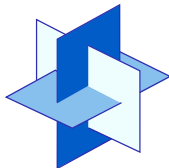


Alternating linear scheme ALS and MALS and the DMRG algorithm

R. Schneider (TUB Matheon)

John von Neumann Lecture – TU Munich, 2012



Setting - Tensors

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

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

The function $U \in \mathcal{H}$ will be called an **order d -tensor**.

For notational simplicity, we often consider $n_i = n$. Here $x_1, \dots, x_d \in \{1, \dots, n\}$ will be called **variables** or **indices**.

$$\mathbf{k} \mapsto U(k_1, \dots, k_d) = (U_{k_1, \dots, k_d}) , \quad k_i = 1 \dots, n_i .$$

Or in index (vectorial) notation

$$\mathbf{U} = (U_{k_1, \dots, k_d})_{k_i=0, 1 \leq i \leq d}^{n_i}$$

$\dim \mathcal{H} = n^d$ curse of dimensions!!!

E.g. wave function $\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)$

TT - Tensors - Matrix product representation

Noteable special case of HT:

TT format (Oseledets & Tyrtshnikov, 2009)

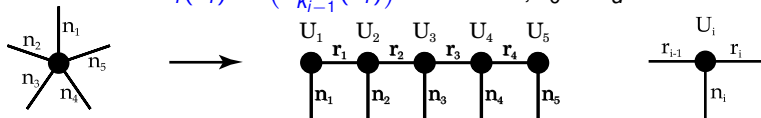
(matrix product states (MPS), Östlund & Rommer (1995), Vidal 2003, Schöllwock et al.)

Tensor U in the TT-representation written as matrix product

$$U(\mathbf{x}) = \mathbf{U}_1(x_1) \cdots \mathbf{U}_\nu(x_\nu) \cdots \mathbf{U}_d(x_d)$$

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

with matrices $\mathbf{U}_i(x_i) = (U_{k_{i-1}}^{k_i}(x_i)) \in \mathbb{R}^{r_{i-1} \times r_i}$, $r_0 = r_d := 1$



- ▶ component tensors $U_i(k_{i-1}, x_i, k_i) \in \mathbb{R}^{r_{i-1} \times n_i \times r_i}$
- ▶ Component functions $\mathbf{U}_i : x_i \mapsto \mathbf{U}_i(x_i) \in \mathbb{R}^{r_{i-1} \times r_i}$

Definitions: TT rank, separation rank

$$\mathbf{U}_i(x_i) : \{1, \dots, n_i\} \rightarrow \mathbb{R}^{r_{i-1} \times r_i},$$

$$\text{left matricisation } \mathbf{L}(\mathbf{U}_i(x_i)) := [\mathbf{U}_{k_{i-1}, x_i}^{k_i}]$$

left orthogonal

$$\mathbf{L}(\mathbf{U}_i(x_i))^T \mathbf{L}(\mathbf{U}_i(x_i)) = \sum_{k_{i-1}, x_i} \mathbf{U}_i(k_{i-1}, x_i, k'_i) \mathbf{U}_i(k_{i-1}, x_i, k_i) = \delta_{k'_i, k_i}$$

$\mathbf{U}_i(x_i) : \{1, \dots, n_i\} \rightarrow \mathbb{R}^{r_{i-1} \times r_i}$ satisfies **full rank condition** iff

$$\text{rank } \mathbf{L}(\mathbf{U}_i(x_i)) = r_i = \text{rank } \mathbf{R}(\mathbf{U}_{i+1}(x_{i+1})).$$

If all component functions $\mathbf{U}_i(x_i)$ of a TT decomposition

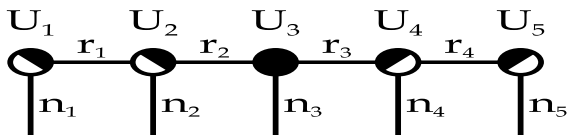
$$U(\mathbf{x}) = \mathbf{U}_1(x_1) \mathbf{U}_2(x_2) \cdots \mathbf{U}_{d-1}(x_{d-1}) \mathbf{U}_d(x_d)$$

satisfy the full rank condition, it has TT rank $\underline{r} = (r_1, \dots, r_{d-1})$.

This is equal to the *i-th separation rank s_i* of U :

$$\text{Rank of } \mathbf{A}_i = U_{x_1, \dots, x_i}^{x_{i+1}, \dots, x_d} = s_i = r_i.$$

Nonlinear Galerkin Ritz formulation



$$\text{TT-tensor } U = \sum_{\mathbf{k}} \bigotimes_{\nu=1}^d \mathbf{u}_{k_{\nu-1}}^{k_{\nu}}$$

Tensor optimization

Approximation of $W \in \mathcal{H}$: by a tensor of (canonical) rank \mathbf{r}

$$U = \operatorname{argmin} \{ F(V) = \frac{1}{2} \langle V - W, V - W \rangle : V \in \mathbb{T}_{\mathbf{r}} \} .$$

Linear elliptic equations

$$U = \operatorname{argmin} \{ F(V) = \frac{1}{2} \langle AV, V \rangle - \langle B, V \rangle : V \in \mathbb{T}_{\mathbf{r}} \} .$$

eigenvalue problems

$$U = \operatorname{argmin} \{ F(V) = \langle AV, V \rangle : \langle V, V \rangle = 1 : V \in \mathbb{T}_{\mathbf{r}} \} .$$

Existence

Setting includes elliptic PDE's in high dimensions.

Weak closedness (Hackbusch & Falco (2010))

$$\mathbb{T}_{\leq \mathbf{r}} = \bigcup_{s_i \leq r_i: 1 \leq i < d} \mathbb{T}_{\mathbf{s}} \subset \mathcal{H} = \text{clos } \mathbb{T}_{\mathbf{r}} \text{ is (weakly) closed!}$$

Multi-linear ansatz destroys **convexity**, but **existence of minimizers** follows from Hackbusch & Falco

Compare: Existence of Hartree Fock (Lieb & Simon, J.L. Lions) and of MCSCF (Frieesecke, Levin) But: even the best rank 1 approximation of U is NP hard to compute! (Lim (2011))

Optimization by relaxation in TT format ALS

Alternating Linear Scheme - ALS

Relaxation (see e.g. Gauss-Seidel, ALS):

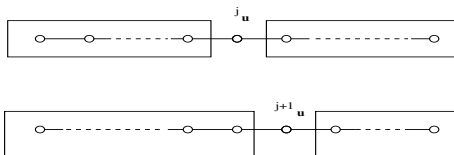
For $j = 1, \dots, d$:

1. fix all component tensors \mathbf{U}_ν , $\nu \in \{1, \dots, d\} \setminus \{j\}$, except index j .



2. Optimize $\mathbf{U}_j(k_{j-1}, x_j, k_j)$, and orthogonalize left

Repeat the relaxation procedure (in the opposite direction.)



Relaxation scheme $\mathbf{U}(x_\nu)_{k_{\nu-1}}^{k_\nu}$

This is NOT exactly the DMRG algorithm!

A modified relaxation algorithm MALS

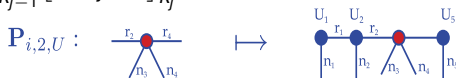
Disadvantage of simple relaxation:

1. \mathbf{U}_j, \dots must be (left)-orthogonalized,
2. $\mathbf{r} = (r_1, \dots, r_d)$ is fixed, but we want an adaptive rank selection

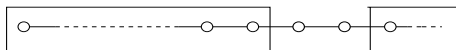
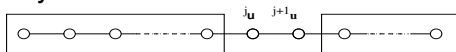
Modified Alternating Linear Scheme MALS (DMRG)

Idea: optimize **comprised components**

$$U_{i,i+1} =: [\mathbf{W}(k_{j-1}, x_j, x_{j+1}, k_{j+1})] \approx \sum_{k_j=1}^{r_j} [\mathbf{U}(x_j)]_{k_{j-1}}^{k_j} [\mathbf{V}(x_{j+1})]_{k_j}^{k_{j+1}} \in \mathbb{R}^{r_{i-1} n_i n_{i+1} r_{i+1}}.$$



system \otimes 2 sites \otimes environment



repeat in opposite direction \leftarrow

Retractions, extensions and density matrices

Let $x_i \mapsto \mathbf{W}_i(x_i) \in \mathbb{R}^{r_{i-1} \cdot n_i \cdot r_i}$ be of the size of a component tensor.

The **extension or dressing operator**

$$\mathbf{P}_{i,1,U} := \mathbf{E}_i := \mathbf{E}_i(U) : X \rightarrow \mathbb{T}_r \subset \mathcal{H},$$

extending \mathbf{W}_i to a TT tensor in $\mathcal{H} = \bigotimes_{i=1}^d \mathbb{R}^{n_i}$, is defined by

$$\mathbf{W}_i(x_i) \mapsto \mathbf{E}_i \mathbf{W}_i(\mathbf{x}) := \mathbf{U}_1(x_1) \cdots \mathbf{W}_i(x_i) \cdots \mathbf{U}_d(x_d).$$

The adjoint operator is \mathbf{E}_i^T . $\mathbf{P}_{i,2,U} := \mathbf{E}_{i,i+1}$ is defined analogously.

The left inverse \mathbf{E}_i^\dagger applied to the tensor $U = \mathbf{U}_1 \cdots \mathbf{U}_i \cdots \mathbf{U}_d \in \mathbb{T}_r$ is given by

$$\mathbf{U}_i = \mathbf{E}_i^\dagger U = (\mathbf{E}_i^T \mathbf{E}_i)^{-1} \mathbf{E}_i^T U.$$

E.g, if \mathbf{U}_i , $i = 1, \dots, d-1$, are left orthogonal, the i -th **density matrix** will be defined by

$$\mathbf{I} \otimes \mathbf{D}_i := (\mathbf{E}_i^T \mathbf{E}_i)^{-1}.$$

Relaxation TT- ALS algorithm

Require: Functional $F : \mathbb{R}^{n_1 \times \dots \times n_d} \rightarrow \mathbb{R}$, component starting vector (U_1, \dots, U_d) .

Right-orthogonalise the components (U_2, \dots, U_d) .

while termination criterion is not fulfilled **do**

for $i = 1, \dots, d - 1$ **do**

 Find $\hat{U}_i = \operatorname{argmin} F \circ \mathbf{E}_i : \mathbb{R}^{r_{i-1} \times n_i \times r_i} \rightarrow \mathbb{R}$.

 Set $U_i, U_{i+1} \leftarrow \hat{U}_i, U_{i+1}$, by applying an orthogonalisation.

end for

for $i = d, \dots, 2$ **do**

 Find $\hat{U}_i = \operatorname{argmin} F \circ \mathbf{E}_i : \mathbb{R}^{r_{i-1} \times n_i \times r_i} \rightarrow \mathbb{R}$.

 Set $U_i, U_{i-1} \leftarrow \hat{U}_i, U_{i-1}$, by applying an orthogonalisation.

end for

end while

Modified relaxation TT- MALS algorithm (DMRG).

Require: Functional : $\mathbb{R}^{n_1 \times \dots \times n_d} \rightarrow \mathbb{R}$, component starting vector (U_1, \dots, U_d) .

Right-orthogonalise the components (U_2, \dots, U_d) .

while termination criterion is not fulfilled **do**

for $i = 1, \dots, d - 1$ **do**

 (Optimization step) Find

$(k_{i-1}, x_i, x_{i+1}, k_{i+1}) \mapsto W_i(k_{i-1}, x_i, x_{i+1}, k_{i+1})$ where

$$W_i = \operatorname{argmin} F \circ \mathbf{W}_{i,i+1} : \mathbb{R}^{r_{i-1} \times n_i n_{i+1} \times r_{i+1}} \rightarrow \mathbb{R}.$$

 (Decimation step) Approximate with low rank r_i

$$\mathbf{w}_{k_{i-1}, x_i}^{x_{i+1}, k_{i+1}} \approx \sum_{k_j} \mathbf{u}_{k_{i-1}, x_i}^{k_j} \mathbf{v}_{k_j}^{x_{i+1}, k_{i+1}}.$$

end for

 Repeat with reverse order

end while

ALS - Stationary conditions (equations)

For fixed j , and $Ax = B$ the **stationary condition** is

$$\tilde{\mathbf{A}}_j \mathbf{U}_j = \mathbf{E}_j^T \mathbf{A} \mathbf{E}_j \mathbf{U}_j = \mathbf{b}_j = \mathbf{E}_j \mathbf{B} .$$

For the **eigenvalue problem**

$$\tilde{\mathbf{A}}_j \mathbf{U}_j = \mathbf{E}_j^T \mathbf{A} \mathbf{E}_j \mathbf{U}_j = \lambda \tilde{\mathbf{M}}_j \mathbf{U}_j = \lambda \mathbf{E}_j^T \mathbf{E}_j \mathbf{U}_j .$$

For MALS we obtain

$$\tilde{\mathbf{A}}_j = \mathbf{E}_{j,j+1}^T \mathbf{A} \mathbf{E}_{j,j+1} \quad , \quad \tilde{\mathbf{M}}_j = \mathbf{E}_{j,j+1}^T \mathbf{E}_{j,j+1}$$

Lemma (Rohw. & Holtz & S.)

If $\mathbf{U}_j, j < i$ are *left-*, and $\mathbf{U}_j, i < j$ are *right-orthogonal* then

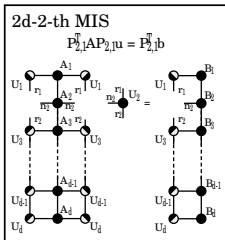
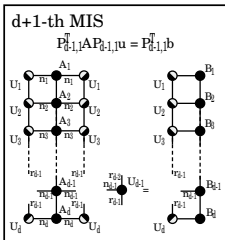
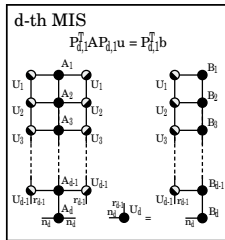
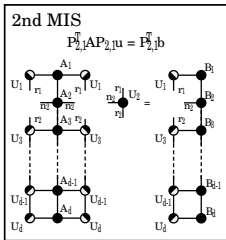
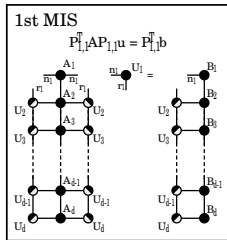
$$\text{cond } \tilde{\mathbf{A}}_i \leq \text{cond } A \quad \text{and} \quad \tilde{\mathbf{M}}_i = \mathbf{I} !$$

Remark: $\det \mathbf{A}_i \neq 0$ can hold, for some A when $\det A = 0$
(-RIP, compressed sensing)

ALS - Solution scheme for linear equations $Ax = b$

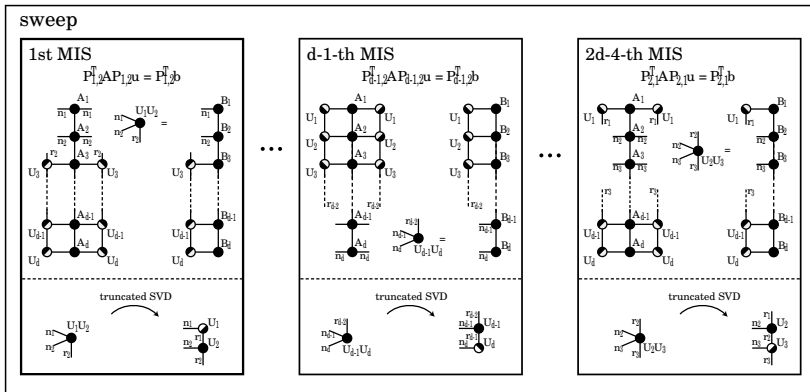
- ▶ Proceeding in “sweeps”: Solution of successive linear systems
- ▶ additional orthogonalisation

sweep



MALS - Solution scheme for linear equations $Ax = b$ (inspired by the DMRG algorithm)

- ▶ Proceeding in “sweeps”: Solution of successive linear systems for **comprised components** $U_{i,i+1} \in \mathbb{R}^{r_{i-1}n_i n_{i+1}r_{i+1}}$, other components fixed
- ▶ subsequent **SVDs** restore orthogonal TT-format



ALS - Stationary conditions

$$\mathbf{G}_j := [\mathbf{g}_{k_{j-1}}^{k'_j}] = \langle \mathbf{L}^{k_{j-1}}, \mathbf{A}_L \mathbf{L}^{k'_j} \rangle ,$$

$$\mathbf{H}_j := [\mathbf{h}_{k_j}^{k'_j}] = \langle \mathbf{R}_{k_j}, \mathbf{A}_R \mathbf{R}_{k'_j} \rangle ,$$

$$\mathbf{X}_j = [\mathbf{p}_{k_{j-1}}^{k'_j}] = \langle \mathbf{L}^{k_{j-1}}, \mathbf{L}^{k'_j} \rangle , \quad \mathbf{Q}_j := [\mathbf{q}_{k_j}^{k'_j}] = \langle \mathbf{R}_{k_j}, \mathbf{R}_{k'_j} \rangle .$$

$$\langle \mathbf{e}_i, \mathbf{b}_{k_{j-1}}^{k_j} \rangle := \langle B, \mathbf{L}^{k_{j-1}} \otimes \mathbf{e}_i \otimes \mathbf{R}_{k_j} \rangle ,$$

and let $\mathbf{u}_j := [u_{k_{j-1}, x_j, k_j}]$, $\mathbf{b} := [b_{k_{j-1}, x_j, k_j}]$

For fixed j , and $\mathbf{A}\mathbf{x} = \mathbf{B}$ the **stationary condition** is

$$\tilde{\mathbf{A}}_j \mathbf{u}_j = (\mathbf{G}_j \otimes \mathbf{a}_j \otimes \mathbf{H}_j) \mathbf{u}_j = \mathbf{b}_j . \quad *$$

and for the **eigenvalue problem**

$$\tilde{\mathbf{A}}_j \mathbf{u}_j := (\mathbf{G}_j \otimes \mathbf{a}_j \otimes \mathbf{H}_j) \mathbf{u}_j = \lambda \tilde{\mathbf{M}}_j \mathbf{u}_j = \lambda (\mathbf{X}_j \otimes \mathbf{I} \otimes \mathbf{Q}_j) \mathbf{u}_j . \quad *$$

* Rank $\mathbf{A} = 1$ i.e. $\mathbf{A} = \bigotimes_{\nu=1}^d \mathbf{a}_\nu$.

Stationary condition

Defining the matrices

$$\mathbf{G} := ((\mathbf{g})_{k_{j-1}}^{k'_j}) = \langle \ell^{k_{j-1}}, \mathbf{A}_L \ell^{k'_j} \rangle,$$

$$\mathbf{H} := ((\mathbf{h})_{k_j}^{k'_j}) = \langle \mathbf{r}_{k_j}, \mathbf{A}_{RR} \mathbf{r}_{k'_j} \rangle,$$

and let $\mathbf{u} := (u_{k_{j-1}, x_j, k_j})$, $\mathbf{b} := (b_{k_{j-1}, x_j, k_j})$

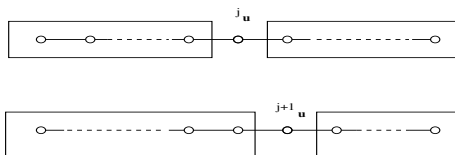
For fixed j , the stationary condition is

$$\tilde{\mathbf{A}} \mathbf{u} = (\mathbf{G} \otimes \mathbf{a} \otimes \mathbf{H}) \mathbf{u} = \mathbf{b}.$$

*

* Rank $\mathbf{A} = 1$

Update $j \rightarrow j + 1$



$j \rightarrow j + 1$: update ${}^j \mathbf{G} \rightarrow {}^{j+1} \mathbf{G} =: \mathbf{G}$,

$$g_{k_j}^{k'_j} = \sum_{x_j, y_j=1}^n \sum_{k_{j-1}, k'_{j-1}=0}^{r_{j-1}} u_{x_j, k_{j-1}}^{k_j} g_{k_{j-1}}^{k'_{j-1}} a_{x_j}^{y_j} u_{y_j, k'_{j-1}}^{k'_j} . \quad *$$

shortly

$${}^{j+1} \mathbf{G} := \mathbf{u}^T ({}^j \mathbf{G} \otimes \mathbf{a}) \mathbf{u} , \quad (u_{x_j, k_{j-1}}^{k_j}) := (u_{x_j, k_{j-1}}^{k'_j}) \quad *$$

For $\nu = 1, \dots, d$, ${}^\nu \mathbf{H}$ are computed once and stored.

* Rank $\mathbf{A} = 1$

Theorem (reconstruction)

Let the initial guess $U^{(0)} \in \mathbb{T}^r$ does not provide rank deficiencies, e.g. be sufficiently close to

$$U = \operatorname{argmin} \{F(V) := \frac{1}{2} \langle V - b, V - b \rangle : V \in \mathbb{T}^r\} .$$

And if $U = b$, with $U \in \mathbb{T}^r$ (reconstruction), then the ALS scheme $U^{(n)} \in \mathbb{T}^r$ terminates after the first half sweep.

Theorem (approximation)

Let the initial guess $U^{(0)} \in \mathbb{T}^r$ be sufficiently close to

$$U = \operatorname{argmin} \{F(V) := \frac{1}{2} \langle AV - 2f, V \rangle : V \in \mathbb{T}^r\} .$$

Let $A : \mathcal{H} \rightarrow \mathcal{H}$ be SPD, and $W \in \mathcal{H}$ solving $AW = f$, and $\|W - U\| \leq \delta$ sufficiently small. Then the ALS converges linearly

$$\|U - U^{(n+1)}\| \leq L \|U - U^{(n)}\| , \quad 0 < L < 1 .$$

Sketch of proof

Suppose $U \in N_\delta(0)$ and let $N_\delta(0)$ be a suitable neighbourhood of $U^{(n)}$. Let $\mathbf{u} := (\mathbf{w}_1, \dots, \mathbf{w}_d) \in X := X_r$ and assume a (local) parametrization $U = T(\mathbf{u})$ (details later).

We assume $A : \mathcal{H} \rightarrow \mathcal{H}$ is SPD and consider

$$J(\mathbf{u}) := F(T(\mathbf{u})) = \frac{1}{2} \langle AT(\mathbf{u}), T(\mathbf{u}) \rangle - \langle f, T(\mathbf{u}) \rangle .$$

First order condition reads as

$$[J'(\mathbf{u})](\mathbf{v}) = \langle AT(\mathbf{u}) - f, [T'(\mathbf{u})](\mathbf{v}) \rangle = 0 \quad \forall \mathbf{v} \in X,$$

For ALS, i.e. non-linear Gauß-Seidel, we have to show that the corresp. Hessian is SPD:

$$\begin{aligned} [J''(\mathbf{u})](\mathbf{v}, \mathbf{v}) &= \langle A[T'(\mathbf{u})](\mathbf{v}), [T'(\mathbf{u})](\mathbf{v}) \rangle \\ &\quad + \langle AT(\mathbf{u}) - f, [T''(\mathbf{u})](\mathbf{v}, \mathbf{v}) \rangle \geq \gamma \|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in X. \end{aligned}$$

If the residual $F'(T(\mathbf{u})) = AT(\mathbf{u}) - f = AU - f \in \mathcal{H}$ is sufficiently small, e.g. $= 0$ (reconstruction), this remains to be true, due to the SPD operator A .

Complexity - rough estimates

MALS \approx DMRG:

- ▶ storage components $\mathcal{O}(n^2 dr^2)$
- ▶ storage matrices $\mathcal{O}(n^2 dR_c r^2)$
- ▶ computation $\mathcal{O}(n^2 dR_c r^3)$

if R_c is a canonical rank of A .

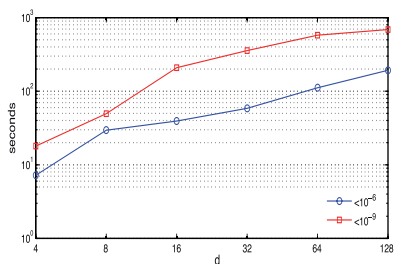
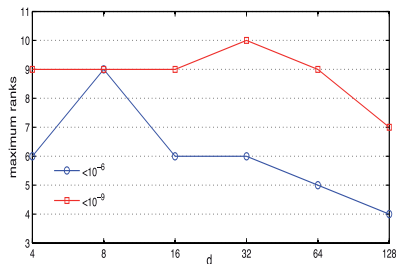
If A is represented in TT with maximal TT rank R . then

- ▶ storage components $\mathcal{O}(n^2 dr^2)$
- ▶ storage matrices $\mathcal{O}(n^2 dR^2 r^2)$
- ▶ computation $\mathcal{O}(n^2 dR^3 r^3)$

For ALS (one-site DMRG) n^2 is replaced by n .

Solution of $-\Delta U = b$

- ▶ Dimension $d = 4, \dots, 128$ varying
- ▶ Gridsize $n = 10$
- ▶ Right-hand-side b of rank 1
- ▶ Solution U has rank 13



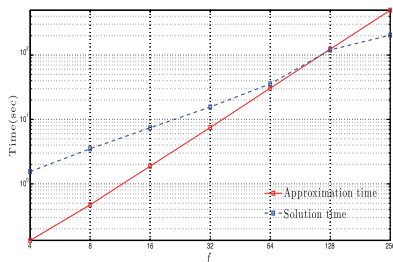
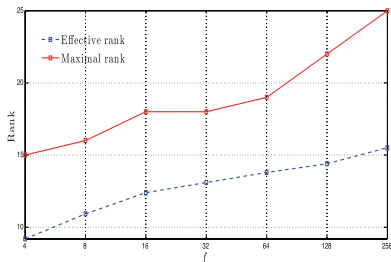
Example: Eigenvalue problem in QTT

in courtesy of B. Khoromskij, I. Oseledets, *QTT: Toward bridging high-dimensional quantum molecular dynamics and DMRG methods*,

$$H\Psi = \left(-\frac{1}{2}\Delta + V\right)\Psi = E\Psi$$

with potential energy surface given by Henon-Heiles potential

$$V(q_1, \dots, q_f) = \frac{1}{2} \sum_{k=1}^f q_k^2 + \lambda \sum_{k=1}^{f-1} \left(q_k^2 q_{k+1} - \frac{1}{3} q_k^3 \right).$$



Dimensions $f = 4, \dots, 256$; 1-D grid size $n = 128 = 2^7 = 2^d$; \rightsquigarrow

$$\text{QTT-tensors} \in \bigotimes_{i=1}^{7f} \mathbb{R}^2 = \mathbb{R}^{\underbrace{2 \times \dots \times 2}_{7f=1792}}.$$

Thank you for your attention.

References:

- G.K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **118**, 8551 (2003)
- M. Espig, W. Hackbusch, S. Handschuh and R. Schneider *Optimization Problems in Contracted Tensor Networks*, MPI, MIS Preprint 66/2011, submitted to *Numer. Math.*
- S. Holtz, Th. Rohwedder, R. Schneider, *On manifolds of tensors with fixed TT rank*, *Num. Math.* DOI: 10.1007/s00211-011-0419-7, 2011.
- S. Holtz, Th. Rohwedder, R. Schneider, *The alternating linear scheme for tensor optimization in the TT format*, to appear in *SISC*, 2011.
- T.Huckle, T. Schulte-Herbrcken and K. Waldherr, *Computations in Physical Tensor Networks*, to appear in *Linear Algebra and its Applications*, Special Issue on Tensors and Multilinear Algebra.
- B. Khoromskij, I. Oseledets, *DMRG+QTT approach to computation of the ground state of the molecular Schrödinger operator*, submitted to *Num. Math.*, 2011.
- D. Kressner and C. Tobler, *htucker - A Matlab toolbox for tensors in hierarchical Tucker format* SAM, ETHZ Preprint, 2012.
- J. M. Landsberg, Yang Qi, Ke Ye , *On the geometry of tensor network states* arXiv:1105.4449v2 [math.AG] to appear *Quantum Information Computation*
- O. Legeza, T. Rohwedder and R. Schneider, *Numerical approaches for high-dimensional PDE's for quantum chemistry* , in *ENCYCLOPEDIA OF APPLIED AND COMPUTATIONAL MATHEMATICS*, B. Engquist ed., Springer, to appear
- Martin J. Mohlenkamp, *Musings on Multilinear Fitting* , to appear in *Linear Algebra and its Applications*.
- Th. Rohwedder, A. Uschmajew, *Convergence of the Alternating Linear Scheme for optimization of tensors in the TT format*, submitted to *SINUM*, 2011.
- U. Schollwoeck, *The density-matrix renormalization group*, *Rev. Mod. Phys.* **77**, 259 (2005)
- S.R. White, *Phys. Rev. Lett* **69**, 2863 (1992)