_1), \ldots, (\mu_d, \nu_d)) \mapsto A(\theta) := a(k(\theta)) \in \bigotimes_{j=1}^d K^{2 \times 2} , \]

Example:

1. Bit reversal is a permutation of tensor variables.
2. Identity matrix

\[ I = I_{2 \times 2} \otimes \cdots \otimes I_{2 \times 2} . \]

3. The Hadamard-Walsh transform is given by

\[ W = \bigotimes_{j=1}^d \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \]

4. finite differences, e.g.

\[ \frac{d^2 f}{dx^2} \approx \frac{f(x_{k-1}) - 2f(x_k) + f(x_{l+1})}{2h^2} , \quad h = 2^{-d} , \]

has ranks \( r_j \leq 3 \).

5. (quantum) Fourier transform
Examples: TT approximation of tensorized functions (QTT)

Airy function: \( f(x) = x^{1/4} \sin \frac{2x^{2/3}}{3} \), chirp: \( f(x) = \sin \frac{x}{4} \cos(x^2) \) and \( f(x) = \sin \frac{1}{x} \)
TT ranks of the tensorized Hartree potential $V_H$.

[Khoromskaia, Khoromskij, R.Schneider '11]

$$\bar{r} = \sqrt{\frac{1}{D} \sum_{\ell=1}^{D} r_{\ell-1} r_{\ell}}, \ D = 3L.$$ 

The average QTT ranks $\bar{r}$ of $V_H$ for molecules CH$_4$, H$_2$O and C$_2$H$_6$.

Comparison of times: QTT vs. canonical format (CH$_4$).

Grid-based calculations on $n \times n \times n$ ($n^3$) 3D Cart. grids, $n = 2^L$, $L = 5, \ldots, 12$, (in Matlab). For $L = 12$, $n^3 = 6.8 \cdot 10^{10}$. 
Ranks of the tensorized two-electron integrals (TEI) tensor.

[Khoromskaia, Khoromskij, R.Schneider ’12]

Given the “basis sampling“ tensor \( G = [G_{\mu\nu}] \in \mathbb{R}^{N_b \times N_b \times n^3} \) and the convolution \( H = [P_N *_{n^3} G_{\kappa\lambda}] \in \mathbb{R}^{N_b \times N_b \times n^3} \), the TEI tensor is computed by

\[
B = [b_{\mu\nu\kappa\lambda}] \approx G \times_{n^3} (P_N *_{n^3} G) = \langle G, P_N *_{n^3} G \rangle_{n^3} = \langle G, H \rangle_{n^3}.
\]

Pseudopotential of CH\(_4\): average QTT ranks of product basis functions, \( G_{\mu\nu} \), (left) and their Newton potential, \( H_{\kappa\lambda} \), (right), \( \varepsilon = 10^{-6} \), \( \mu, \nu, \kappa, \lambda = 1, ..., N_b \), \( N_b = 50 \), \( n = 8192 \).
Examples for linear operators - matrices

\[
((\mu_0, \nu_0), \ldots, (\mu_{d-1}, \nu_{d-1})) \mapsto A(\theta) \in \bigotimes_{j=1}^{d} \mathbb{K}^{2 \times 2},
\]

1. Identity matrix

\[ I = I_{2 \times 2} \otimes \cdots \otimes I_{2 \times 2}. \]

2. The Hadamard-Walsh transform is given by

\[
W = \bigotimes_{j=1}^{d} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}
\]

3. 1D- Laplace operator (2nd order) is of multi-ranks \( r_i \leq 3 \)

4. Convolutions (Hackbusch 2011)

5. The discrete Fourier transform is a 3\(d\)-fold product of rank \( \leq 2 \) operators - Fast Fourier transform (Quantum Fourier Transform)
Numerics on QTT-Fourier transform versus FFT in 1D

\[ t \mapsto f(t) = \Pi(t) := \chi_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(t) \text{ char. function} \]

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

**Table:** Time for QTT–FFT (in milliseconds) w.r.t. size \( n = 2^d \) and accuracy. \( \text{time}_{\text{QTT}} \) is the runtime of Alg. QTT–FFT, is the runtime of the FFT from the FFTW library, and \( \hat{f} = \text{sinc} \) is the effective QTT–rank of the Fourier image. In courtesy of B. Khoromskij

<table>
<thead>
<tr>
<th>( f = \Pi(t) )</th>
<th>= 10(^{-4})</th>
<th>= 10(^{-8})</th>
<th>= 10(^{-12})</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d )</td>
<td>( \hat{f} )</td>
<td>( \hat{f} )</td>
<td>( \hat{f} )</td>
</tr>
<tr>
<td>16</td>
<td>1.7</td>
<td>4.66</td>
<td>6.85</td>
</tr>
<tr>
<td>18</td>
<td>8.9</td>
<td>4.70</td>
<td>6.86</td>
</tr>
<tr>
<td>20</td>
<td>42.5</td>
<td>4.75</td>
<td>6.85</td>
</tr>
<tr>
<td>22</td>
<td>180</td>
<td>4.77</td>
<td>6.83</td>
</tr>
<tr>
<td>24</td>
<td>810</td>
<td>4.74</td>
<td>6.72</td>
</tr>
<tr>
<td>26</td>
<td>4100</td>
<td>4.62</td>
<td>6.76</td>
</tr>
<tr>
<td>28</td>
<td>26300</td>
<td>4.57</td>
<td>6.80</td>
</tr>
<tr>
<td>30</td>
<td>—</td>
<td>4.72</td>
<td>6.78</td>
</tr>
<tr>
<td>40</td>
<td>—</td>
<td>4.20</td>
<td>6.59</td>
</tr>
<tr>
<td>50</td>
<td>—</td>
<td>3.96</td>
<td>6.45</td>
</tr>
<tr>
<td>60</td>
<td>—</td>
<td>3.69</td>
<td>6.25</td>
</tr>
</tbody>
</table>
Comparison: quantum information theory - $\bigotimes_{i=1}^{d} \mathbb{C}^2$

$\mathcal{H} = \mathbb{C}^{2^d} \approx L_2(\{1, \ldots, 2^d\}, \mathbb{C}) \approx L_2(\{0, 1\}^d, \mathbb{C}) \approx \bigotimes_{i=1}^{d} \mathbb{C}^2$

all equipped with $L_2$ resp. Frobenius norms.

- here we have the same configuration space
- here we do not confine to $U \in \mathcal{H}$ with $\|U\| = 1$
- not necessarily a probabilistic interpretation
- we do not confine to unitary operator $A : \mathcal{H} \to \mathcal{H}$
- But: TT ranks must be moderate!

The treatment of quantum mechanical problems with MPS will be presented later!
Ordering of indices - space filling curves

There are various possibilities for higher-dimensional functions

1. lexicographical ordering \((k_1, k_2, \ldots) \rightarrow \otimes_{i=1}^D \otimes_{l=1}^{n_l} \mathbb{R}^2\)

2. \(Z\)-curve ordering, multi-level octree ordering \((k_1, k_2, \ldots) \rightarrow \otimes \mathbb{R}^{2D}\)

3. Hilbert space filling curves
Appendix: Translation invariant operators and spaces

\[ \mathbf{v} \mapsto L^{(p)} \mathbf{v} := \mathbf{v}^{(p)} + a_{p-1} \mathbf{v}^{(p-1)} + \cdots + a_0. \]

The kernel of the linear differential operator \( L^{(p)} \)

\[ V_p := \{ \mathbf{v} \in H^{p}_{\text{loc}}(\mathbb{R}) : L^{(p)} \mathbf{v} = 0 \}. \]

The space \( V_p \) is translation invariant, i.e.

\[ f(\cdot) \in V_p \Rightarrow f(\cdot - a) \in V_p, \forall a \in \mathbb{R}, \text{ dim } V_p = p + 1. \]

Let \( L^{(p)} \) be a linear partial differential operator of order \( p \),

\[ \mathbf{v} \mapsto L^{(p)} \mathbf{v} := \sum_{|\alpha| \leq p} a_\alpha D^\alpha \mathbf{v}. \]

Then \( V_p := \{ \mathbf{v} \in H^{p}_{\text{loc}}(\mathbb{R}) : L^{(p)} \mathbf{v} = 0 \} \), is translation invariant.

\[ S^p = \bigoplus_{k \in \Delta_j} V_p \big|_{\Omega^j_k} \Rightarrow \text{sim rank} S^p = \text{dim } V_p. \]
Examples

1. \( L^{(2)} v = v'' \) then \( V_p \) contains polynomials of degree \( p = 2 \).
2. \( L^{(1)} v = v' + a, a \in \mathbb{C} \) then \( v(x) = ce^{-ax} \).
3. \( L^{(2)} v = v'' + a^2, a \in \mathbb{R} \), then \( v(x) = c_1 \cos ax + c_2 \sin ax \) or \( v(x) = ce^{i\sqrt{ax}} \).

4. harmonic polynomials for the Laplacian \( L = \Delta \),
5. \( e^{i\langle k, x \rangle} \) or Fourier Bessel function for the Helmholtz operator, (Treftz method)
6. advection equation (Demkovich et al.)

Corollary

- \( S^p := \bigoplus S^p(\Omega_j^k) \),
  \[
  S^p(\Omega_j^k) = \{ f : \Omega_j^k : \mathbb{R} : f \text{ is a polynomial of degree } \leq p \}
  \]
  \[
  p_n = \sum_{k=0}^{p} a_k x^k \text{ are of similarity rank } \leq (p + 1)^D.
  \]
- Assume \( f = \sum_{|k| \leq p} \hat{f}_k e^{2\pi ikx} \), is piecewise bandlimited, and \( l \leq j \) sufficiently large. Then \( f \) has rank \( r_l \leq 2p + 1 \).
Appendix - examples

Let $T_k : \mathbb{R}^D \to \mathbb{R}^D$. We will call a function space $V_T$ of functions $T_k$-invariant if $\forall k \in \Delta_j$ and transformations $T_k$

$f \in V_T \Rightarrow f \circ T_k \in V_T$.

1. If $V_T$ is a $T_k$ invariant space of dimension $\dim V_T$ then the similarity rank of a function $f \in V_T$ is $\leq \dim V_T$.

2. $T_k$-translation invariant space $V_j$ generated by

$\varphi_k(x) = \varphi(x - k)$ has similarity rank

$p \leq \#\{\text{supp } \varphi_k \cap \Omega^j_0 \neq \emptyset\}$.

3. Multiresolution spaces $V^l = V^0 \oplus \bigoplus_{i=0}^{l-1} W^i$, where $l \leq j$ have similarity rank $\leq p$.

Homogenization

1. Let $\phi : \mathbb{R} \to \mathbb{K}$ be $2^{-J}$-periodic, $J \geq j$ then $\phi|_{\Omega}$ is if sim rank $\phi = 1$.

2. Let $f : [0, 1] \to \mathbb{K}$ has sim rank $f = p$, and $\phi$ 1-periodic, than

$x \mapsto g(x) := f(x)\phi(\frac{x}{2^{-j}})$ has similarity rank sim rank $g = p$. 
Two-scale decomposition

Theorem

1. Let us consider

\[ f = \sum_{j=0}^{d-1} \sum_{k \in \mathcal{I}^j} \sum_{i=1}^{p} a_{k,i}^j \phi_{k,i}^j, \]

to be a piecewise polynomial function with polynomial degree \( \leq p - 1 \). Let \( N_j \) be the number of domains \( \Omega_k^j \) where \( f|_{\Omega_k^j} \) is not a polynomial. Then the similarity rank of \( f \) at level \( j \) satisfy

\[ \text{sim rank}_j f \leq N_j + p. \]

2. (Best N-term representation) Let \( f \) be expanded by multiwavelets of degree \( p \) within \( N \) terms, then the similarity rank is bounded by

\[ \text{sim rank}_j f \leq \max\{N, p + 1\}. \]
Thank you for your attention.

References:
I. Oseledets & B. Khoromskij, Quantized TT format – QTT (Oseledets (09), Tyrtishnikov, Khoromskij)  
I. Oseledets, *Constructive Representation of Functions in Tensor Formats*,  
W. Hackbusch, *Tensorisation of Vectors and their Efficient Convolution*,  
W. Hackbusch, L. Grasedyck, R. Schneider *Multilevel tensorisation* (in preparation)  
Chemometrics and Intelligent Laboratory Systems, 110 (2012), 1-19.  
V. Khoromskaia, B. Khoromskij and R. Schneider QTT approximation of two-electron integrals Preprint