# Tensor
# Basic Facts II

# Overview

Hierarchical Tucker

Tensor Train

Generalizing matrix properties

Software

# Tensor Networks

Original tensor $A_{i1,\ldots,id}$: memory $n^d$

CP: dnR, good compression but bad approximation

Tucker: $dnr+r^d$ , good approximation but bad compression

Find better Tensor Networks as a compromise of both:

Recursive binary tree (→ H-Tucker) or

Recursive chain (→ Tensor train)

# Hierarchical Decompositions

$$CP: \quad A = \sum_{j=1}^{r} \bigotimes_{\mu=1}^{d} a_{j,\mu} \quad \in IR^{I_1 \times \cdots \times I_d}$$

←→ Tensor rank

$$\text{H-Tucker}: \quad A = \sum_{j_1}^{r_1} \cdots \sum_{j_q}^{r_q} \prod_{\nu=1}^{p} B_{J_\nu} \bigotimes_{\mu=1}^{d} a_{j_\mu,\mu}, \quad J_\nu \subset \{j_1, \cdots, j_q\}$$

←→ hierarchical rank

$$\text{Tucker}: \quad A = \sum_{j_1}^{r_1} \cdots \sum_{j_d}^{r_d} C_{j_1,\cdots,j_d} \bigotimes_{\mu=1}^{d} a_{j_\mu,\mu}, \quad C \in IR^{r_1 \times \cdots \times r_d}$$

←→ multilinear rank

# Tucker revisited

Special case of Tucker:
Orthogonal Tucker, if all matrices $U^{(\mu)}$ the so called mode frames are orthogonal = HOSVD

$$A = G \times_1 U^{(1)} \times_2 U^{(2)} \ldots \times_N U^{(N)} = [[G; U^{(1)}, U^{(2)}, \ldots, U^{(N)}]]$$

$$A_{i_1 i_2 \ldots i_N} = \sum_{k_1=1}^{R_1} \sum_{k_2=1}^{R_2} \ldots \sum_{k_N=1}^{R_N} G_{k_1 k_2 \ldots k_N} u^{(1)}_{i_1 k_1} u^{(2)}_{i_2 k_2} \ldots u^{(N)}_{i_N k_N}, \quad i_n = 1, \ldots, I_n$$

# Best Appoximation

For given orthogonal $U^{(\mu)}$ , $\mu$=1,…,d, it holds

$$\min_G \left\| A - G \times_1 U^{(1)} \times_2 U^{(2)} \ldots \times_N U^{(N)} \right\| \Leftrightarrow$$

$$G = A \times_1 U^{(1)^T} \times_2 U^{(2)^T} \ldots \times_N U^{(N)^T}$$

Truncation of tensor A to Tucker rank ($k_1$,…,$k_N$) via SVD:

$$A_{(\mu)} = U_\mu \Sigma_\mu V_\mu^T, \qquad \Sigma_\mu = diag(\sigma_{\mu,1},\ldots,\sigma_{\mu,n_\mu}), \qquad \sigma_{\mu,1} \geq \ldots \geq \sigma_{\mu,n_\mu}$$

Define $\tilde{U}_\mu := U_\mu(:,1:k_\mu)$, the Tucker truncation of A by

$$T_{(k_1,\ldots,k_N)}(A) := \left( A \times_1 \tilde{U}_1^T \times_2 \ldots \times_N \tilde{U}_N^T \right) \times_1 \tilde{U}_1 \times_2 \ldots \times_N \tilde{U}_N =$$

$$= A \times_1 \left( \tilde{U}_1 \tilde{U}_1^T \right) \times_2 \ldots \times_N \left( \tilde{U}_N \tilde{U}_N^T \right)$$

# Tucker Truncation Error

It holds: $$\left\| A - T_{(k_1,\ldots,k_N)}(A) \right\| \le \sqrt{\sum_{\mu=1}^{N} \sum_{i=k_\mu+1}^{n_\mu} \sigma_{\mu,i}^2} \le \sqrt{N} \left\| A - A^{best} \right\|$$

where $A^{best}$ is the best possible approximation in Tucker($k_1,\ldots,k_N$).

Beware!
Change of the notation: Unfolding $A_{(n)} \rightarrow A^{(\mu)}$, $N \rightarrow d$ !!!

# Hierarchical Rank Model

$$A \in IR^{I}, \qquad I = I_1 \times \cdots \times I_d$$

$$\Downarrow$$

$$A \in IR^{(I_1 \times \cdots \times I_q) \times (I_{q+1} \times \cdots \times I_d)},$$

Matricification by reduction to two indices → matrix A

Compute SVD of A:

$$A = \sum_{i=1}^{d} U_i \Sigma_{ii} V_i^T, \qquad U_i \in IR^{I_1 \times \cdots \times I_q}, \quad V_i \in IR^{I_{q+1} \times \cdots \times I_d},$$

$U_i$, $V_i$ seen as vectors/tensors of lower dimensional.

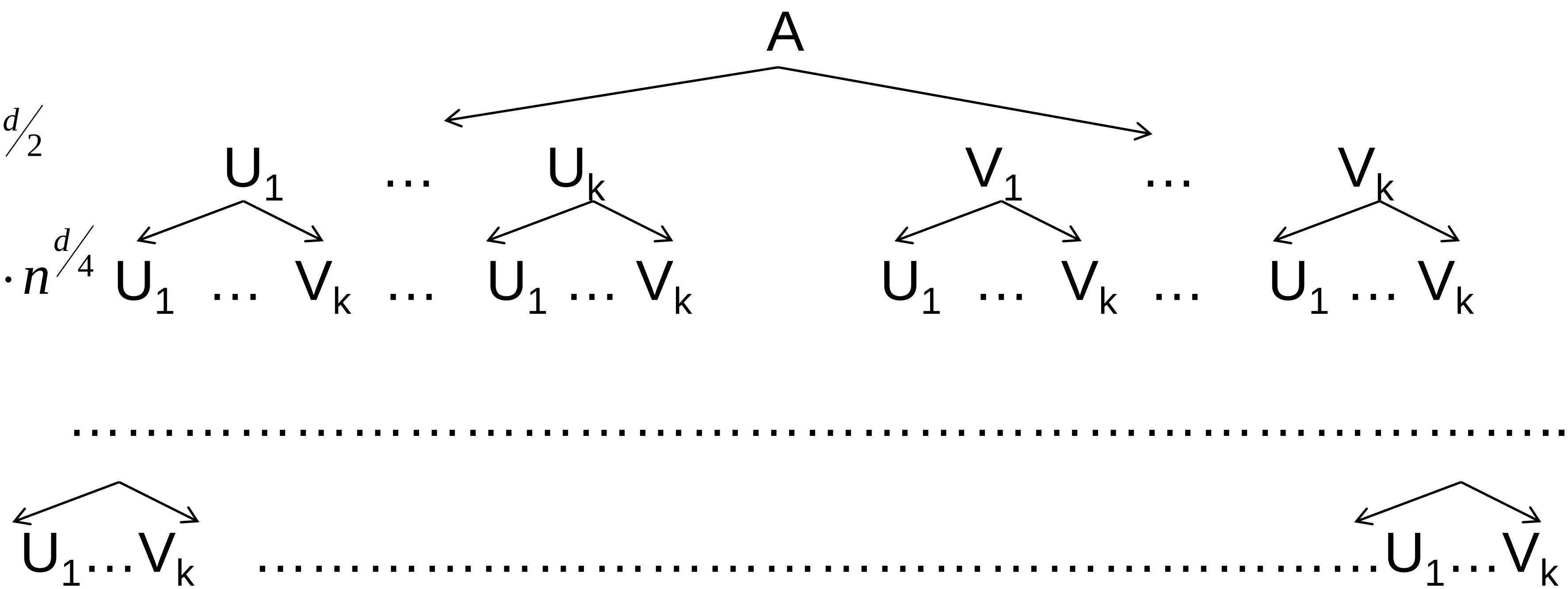
Repeat for tensor matricisation $U_l$:

$$U_l = \sum_{i=1}^{r} W_i^{(l)} \hat{\Sigma}_{ii} \left( X_i^{(l)} \right)^T, \qquad W_j^{(l)} = \sum_{i=1}^{r} Y_i^{(l,j)} \tilde{\Sigma}_{ii} \left( Z_i^{(l,j)} \right)^T, \quad etc.$$

# Memory Costs

$$2k \cdot n^{d/2}$$

$$(2k)^2 \cdot n^{d/4}$$




Data complexity of the last row: $O\left((2k)^{\log(d)} \cdot n^{d/d}\right) = O\left(dnk^{\log(d)}\right)$

# Hierarchical uniform subspaces

$$A \in IR^{(I_1 \times \cdots \times I_q) \times (I_{q+1} \times \cdots \times I_d)},$$

$$A = \sum_{i=1}^{d} U_i \Sigma_{ii} V_i^T, \qquad U_i \in IR^{I_1 \times \cdots \times I_q}, \quad V_i \in IR^{I_{q+1} \times \cdots \times I_d},$$

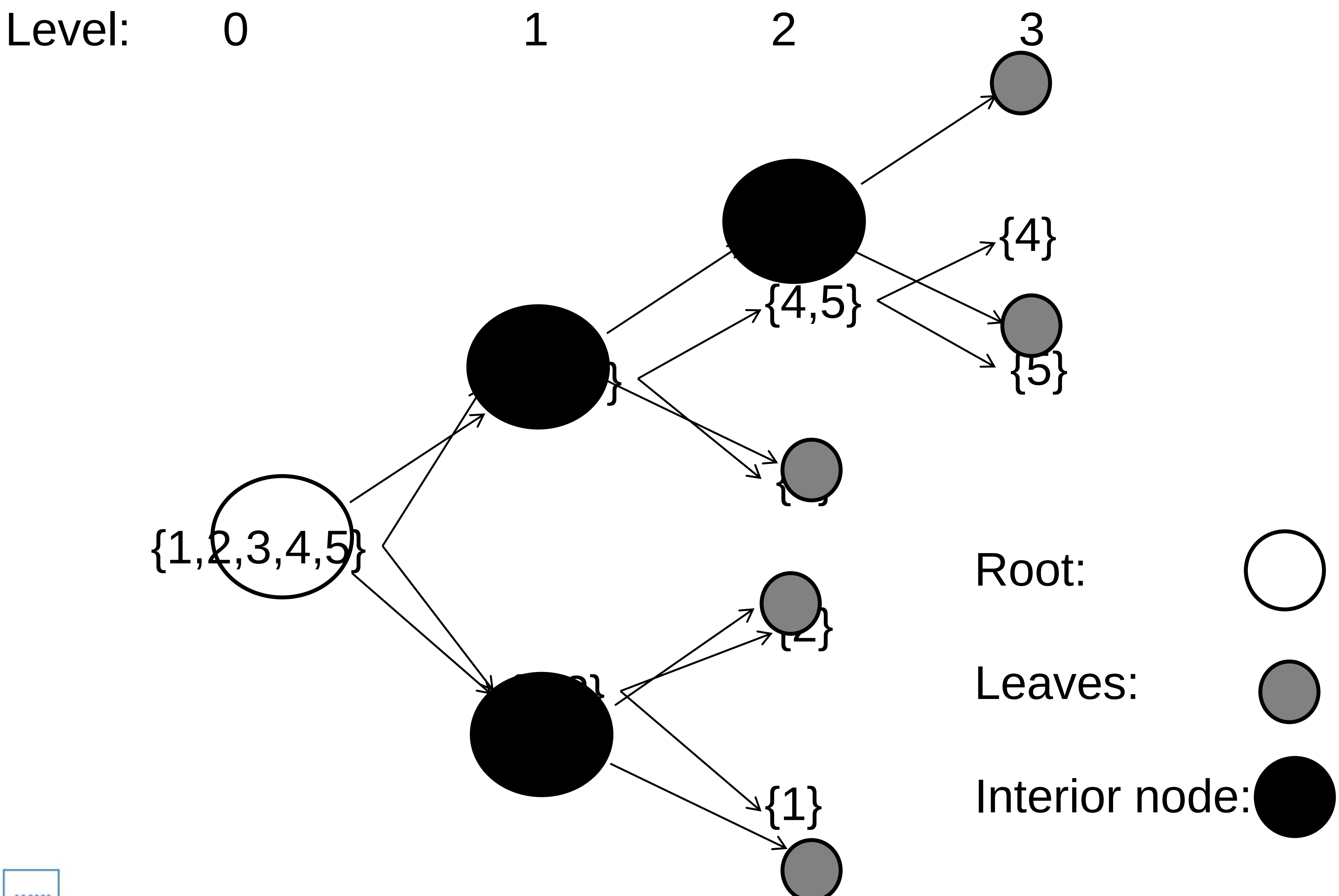
$U_i$, $V_i$ seen as vectors/tensors are lower dimensional.

Better and cheaper recursion in binary tree:
Repeat for tensor $U_l$:

$$U_l = \sum_{i=1}^{r} \sum_{j=1}^{r} W_i \, B_{l,i,j} X_j^T, \qquad W_l = \sum_{i=1}^{r} \sum_{j=1}^{r} Y_i \, \tilde{B}_{l.i.j} Z_j^T, \quad etc.$$

Leads to smaller data complexity $O(dnk + dk^3)$

# Dimension Tree



# Dimension Tree

Definition: A dimension tree $T_I$ for dimension d ϵ IN is a tree with root {1,…,d} and depth $p = \lceil \log_2(d) \rceil$ such that each node t ϵ $T_d$ is either

1. a leaf and singleton t={μ } on level l ϵ {p-1,p}
2. the union of two disjoint successors S(t) = {$s_1$,$s_2$}: $t = s_1 \dot{\cup} s_2$

The level l of the tree is defined as the set of all nodes having a distance of exactly l to the root.

$$T_I^l := \{ t \in T_I \mid level(t) = l \}.$$

Set of leaves: $L(T_I)$. Set of interior nodes: $I(T_I)$.
A node of the tree is a mode cluster = union of modes.

# Properties

Up to last level: complete binary tree.

For a canonical dimension tree each interior node has two successors:

$$t = \{\mu_1, \dots, \mu_q\}, \qquad q > 1$$

$$t_1 := \{\mu_1, \dots \mu_r\}, \qquad r := \lfloor q/2 \rfloor, \qquad t_2 := \{\mu_{r+1}, \dots, \mu_q\}$$

Total number of nodes: 2d-1
Number of leaves: d
Number of interior nodes: d-1

# Hierarchical Rank

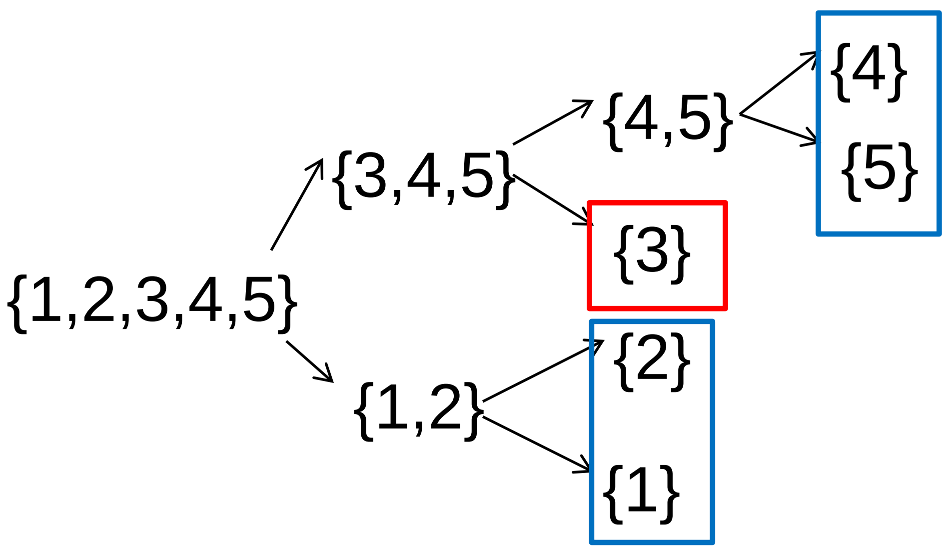
For a mode cluster t in a dimension tree $T_I$ we define the complementary cluster $t' := \{1,\dots,d\} \setminus t$, with

$$I_t := \times_{\mu \in t} I_\mu, \qquad I_{t'} := \times_{\mu \in t'} I_\mu,$$

and the related t-matricization

$$M_t : IR^I \to IR^{I_t \times I_{t'}}, \qquad \left(M_t(A)\right)_{(i_\mu)_{\mu \in t}, (i_{\mu'})_{\mu' \in t'}} := A_{(i_1,\dots,i_d)}$$

$$Notation: \qquad M_t(A) = A^{(t)}$$




$$A^{(\{3\})} = A_{\{3\},\{1,2,4,5\}}$$

# Example

$$A = a \otimes b \otimes c \otimes d \in IR^{I_1 \times I_2 \times I_3 \times I_4}$$

$$A^{(\{1,2\})} = (a \otimes b)(c \otimes d)^T \in IR^{(I_1 \times I_2) \times (I_3 \times I_4)}$$

$$A^{(\{3,4\})} = (c \otimes d)(a \otimes b)^T \in IR^{(I_3 \times I_4) \times (I_1 \times I_2)}$$

$$A^{(\{2,3\})} = (b \otimes c)(a \otimes d)^T \in IR^{(I_2 \times I_3) \times (I_1 \times I_4)}$$

$$A^{(\{1\})} = a(b \otimes c \otimes d)^T \in IR^{I_1 \times (I_2 \times I_3 \times I_4)}$$

For tensor A , dimension tree $\mathrm{T}$, node t with complementary cluster t‘ it holds

$$A^{(t')} = A^{(t)^T}$$

# Hierarchical Rank

For dimension tree $T_I$ the hierarchical rank $(k_t)_{t \in T_I}$ of a tensor $A \in IR^I$ is defined by

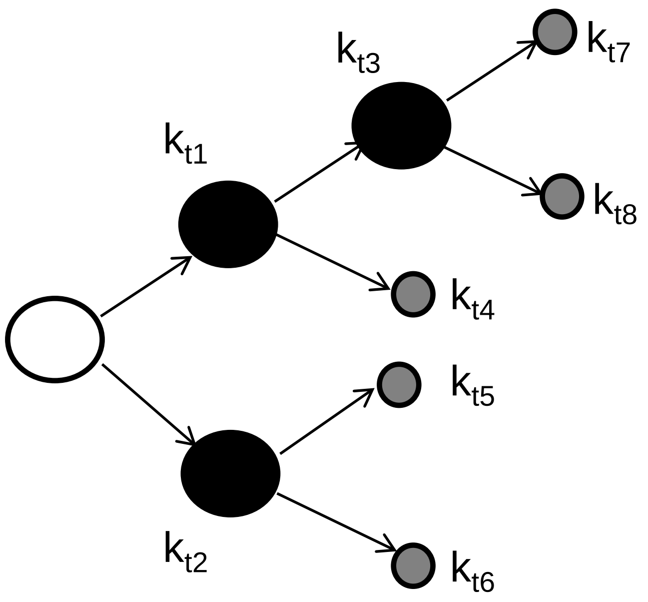
$$\forall t \in T_I : \qquad k_t := rank(A^{(t)})$$

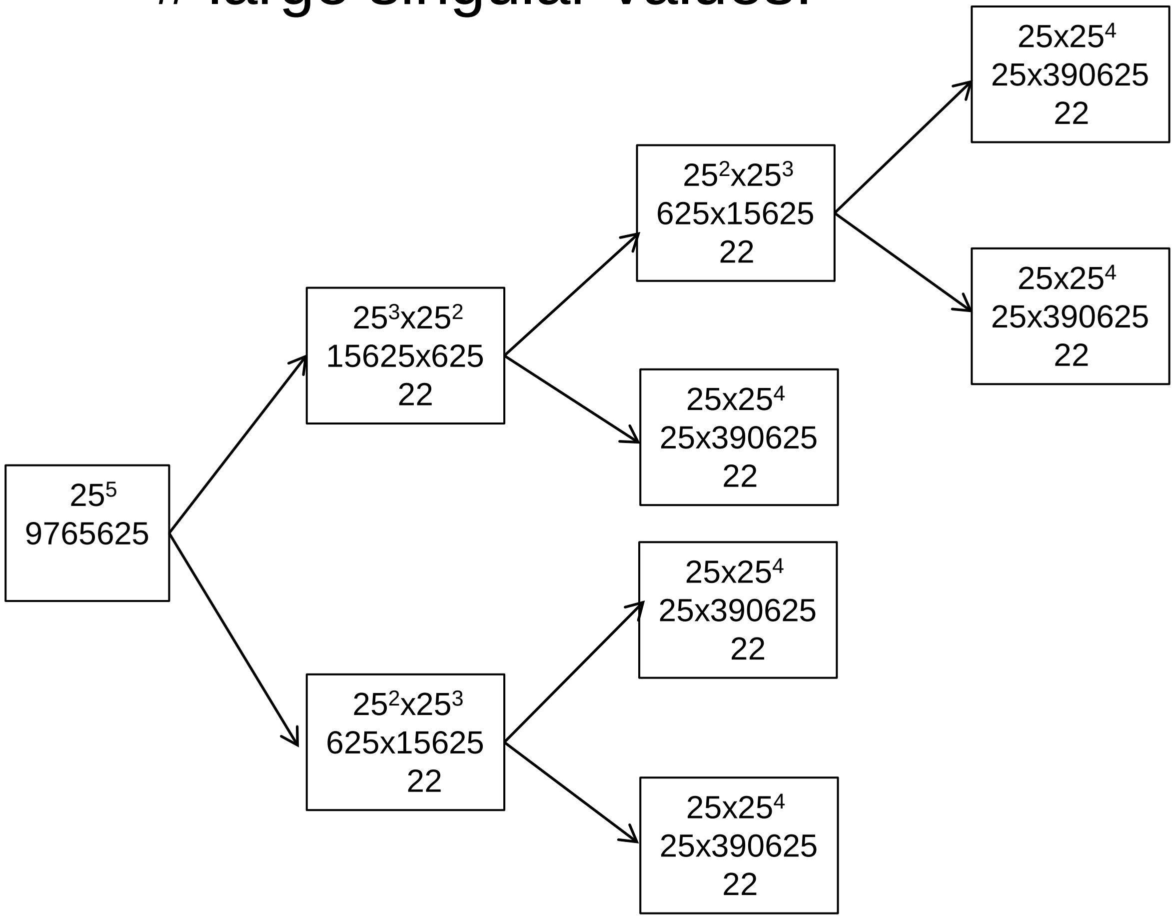
The set of tensors of hierarchical rank at most $(k_t)_{t \in T_I}$ (node-wise) is denoted by

$$H-\text{Tucker}(k_t)_{t \in T_I} := \left\{ A \in IR^I \mid \forall t \in T_I : rank(A^{(t)}) \leq k_t \right\}.$$

# SVD of $A^{(t)}$, d=5, $n_i$=25 → $25^5$

Matrix sizes of $A^{(t)}$ and
# large singular values:



# Nestedness of Matricizations

Let T be a dimension tree and $A \in IR^{I}$ a tensor with hierarchical rank $(k_t)_{t \in T_I}$ .
Let $t \in T$ be a node with sons $s_1$, $s_2$.

Let $(U_t)_i, i = 1,\ldots,k_t$ *be a basis of image*$(A^{(t)})$

$(U_{s_1})_j, j = 1,\ldots,k_{s_1}$ *be a basis of image*$(A^{(s_1)})$

$(U_{s_2})_l, l = 1,\ldots,k_{s_2}$ *be a basis of image*$(A^{(s_2)})$

Then there exist coefficients $(B_t)_{i,j,l}$ such that

$$(U_t)_i = \sum_{j=1}^{k_{s_1}} \sum_{l=1}^{k_{s_2}} (B_t)_{i,j,l} (U_{s_1})_j \otimes (U_{s_2})_l .$$

If the bases $U_{s_1}, U_{s_2}$ are orthogonal then

$$(B_t)_{i,j,l} = \left\langle (U_t)_i , (U_{s_1})_j \otimes (U_{s_2})_l \right\rangle .$$

# Proof:

Consider one column $\left(A^{(t)}\right)_{:,(j_\mu)_{\mu\in t'}}$ of the matricization $A^{(t)}$.

For index $(j_\mu)_{\mu\epsilon t'}$ this column defines a matrix

$$Y_{(j_\mu)_{\mu\in s_1},(j_\mu)_{\mu\in s_2}} := A_{(j_1,\ldots,j_d)}.$$

By assumption the rows and columns of Y are in the span of $\left(U_{s_1}\right)_j, \left(U_{s_2}\right)_l$ :

$$Y = \sum_{j=1}^{k_{s_1}} \sum_{l=1}^{k_{s_2}} c_{j,l} \left(U_{s_1}\right)_j \left(U_{s_2}\right)_l^T$$

for some coefficients $c_{j,l}$.

Therefore, every column of $A^{(t)}$ is a linear combination of

$$\left(U_{s_1}\right)_j \otimes \left(U_{s_2}\right)_l.$$

# t-frame, frame tree:

Let $t \in T$ be a mode cluster and $(k_t)_{t \in T}$ a family of non-negative integers.

We call a matrix $U_t \in IR^{I_t \times k_t}$ a t-frame and a tuple $(U_s)_{s \in T_I}$ of frames a frame tree.

A frame is called orthogonal if the columns are orthogonal.

A frame tree is called orthogonal if each frame (except the root) is orthogonal.

# transfer tensor

A frame tree is nested if for each interior mode cluster t with successors S(t) = $\{t_1, t_2\}$ the following relation holds:

$$span\{(U_t)_i \mid 1 \le i \le k_t\} \subset span\{(U_{t_1})_i \otimes (U_{t_2})_j \mid 1 \le i \le k_{t_1}, 1 \le j \le k_{t_2}\}.$$

The corresponding tensor $B_t \in IR^{k_t \times k_{t_1} \times k_{t_2}}$ relative to the represen-tation of $(U_t)_i$ by $U_{t_1}, U_{t_2}$ is called the *transfer tensor* :

$$(U_t)_i = \sum_{j=1}^{k_{t_1}} \sum_{l=1}^{k_{t_2}} (B_t)_{i,j,l} (U_{t_1})_j \otimes (U_{t_2})_l.$$

# Hierarchical Tucker Format

T is a dimension tree, $(k_t)_{t \in T}$ a family of non-negative integers,

$A \in H - \mathrm{Tucker}\left((k_t)_{t \in T}\right)$ .

Let $(U_t)_{t \in T}$ be a nested frame tree

with transfer tensors $(B_t)_{t \in I(T)}$ and

$$\forall t \in T_I : image\left(A^{(t)}\right) = image(U_t), \qquad A = U_{(1,\dots,d)} .$$

Then the representation $\left((B_t)_{t \in I(T)}, (U_t)_{t \in L(T)}\right)$ is a hierarchical Tucker representation.

The family $(k_t)_{t \in T}$ is the hierarchical representation rank.

Note that the columns of $U_t$ need not to be linear independent!

This representation with orthogonal frame tree is unique upto orthogonal transformations of the t-frames.

# Storage complexity

Again T dimension tree with $A \in H - \mathrm{Tucker}\left((k_t)_{t \in T}\right)$ given in hierarchical Tucker representation $\left((B_t)_{t \in I(T)}, (U_t)_{t \in L(T)}\right)$ and $B_t \in IR^{k_t \times k_{t_1} \times k_{t_2}}$ for $S(t)=\{t_1,t_2\}$, $B_t$ of minimal size.

Then the total storage for all transfer tensors $(B_t)_{t \in I(T)}$ and leaf-frames $(U_t)_{t \in L(T)}$ in terms of number of entries is bounded by

$$Storage\left((B_t)_{t \in I(T)}, (U_t)_{t \in L(T)}\right) \leq (d-1)k^3 + k\sum_{\mu=1}^{d} n_\mu,$$

where $k := \max_{t \in T} k_t$,

is linearly bounded in the dimension d (provided the representation parameter k is uniformly bounded).

# Proof:

For each leaf t={μ} of the dimension tree we have to store the t-frame $U_t \in IR^{n_\mu \times k_t}$
which yields the second term $k\sum_{\mu=1}^{d} n_\mu$.

For all d-1 interior mode clusters we have to store the transfer tensors $B_t \in IR^{k_t \times k_{t_1} \times k_{t_2}}$

Each has at most $k^3$ entries.

# Hierarchical Truncation error

Def.: $T$ a dimension tree, $t \in T$ and $U_t$ an orthogonal t-frame.
Orthogonal frame projection $\pi_t : \mathrm{IR}^I \rightarrow \mathrm{IR}^I$ is defined as

$$(\pi_t A)^{(t)} := U_t U_t^T A^{(t)} \qquad for \quad t \neq \{1,\ldots,d\}$$

$$\pi_{\{1,\ldots,d\}} A := A.$$

Theorem (Hierarchical truncation error):
Dimension tree $T$, $A \in \mathrm{IR}^I$.
Let $A^{best}$ be the best approximation of A in H-Tucker$((k_t)_{t\in T})$,
$\pi_t$ the ortogonal frame projection for t-frame $U_t$ that consists of the left singular vectors of $A^{(t)}$ corresponding to the $k_t$ largest singular values $\sigma_{t,i}$ of $A^{(t)}$.
Then it holds

$$\left\| A - \prod_{t\in T} \pi_t A \right\| \leq \sqrt{\sum_{t\in T} \sum_{i>k_t} \sigma_{t,i}^2} \leq \sqrt{2d-2} \left\| A - A^{best} \right\|.$$

# Proof:

Lemma: It holds $\left\|A-\pi_t\pi_s A\right\|^2 \le \left\|A-\pi_t A\right\|^2 + \left\|A-\pi_s A\right\|^2$

$$\left\|A-\prod_{t\in T}\pi_t A\right\|^2 \le \sum_{t\in T}\left\|A-\pi_t A\right\|^2$$

Proof of Lemma: $\left\|A-\pi_t\pi_s A\right\|^2 = \left\|\left(A-\pi_t A\right)+\pi_t\left(A-\pi_s A\right)\right\|^2 =$

$$\left\|A-\pi_t A\right\|^2 + \left\|\pi_s\left(A-\pi_s A\right)\right\|^2 \le \left\|A-\pi_t A\right\|^2 + \left\|A-\pi_s A\right\|^2$$

Proof of Theorem: It holds $\left\|A-\pi_t A\right\|^2 \le \sum_{i>k_t}\sigma_{t,i}^2 \le \left\|A-A^{best}\right\|^2$

Applying the above Lemma and the result on the number of nodes of a dimension tree yields:

$$\left\|A-\prod_{t\in T}\pi_t A\right\|^2 \le \sum_{t\in T}\sum_{i>k_t}\sigma_{t,i}^2 \le \left(2d-2\right)\left\|A-A^{best}\right\|^2$$

Can be improved to $(2d - 3)$

# Properties of H-Tucker Format

The set H-Tucker($(k_t)_{t\in T}$) is
- a closed set in $IR^I$ , but
- is not a linear space (the rank increases by linear combinations)

Storage complexity: With $n := \max_{\mu=1,\ldots,d} n_\mu$ and $k := \max_{t\in T} k_t$

$$Storage(A) \le \sum_{\mu=1}^{d} n_\mu k_\mu + \sum_{t\in I(T), sons(t)=\{t_1,t_2\}} k_t k_{t_1} k_{t_2} \le dnk + dk^3$$

All tensors of canonical rank k are contained in H-Tucker($(k_t)_{t\in T}$) (also all tensors of boarder rank k).
The set H-Tucker($(k_t)_{t\in T}$) is much thinner than the Tucker format because we impose additional rank conditions.

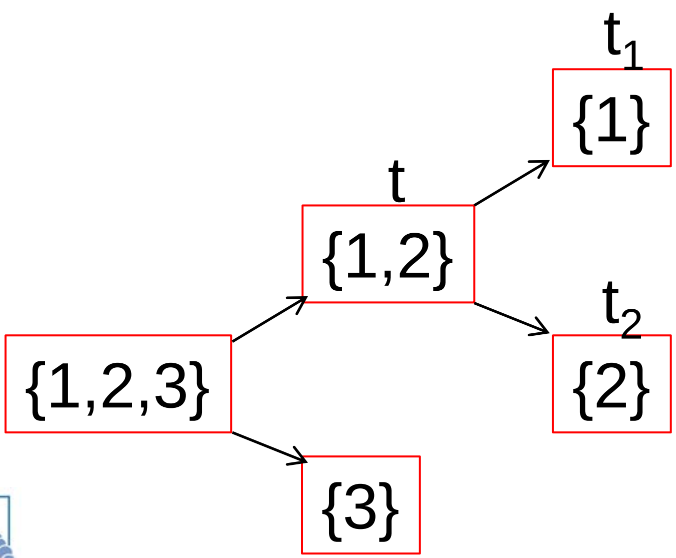
# Example

$A \in IR^{3\times3\times3}$ :  $A^{(\{1,2\})} := [u_1 \otimes q_1 \quad | \quad u_2 \otimes q_2 \quad | \quad u_1 \otimes q_2],$

$$A^{(\{1,2\})}_{(i,j),l} = \begin{cases} (u_1 \otimes q_1)_{i,j} & if \quad l=1 \\ (u_2 \otimes q_2)_{i,j} & if \quad l=2 \\ (u_1 \otimes q_2)_{i,j} & if \quad l=3 \end{cases}$$

9 x 3 - matrix

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; q_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \; q_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$



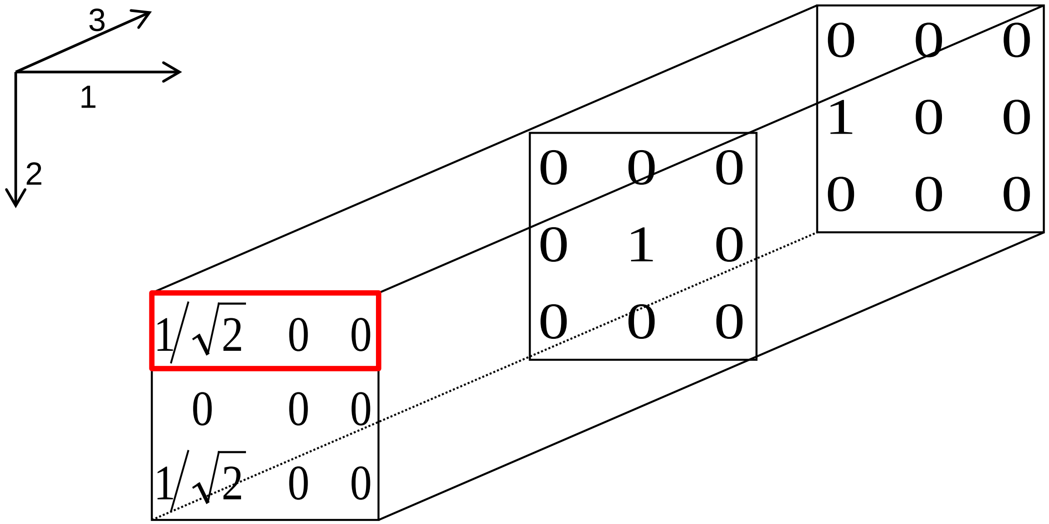

Consider orthogonal mode frames:

$$U_t := [u_1 \otimes q_1 \,|\, u_2 \otimes q_2 \,|\, q_1 \otimes q_2],$$

$$U_{t_1} := [u_1 \,|\, u_2], \; U_{t_2} := [q_1 \,|\, q_2]$$

$$A^{(\{1,2\})} = \begin{pmatrix} u_1 \otimes q_1 & u_2 \otimes q_2 & u_1 \otimes q_2 \end{pmatrix} =$$

$$= \begin{pmatrix} 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$




$$u_1 \otimes q_2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$u_1 \otimes q_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \qquad u_2 \otimes q_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$A^{(1)} = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(\pi_{t_1} A)^{t_1} = U_{t_1} U_{t_1}^T A^{(t_1 = 1)} = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} u_1^T \\ u_2^T \end{pmatrix} A^{(1)} =$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} A^{(1)} \implies rank\left((\pi_{t_1} A)^{t_1}\right) = 2.$$

$$\pi_{t_1}\pi_{t_2} \Leftrightarrow U_{t_1}, U_{t_2} \Leftrightarrow U_{t_1}U_{t_1}^T, U_{t_2}U_{t_2}^T \Leftrightarrow$$

$$\Leftrightarrow Q = U_{t_1}U_{t_1}^T \otimes U_{t_2}U_{t_2}^T = \left(u_1u_1^T + u_2u_2^T\right) \otimes \left(q_1q_1^T + q_2q_2^T\right)$$

$$QU_t = \left(\left(u_1u_1^T + u_2u_2^T\right) \otimes \left(q_1q_1^T + q_2q_2^T\right)\right)\left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad q_1 \otimes q_2\right) =$$

$$= \left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad \frac{1}{\sqrt{2}} u_1 \otimes q_2\right)$$

$$\left(\pi_t\pi_{t_1}\pi_{t_2}A\right)^{(\{1,2\})} = U_tU_t^T QA^{(\{1,2\})} = U_t\left(QU_t\right)^T A^{(\{1,2\})}$$

$$= U_t\left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad \frac{1}{\sqrt{2}} u_1 \otimes q_2\right)^T \left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad u_1 \otimes q_2\right) =$$

$$= \left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad q_1 \otimes q_2\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/\sqrt{2} \end{pmatrix} = \left(u_1 \otimes q_1 \quad u_2 \otimes q_2 \quad \frac{1}{\sqrt{2}} q_1 \otimes q_2\right)$$

3 x 9 matrix

$$rank\left(\left(\pi_t \pi_{t_1} \pi_{t_2} A\right)^{(1)}\right) = rank\left(\begin{matrix} u_1 q_1^T & u_2 q_2^T & q_1\left(\frac{1}{\sqrt{2}} q_2\right)^T \end{matrix}\right) = 3$$

because $u_1$, $u_2$, $q_1$ are linearly independent.

The first projection $\pi_{t_1}\pi_{t_2}$ maps A into Tucker(2,2,3), but after the coarser projection $\pi_t$ the 1-mode rank is 3 and thus $\pi_t \pi_{t_1} \pi_{t_2} A \notin \text{Tucker}(2,2,3)$.
This is because $\pi_t$ mixes the $t_1$-frame and the $t_2$-frame.

Tucker(2,2,3) means ranks in the standard Tucker format.

# Root-to-Leaves Truncation

*Input: tensor A, dimension tree* $T_I$*, target rank* $((k_t)_{t \in T})$.
*For each singelton* $t \in L(T_I)$ *do*

*Compute SVD of* $A^{(t)}$*, store dominant* $k_t$ *left singular vectors in the columns of the t-frame* $U_t$.

*For l=p-1,…,0 do*

*For each mode cluster* $t \in L(T_I)$ *on level l do*

*Compute SVD of* $A^{(t)}$*, store dominant* $k_t$ *left singular vectors in the columns of the t-frame* $U_t$.

*Let* $U_{t_1}$ *and* $U_{t_2}$ *denote the frames for the successors of t on level l+1.*

*Compute the entries of the transfer tensor*

$$(B_t)_{i,j,\nu} := \left\langle (U_t)_i , (U_{t_1})_j \otimes (U_{t_2})_\nu \right\rangle$$

*Compute the entries of the root (with sons $t_1$, $t_2$ ) transfer tensor:*

$$\left(B_{\{1,\dots,d\}}\right)_{1,j,\nu} := \left\langle A, \left(U_{t_1}\right)_j \otimes \left(U_{t_2}\right)_\nu \right\rangle$$

*Return H-Tucker representation* $\left(\left(U_t\right)_{t\in L(T_I)}, \left(B_t\right)_{t\in I(T_I)}\right)$

*for* $A_H \in H - \text{Tucker}\left(\left(k_t\right)_{t\in T_I}\right).$

Complexity: $O\left(\left(n_1 \cdot \ldots \cdot n_d\right)^{3/2}\right)$

# Brothers and related matricizations

For dimension tree $T_I$ and non-root mode cluster t with father f we define the unique mode cluster $\bar{t}$ as the brother of t such that $f = t \,\dot{\cup}\, \bar{t}$ .

Let $T_I$ a dimension tree with interior node $t = t_1 \,\dot{\cup}\, t_2$ .
Assume the matricization $A^{(t)} = \sum_{\nu=1}^{k} u_\nu v_\nu^T$

and the representation

$$u_\nu = \sum_{j=1}^{k_1} \sum_{l=1}^{k_2} c_{\nu,j,l} x_j \otimes y_l, \qquad x_j \in IR^{I_{i_1}}, y_l \in IR^{I_{i_2}}, \quad \nu = 1,\dots,k$$

This gives the matricization

$$A^{(t_1)} = \sum_{j=1}^{k_1} x_j \left( \sum_{\nu=1}^{k} \sum_{l=1}^{k_2} c_{\nu,j,l} y_l \otimes v_\nu \right)^T$$

# Proof:

$$A_{(i_1,\dots,i_d)} = A^{(t)}_{(i_\mu)_{\mu\in t},(i_{\mu'})_{\mu'\in t'}} = \sum_{\nu=1}^{k}\sum_{j=1}^{k_1}\sum_{l=1}^{k_2} c_{\nu,j,l}\left(x_j\right)_{(i_{\mu_1})_{\mu_1\in t_1}}\left(y_l\right)_{(i_{\mu_2})_{\mu_2'\in t_2}}\left(v_\nu\right)_{(i_{\mu'})_{\mu'\in t'}}$$

$$= \sum_{j=1}^{k_1}\left(x_j\right)_{(i_{\mu_1})_{\mu_1\in t_1}}\left(\sum_{\nu=1}^{k}\sum_{l=1}^{k_2} c_{\nu,j,l}\left(y_l\right)_{(i_{\mu_2})_{\mu_2\in t_2}}\otimes\left(v_\nu\right)_{(i_{\mu'})_{\mu'\in t'}}\right)$$

$$= \sum_{j=1}^{k_1}\left(x_j\right)_{(i_{\mu_1})_{\mu_1\in t_1}}\left(\sum_{\nu=1}^{k}\sum_{l=1}^{k_2} c_{\nu,j,l}\, y_l\otimes v_\nu\right)_{(i_{\mu'})_{\mu'\in t'}}.$$

# Matricization in H-Tucker format

$T_I$ dimension tree, $A \in$ H-Tucker$((k_t)_{t \in I})$ with nested orthogonal frame tree $(U_t)_{t \in T_I}$ and transfer tensors $(B_t)_{t \in T_I}$ .

For p>0 let $\mathrm{Root}(T_I) = t_0, t_1, \ldots, t_{p-1}, t_p = t$ a path of length p.

Let $\overline{U}^1, \ldots, \overline{U}^p$ denote the frames of the corresponding brothers,
$B^0, \ldots, B^{p-1}$ the corresponding transfer tensors, and
$k_0, \ldots, k_p$ the corresponding representation ranks.

We always assume that the brother is always first

$$\left(U_{t_l}\right)_\nu = \sum_i \sum_j B^l_{\nu,i,j} \overline{U}_i^{l+1} \otimes \left(U_{t_{l+1}}\right)_j$$

Then it holds $A^{(t)} = \sum_{\nu=1}^{k_1} (U_t)_\nu (V_t)_\nu^T = U_t V_t^T$ with complementary frame

$$(V_t)_{j_p} = \sum \sum \cdots \sum \sum \sum B^0_{1,i_1,j_1} \cdots B^{p-1}_{j_{p-1},i_p,j_p} \overline{U}^1_{i_1} \otimes \cdots \otimes \overline{U}^p_{i_p}.$$

# Accumulated Transfer Tensors

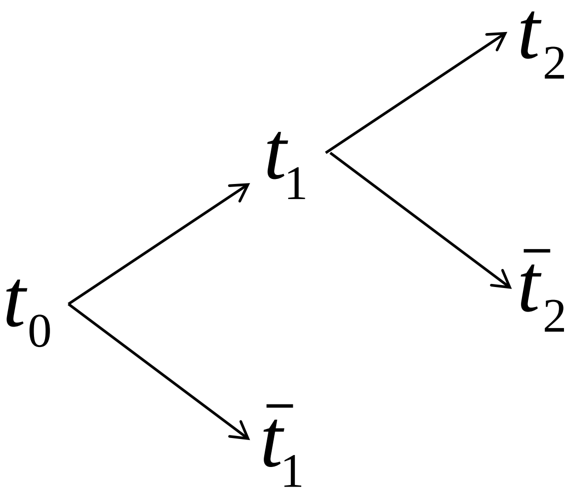
$$\left(\hat{B}^1\right)_{j_1,s_1} := \sum_{i_1=1}^{\bar{k}_1} B^0_{1,i_1,j_1} B^0_{1,i_1,s_1},$$

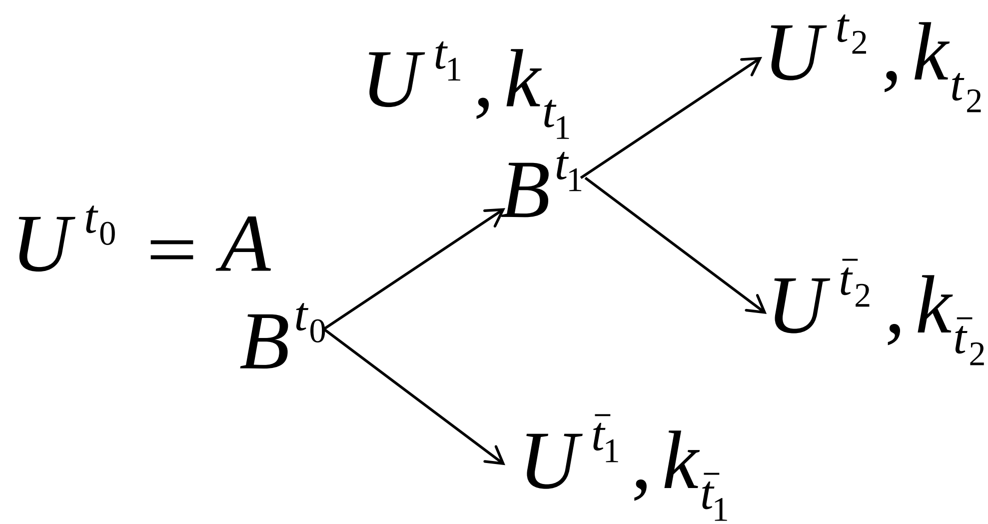
$$\left(\hat{B}^l\right)_{j_l,s_l} := \sum_{s_{l-1}=1}^{k_{l-1}} \sum_{i_l=1}^{\bar{k}_l} \left( \sum_{j_{l-1}}^{k_{l-1}} \left(\hat{B}^{l-1}\right)_{j_{l-1},s_{l-1}} B^{l-1}_{j_{l-1},i_l,j_l} \right) B^{l-1}_{s_{l-1},i_l,s_l}, \qquad l = 2,\ldots,p,$$

$$\hat{B}_t := \hat{B}^p.$$

are useful for computing all matricizations out of the transfer tensors.

# H-Tucker




$$(U_t)_i = \sum_{j=1}^{k_{t_1}} \sum_{l=1}^{k_{t_2}} (B_t)_{i,j,l} (U_{t_1})_j \otimes (U_{t_2})_l$$

# (non-orthogonal) H-Tucker for CP

CP-tensor: $A = \sum_{i=1}^{k} \bigotimes_{\mu=1}^{d} a_{i,\mu}, \qquad a_{i,\mu} \in IR^{I_\mu}$ .

H-Tucker representation:

Leaves: $\forall t = \{\mu\} \in L(T_I): \qquad (U_t)_i := a_{i,\mu}, \quad i = 1,\dots,k, \quad k_\mu := k,$

Interior nodes, transfer tensors:

$$\forall t \in I(T_I) \setminus Root(T_I): \qquad (B_t)_{i,j,l} := \begin{cases} 1 & if \ \ i = j = l \\ 0 & otherwise \end{cases}, \quad B_t \in IR^{k \times k \times k}, \quad k_t := k.$$

Root transfer tensor:

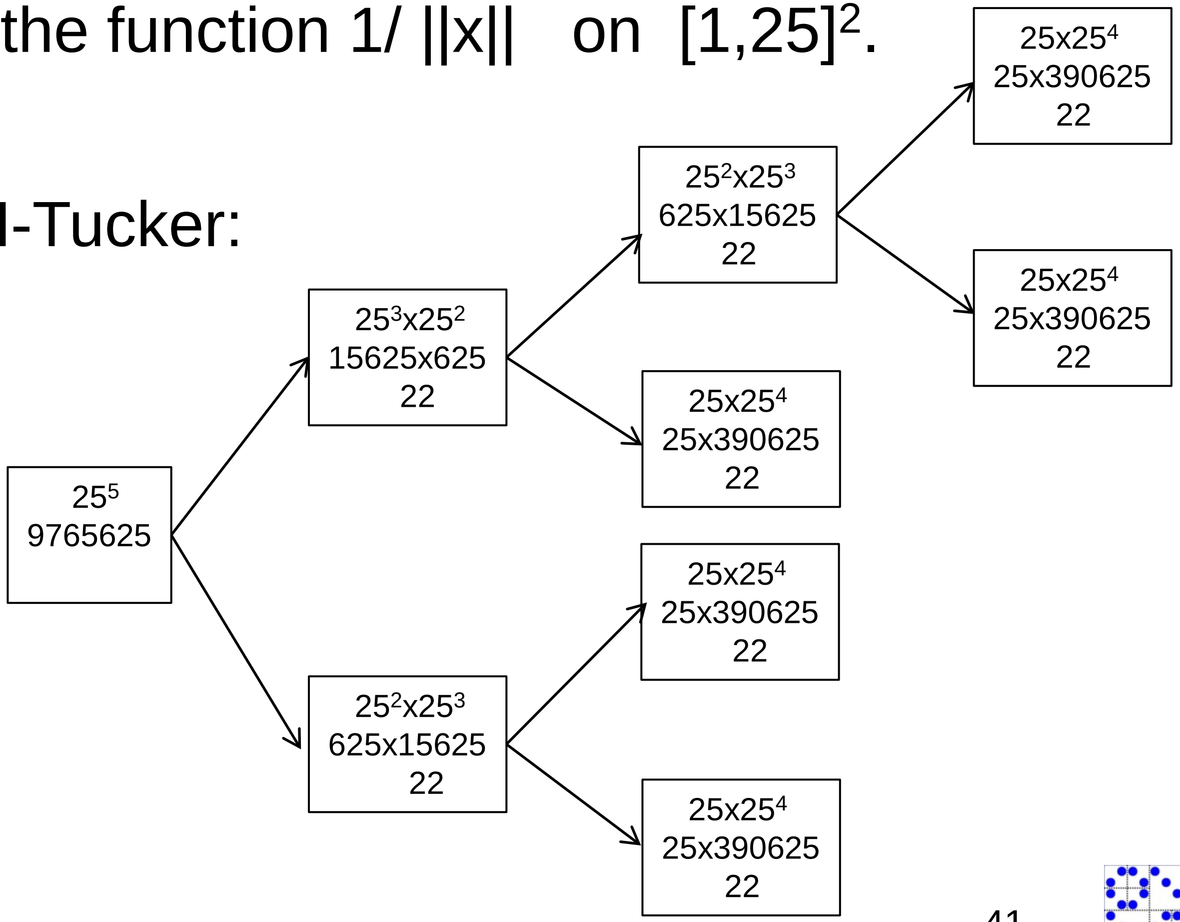
$$(B_{\{1,\dots,d\}})_{i,j,l} := \begin{cases} 1 & if \ \ j = l \\ 0 & otherwise \end{cases}, \quad B_{\{1,\dots,d\}} \in IR^{1 \times k \times k}, \quad k_{\{1,\dots,d\}} := 1.$$

# Examples

Consider tensor $A_{(i_1,\dots,i_d)} := \left(\sum_{\mu=1}^{d} i_\mu^2\right)^{-1/2}, \quad d = 5, \quad n_\mu = 25$

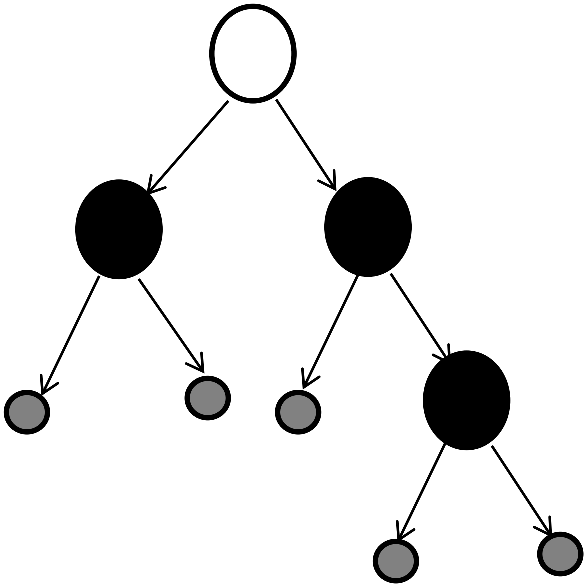
as discretization of the function 1/ ||x|| on $[1,25]^2$.

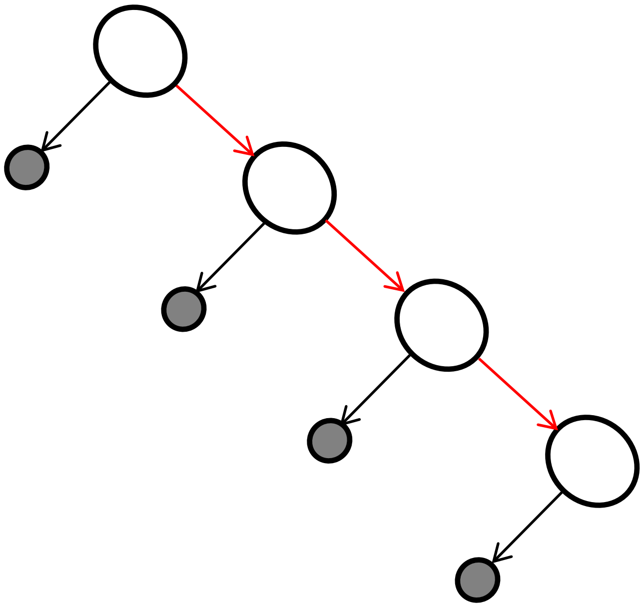
Approximation by H-Tucker:



# Tensor Train

Instead of complete binary tree we can also consider a linear list:





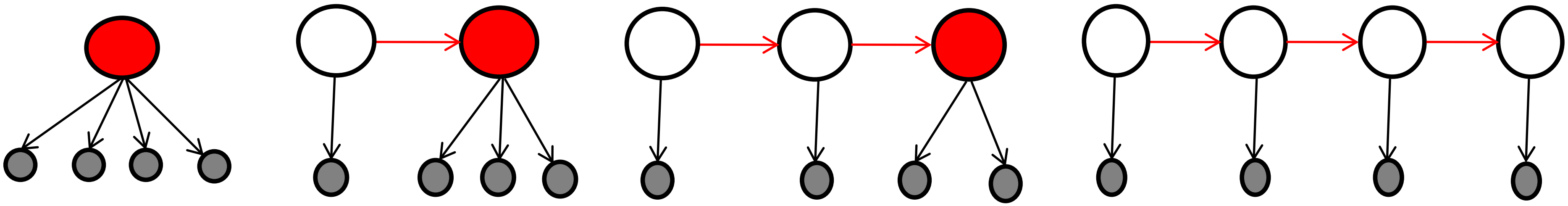
Partitioning the index sets in half:
$(i_1,\ldots,i_{2d}) \rightarrow (i_1,\ldots,i_d)(i_{d+1},\ldots,i_{2d})$

Partitioning the index sets in one and rest:
$(i_1,\ldots,i_{d+1}) \rightarrow (i_1,\ldots,i_d)(i_{d+1})$

# Tensor train by recursive splitting:

CP: $$A_{(i_1,\dots,i_d)} = \sum_{j=1}^{D} u_{1;i_1,j}\, u_{2;i_2,j} \cdots u_{d;i_d,j}$$

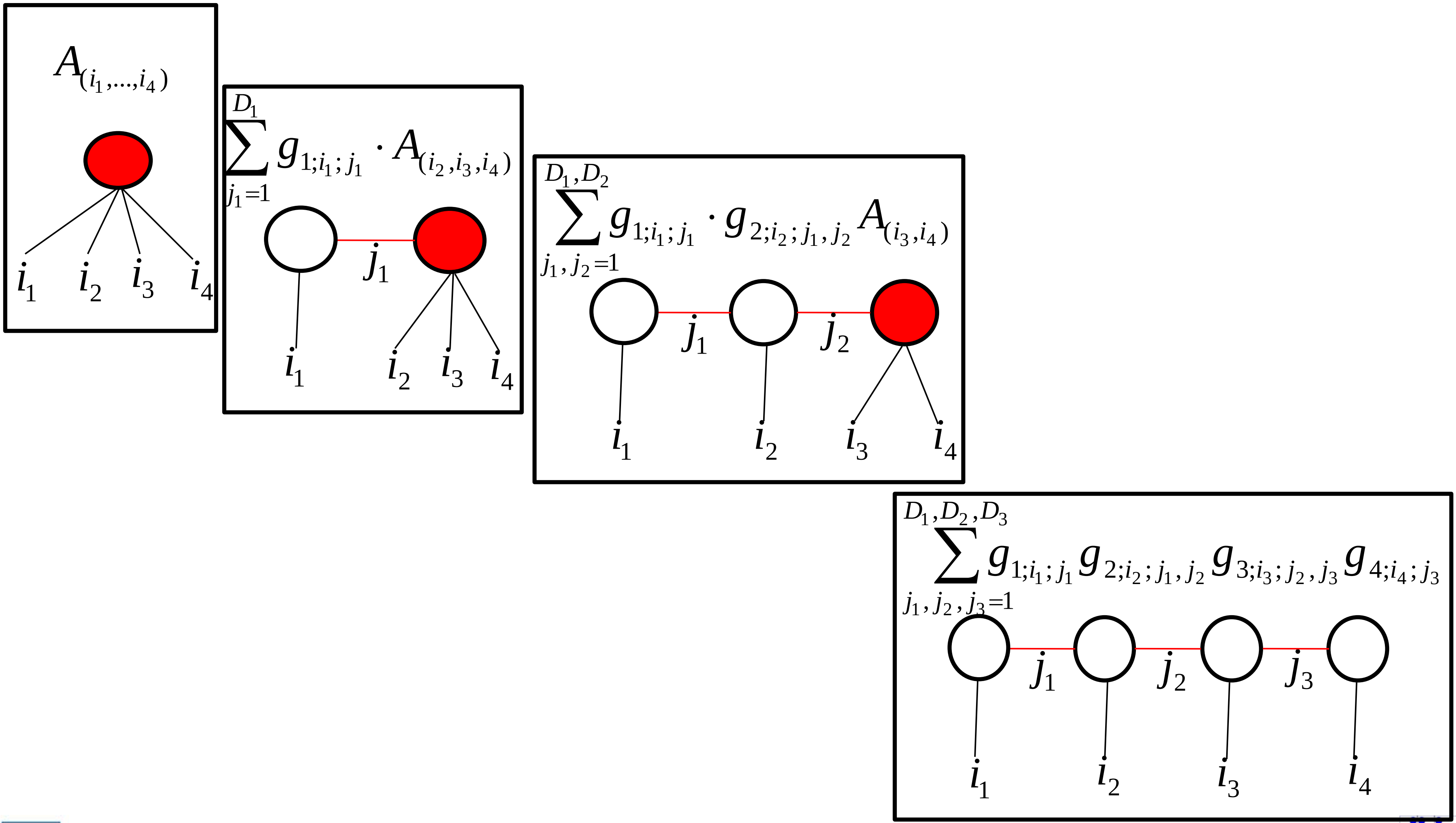
TT: $$A_{(i_1,\dots,i_d)} = \sum_{j_1,j_2,\dots,j_d=1}^{D_1,D_2,\dots D_d} g_{1;i_1;j_1} \cdot g_{2;i_2;j_1,j_2} \cdots g_{d-1;i_{d-1};j_{d-2},j_{d-1}} \cdot g_{d;i_d;j_d}$$



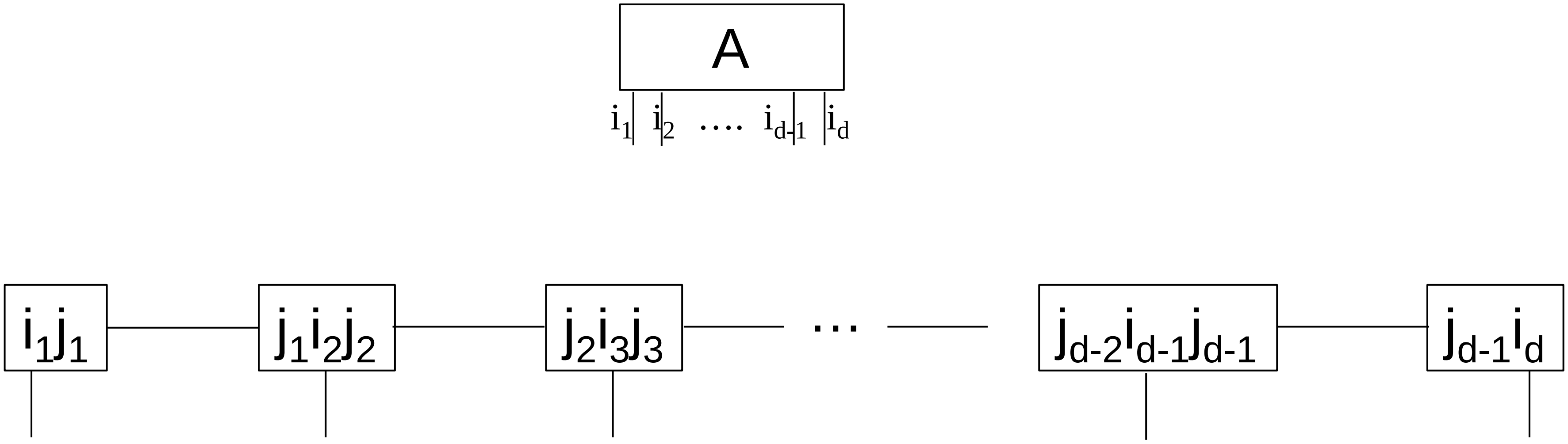
$$A_{(i_1,\dots,i_d)} = \sum_{j_1=1}^{D_1} g_{1;i_1;j_1} \cdot A_{j_1,(i_2,\dots,i_d)} = \sum_{j_1=1}^{D_1}\sum_{j_2=1}^{D_2} g_{1;i_1;j_1} \cdot g_{2;i_2;j_1,j_2} \cdot A_{j_2,(i_3,\dots,i_d)} =$$

$$= \sum_{j_1=1}^{D_1}\sum_{j_2=1}^{D_2} \cdots \sum_{j_{d-2}=1}^{D_{d-2}}\sum_{j_{d-1}=1}^{D_{d-1}} g_{1;i_1;j_1} \cdot g_{2;i_2;j_1,j_2} \cdots\cdot\; g_{d-2;i_{d-1};j_{d-2},j_{d-1}} \cdot g_{d;i_d;j_{d-1}}$$

# Tensor train by recursive splitting:



# Tensor Train Network



# Matrix Formulation

$$A_{(i_1,\ldots,i_d)} = \sum_{j_1=1}^{D_1} g_{1;i_1;j_1} \cdot A_{(i_2,\ldots,i_d)} = \sum_{j_1=1}^{D_1}\sum_{j_2=1}^{D_2} g_{1;i_1;j_1} \cdot g_{2;i_2;j_1,j_2} \cdot A_{(i_3,\ldots,i_d)} =$$

$$= \sum_{j_1=1}^{D_1}\sum_{j_2=1}^{D_2}\cdots\sum_{j_{d-2}=1}^{D_{d-2}}\sum_{j_{d-1}=1}^{D_{d-1}} g_{1;i_1;j_1} \cdot g_{2;i_2;j_1,j_2} \cdot\cdots\cdot g_{d-1;i_{d-1};j_{d-2},j_{d-1}} \cdot g_{d;i_d;j_{d-1}} =$$

$$= G_1^{i_1} \cdot G_2^{i_2} \cdot\cdots\cdot G_{d-1}^{i_{d-1}} \cdot G_d^{i_d}$$

with $D_{j-1}$ x $D_j$ – matrices $G_j^{i_j}$ as core tensors.

The j indices are related to the matrix product.

$i_j = 1,\ldots,n_j;\quad G_j^{i_j} \in IR^{D_{j-1}\times D_j} \quad with \quad D_0 = D_d = 1;$

$G_1^{n_1}$ and $G_d^{n_d}$ are vectors (row, resp. column vector).

# Visualization

$$A = G_1^{i_1} G_2^{i_2} \cdots G_{d-1}^{i_{d-1}} G_d^{i_d}$$

Visualization:

$$A_{i_1,\dots,i_d} = \begin{array}{c|c|c|c|c}
G_1^1 & G_2^1 & \cdots & G_{d-1}^1 & G_d^1 \\
\vdots & \vdots & \cdots & \vdots & \vdots \\
G_1^{n_1} & G_2^{n_2} & \cdots & G_{d-1}^{n_{d-1}} & G_d^{n_d} \\
\hline
i_1 & i_2 & \cdots & i_{d-1} & i_d \\
1 & \cdot & \cdots\cdots & \cdot & n_d
\end{array}$$

# Periodic case:

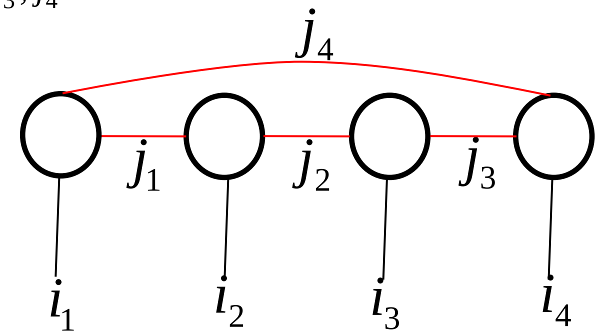
$$A_{(i_1,\dots,i_d)} = \sum_{j_1=1}^{D_1}\sum_{j_2=1}^{D_2}\cdots\sum_{j_{d-1}=1}^{D_{d-1}}\boxed{\sum_{j_d=1}^{D_d}} g_{1;i_1;\boxed{j_d},j_1}\cdot g_{2;i_2;j_1,j_2}\cdot\dots\cdot g_{d-1;i_{d-1};j_{d-2},j_{d-1}}\cdot g_{d;i_d;j_{d-1},\boxed{j_d}} =$$

$$= trace\left(G_1^{i_1}\cdot G_2^{i_2}\cdot\dots\cdot G_{d-1}^{i_{d-1}}\cdot G_d^{i_d}\right)$$

With additional index $j_d$, and summation over $j_d$ represented by trace summation.

$$i_j = 1,\dots,n_j; \qquad G_j^{i_j}\in IR^{D_{j-1}\times D_j} \qquad for \quad j=1,\dots,d, \qquad D_0 = D_d$$

$$\sum_{j_1,j_2,j_3,j_4=1}^{D_1,D_2,D_3,D_4} g_{1;i_1;j_4,j_1}\, g_{2;i_2;j_1,j_2}\, g_{3;i_3;j_2,j_3}\, g_{4;i_4;j_3,j_4}$$



# Rank - Unfolding

$$A_p := A_{\{i_1,\ldots,i_p\},\{i_{p+1},\ldots,i_d\}}, \qquad r_p := rank(A_p)$$

$D_p$ are called compression ranks = size of matrices

Theorem: If for each unfolding rank($A_p$) = $r_p$ = $D_p$,
then there exists a TT decomposition with
compression ranks not higher than $D_p$.

Proof: $A_1 = UV^T$ : $\quad A_{i_1,(i_2\ldots i_d)} = \sum_{j_1=1}^{n_1} U_{i_1,j_1} \cdot V_{j_1,\{i_2\ldots i_d\}}$

Consider **V** also as a d-1 tensor with indices $V_{\{j_1 i_2\},\{i_3,\ldots,i_d\}}$ with „long index“ $j_1 i_2$ varying from 1 to $n_1 n_2$ and consider all unfoldings of **V** resulting in $V_2,\ldots,V_d$ . We show: rank($V_p$) ≤ $r_p$ .

# Proof

To prove rank($V_p$) ≤ $r_p$ we express V as

componentwise

$$A_1 = UV^T \Rightarrow V = A_1^T U^T \left(U^T U\right)^{-1} = A_1^T W \qquad \Rightarrow V_{\{j_1,i_2,i_3,\ldots,i_d\}} = \sum_{i_1=1}^{n_1} A_{\{i_1,\ldots,i_d\}} W_{i_1,j_1}$$

Because the p-th mode has compression rank $r_p$ it holds

componentwise

$$A_{i_1,\ldots,i_d} = \sum_{\beta=1}^{r_p} F_{i_1,\ldots,i_p,\beta} G_{\beta,i_{p+1},\ldots,i_d} \Rightarrow$$

$$V_p = V_{\{j_1 i_2,\ldots,i_p\}\{i_{p+1},\ldots,i_d\}} = \sum_{i_1=1}^{n_1} \sum_{\beta=1}^{r_p} W_{i_1,j_1} F_{i_1,\ldots,i_p,\beta} G_{\beta,i_{p+1},\ldots,i_d} =$$

$$= \sum_{\beta=1}^{r_p} H_{j_1,i_2,\ldots,i_p,\beta} G_{\beta,i_{p+1},\ldots,i_d} \qquad \text{with} \qquad H_{j_1 i_2,i_3,\ldots,i_p,\beta} = \sum_{i_1=1}^{n_1} F_{i_1,\ldots,i_p,\beta} \cdot W_{i_1,j_1}$$

resulting in rank($V_p$) ≤ $r_p$.

Repeat what we have started with A and $i_1$ now for V and $i_2$. …

# Complexity

Following the proof the TT form of a general tensor A can be derived by successive SVD of matrix unfoldings.

The number of parameters in the TT format is bounded by

$$(d-2)nD^2 + 2nD$$

where $n = \max(n_1,\ldots,n_d)$, $D = \max(r_1,\ldots,r_d)$.

Proof: The number of core tensor matrices is bounded by n and the number of entries is bounded by $D^2$ for interior core tensors, resp. D for the first/last vector.

In the periodic case the bound is given by $dnD^2$.

# Approximation

Suppose that the unfolding matrices are only approximated by low rank terms

$$A_p = R_p + E_p, \qquad rank(R_p) = r_p, \quad \|E_p\|_F = \varepsilon_p, \quad p = 1,\ldots,d-1$$

Theorem: With the algorithm we can compute for a given Tensor A an TT-tensor B with ranks $r_p$ and

$$\|A-B\|_F \leq \sqrt{\sum_{p=1}^{d-1} \varepsilon_p^2}$$

$$\|A-B\|_F \leq \sqrt{d-1}\,\|A-A^{best}\|_F$$

# Recompression TT → TT

Let us assume that we have already given a tensor in the TT format

$$A = G_1^{i_1} G_2^{i_2} \cdots G_{d-1}^{i_{d-1}} G_d^{i_d}$$

We want the derive minimal ranks, resp. compute approximations with smaller ranks.
More tomorrow.

# Basic Operations

Addition: $C_1^{i_1} \cdots C_d^{i_d} = C = A + B = A_1^{i_1} \cdots A_d^{i_d} + B_1^{i_1} \cdots B_d^{i_d} =$

$$= \begin{pmatrix} A_1^{i_1} & B_1^{i_1} \end{pmatrix} \begin{pmatrix} A_2^{i_2} & 0 \\ 0 & B_2^{i_2} \end{pmatrix} \cdots \begin{pmatrix} A_{d-1}^{i_{d-1}} & 0 \\ 0 & B_{d-1}^{i_{d-1}} \end{pmatrix} \begin{pmatrix} A_d^{i_d} \\ B_d^{i_d} \end{pmatrix}$$

Periodic: $trace\left(C_1^{i_1} \cdots C_d^{i_d}\right) = C = A + B = trace\left(A_1^{i_1} \cdots A_d^{i_d} + B_1^{i_1} \cdots B_d^{i_d}\right) =$

$$= trace\left( \begin{pmatrix} A_1^{i_1} & 0 \\ 0 & B_1^{i_1} \end{pmatrix} \cdots \begin{pmatrix} A_d^{i_d} & 0 \\ 0 & B_d^{i_d} \end{pmatrix} \right)$$

Scalar multiplication:

$$\alpha A_{i_1 \ldots i_d} = \alpha A_1^{i_1} \cdots A_d^{i_d} = \left(\alpha A_1^{i_1}\right) A_2^{i_2} \cdots A_d^{i_d}$$

# Hadamard product

$$C_1^{i_1} \cdots C_d^{i_d} = C = A \circ B = \left(A_1^{i_1} \cdots A_d^{i_d}\right)\left(B_1^{i_1} \cdots B_d^{i_d}\right) =$$
$$= \left(A_1^{i_1} \cdots A_d^{i_d}\right) \otimes \left(B_1^{i_1} \cdots B_d^{i_d}\right) =$$
$$= \left(A_1^{i_1} \otimes B_1^{i_1}\right) \cdots \left(A_d^{i_d} \otimes B_d^{i_d}\right)$$

Inner product: $\langle A, B \rangle = \sum_{i_1,\dots,i_d} A_{i_1 \dots i_d} B_{i_1 \dots i_d} = \sum_{i_1,\dots,i_d} C_{i_1 \dots i_d}$

$$= \sum_{i_1,\dots,i_d} C_1^{i_1} C_2^{i_2} \cdots C_{d-1}^{i_{d-1}} C_d^{i_d}$$

Take the Hadamard product to derive C in TT-format, and then compute the contractions with vectors of all ones.

# Inner Product:

$$\langle A,B\rangle = \sum_{i_1,\dots,i_d} A_{i_1\dots i_d} B_{i_1\dots i_d} = \sum_{i_1,\dots,i_d} C_{i_1\dots i_d}$$

$$= \sum_{i_1,\dots,i_d} C_1^{i_1} C_2^{i_2} \cdots C_{d-1}^{i_{d-1}} C_d^{i_d} =$$

$$= \sum_{i_1,\dots,i_d} \sum_{j_1,\dots,j_{d-1}} C_{1,j_1}^{i_1} C_{2,j_1 j_2}^{i_2} \cdots C_{d-1,j_{d-2}j_{d-1}}^{i_{d-1}} C_{d,j_{d-1}}^{i_d} =$$

$$= \sum \left( \underbrace{\sum_{j_1}}_{D_1 D_2} \left( \underbrace{\sum_{i_1} C_{1,j_1}^{i_1}}_{D_1 n_1} \right) \left( \underbrace{\sum_{i_2} C_{2,j_1 j_2}^{i_2}}_{D_1 D_2 n_2} \right) \right) \cdots$$

$$O(dD^2 n)$$

# Equivalent formulation for Vector

$$x = \sum_{i_1 \ldots i_d} A_{i_1 \ldots i_d} e_{i_1 \ldots i_d} =$$

$$= \sum_{i_1 \ldots i_d} \left( \sum_{j_1 \ldots j_{d-1}} A_{1,j_1}^{i_1} A_{2,j_1 j_2}^{i_2} \cdots A_{d-1,j_{d-2} j_{d-1}}^{i_{d-1}} A_{d j_{d-1}}^{i_d} \right) \left( e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_d} \right) =$$

$$= \sum_{j_1 \ldots j_{d-1}} \sum_{i_1 \ldots i_d} \left( A_{1,j_1}^{i_1} e_{i_1} \right) \otimes \left( A_{2,j_1 j_2}^{i_2} e_{i_2} \right) \otimes \cdots \otimes \left( A_{d,j_{d-1}}^{i_d} e_{i_d} \right)$$

$$= \sum_{j_1 \ldots j_{d-1}} \left( \sum_{i_1} A_{1,j_1}^{i_1} e_{i_1} \right) \otimes \cdots \otimes \left( \sum_{i_d} A_{d,j_{d-1}}^{i_d} e_{i_d} \right)$$

$$= \sum_{j_1 \ldots j_{d-1}} u_{1,j_1} \otimes u_{2,j_1 j_2} \otimes \cdots \otimes u_{d-1,j_{d-2} j_{d-1}} \otimes u_{d,j_{d-1}}$$

with vectors $u_{k,j_{k-1} j_k}$ of length $n_k$. Compare CP.

# Inner Product for vector:

$$y^T x = \sum_{i_1'\dots i_d'} B_{i_1'\dots i_d'} e^T_{i_1'\dots i_d'} \sum_{i_1\dots i_d} A_{i_1\dots i_d} e_{i_1\dots i_d} =$$

$$= \sum_{i_1\dots i_d} B_{i_1\dots i_d} A_{i_1\dots i_d} = \sum_{i_1\dots i_d} C_{i_1\dots i_d} =$$

$$= \left( \sum_{j_1'\dots j_{d-1}'} v_{1,j_1'} \otimes v_{2,j_1'j_2'} \otimes \cdots \otimes v_{d,j_{d-1}'} \right)^T \cdot \sum_{j_1\dots j_{d-1}} u_{1,j_1} \otimes u_{2,j_1j_2} \otimes \cdots \otimes u_{d,j_{d-1}}$$

$$= \sum_{j_1 j_1'\dots j_{d-1} j_{d-1}'} \left(v^T_{1,j_1'} u_{1,j_1}\right) \cdot \left(v^T_{2,j_1'j_2'} u_{2,j_1j_2}\right) \cdots \left(v^T_{d,j_{d-1}'} u_{d,j_{d-1}}\right)$$

$$= \sum_{j_1 j_1'\dots j_{d-1} j_{d-1}'} w_{1,j_1'j_1} \cdot w_{2,j_1'j_1j_2'j_2} \cdots w_{d,j_{d-1}'j_{d-1}}$$

# Special Cases

Representing unit vector $e_i$:

$$e_i = e_{1,j_1} \otimes e_{2,j_2} \otimes \cdots \otimes e_{d,j_d} =$$

$$= a_1^{i_1} a_2^{i_2} \cdots a_{d-1}^{i_{d-1}} a_d^{i_d}$$

with $a_k^{i_k} = \delta_{i_k, j_k} = \begin{cases} 1 & if \quad i_k = j_k \\ 0 & if \quad i_k \neq j_k \end{cases}$

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left( a_1^{i_1} a_2^{i_2} \right)_{i_1 i_2}, \qquad a_1^i = a_2^i = \delta_{1,i}$$

# Examples

$$x_{i_1 \dots i_d} = A_1^{i_1} A_2^{i_2} \cdots A_{d-1}^{i_{d-1}} A_d^{i_d}$$

$$\begin{array}{cccc|c} (1) & (1) & \cdots & (1) & A_k^0 \\ \hline (0) & (0) & \cdots & (0) & A_k^1 \end{array} \quad \Rightarrow \quad e_{00..0} = e_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}^T$$

$$\begin{array}{cccc|c} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & A_k^0 \\ \hline \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \cdots & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & A_k^1 \end{array} \quad \Rightarrow \quad \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \end{pmatrix}^T$$

$$\begin{array}{ccc|c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & A_k^0 \\ \hline \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & A_k^1 \end{array} \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}^T$$

# Application to functions

Truncated ANOVA decomposition:

$$f(x_1, x_2, \ldots, x_d) \approx g_0 + \sum_{j_1=1}^{d} g_{1,j_1}(x_{j_1}) + \sum_{j_1,j_2=1}^{d} g_{2,j_1 j_2}(x_{j_1}, x_{j_2}) + \sum_{j_1,j_2,j_3=1}^{d} g_{3,j_1 j_2 j_3}(x_{j_1}, x_{j_2}, x_{j_3}) + \ldots$$

Find functions $g_k$ that allow a good approximation of f.

Tensor train approximation with matrices $G_k$:

$$f(x_1, x_2, \ldots, x_d) \approx g_1(x_1) G_2(x_2) \cdots G_{d-1}(x_{d-1}) g_d(x_d)$$

Similarly we can generalize CP, Tucker, H-Tucker to functions.

$$f(x_1, x_2, \ldots, x_d) \approx \sum_{j=1}^{n} g_{1,j}(x_1) g_{2,j}(x_2) \cdots g_{d-1,j}(x_{d-1}) g_{d,j}(x_d)$$

# Representing Matrices

$$M = A_{\{i_1 \dots i_d\}\{j_1 \dots j_d\}} = G_1^{i_1, j_1} G_2^{i_2, j_2} \cdots G_{d-1}^{i_{d-1}, j_{d-1}} G_d^{i_d, j_d}$$

$$M = \sum_{j_1 \dots j_{d-1}} U_{1, j_1} \otimes U_{2, j_1 j_2} \otimes \cdots \otimes U_{d-1, j_{d-2} j_{d-1}} \otimes U_{d, j_{d-1}}$$

Special case: Laplacian

# Generalizing Matrix Properties

Let A be a symmetric n-dimensional tensor
and $x^m$ a rank-one tensor for vector x ϵ *IR* :

$$A = a_{i_1 \ldots i_m}; \qquad x^m := x_{i_1} \cdots x_{i_m};$$

Define the n-dimensional homogeneous polynomial of degree m

$$f(x) := Ax^m := \sum_{i_1,\ldots,i_m=1}^{n} a_{i_1 \ldots i_m} x_{i_1} \cdots x_{i_m} \in IR$$

A symmetric: invariant under any index permutations.
A positive definite: f(x) > 0 for all x ≠ 0.

# Eigenvalue

Define by $Ax^{m-1} := \left( \sum_{i_2 \dots i_m = 1}^{n} a_{i,i_2,\dots,i_m} x_{i_2} \cdots x_{i_m} \right)_{i=1}^{n} \in IR^n$ a vector.

We call a number $\lambda \in \boldsymbol{C}$ an eigenvalue of A if $\lambda$ and $x \neq 0$ are solutions of the polynomial equation

$$\left(Ax^{m-1}\right)_i = \lambda x_i^{m-1} = \lambda\left(Ix^{m-1}\right)_i$$

$Ax^{m-1} = \lambda I x^{m-1}$, or the vector x is a fixed point of operator A.

Here I is the identity tensor: $a_{i,i,\dots,i}=1$, and $a=0$ otherwise.

# Matrix terms

A Hermitian:

- Invariant subspace: $Ax = \lambda x$
- Rayleigh quotient: $x^T Ax / x^T x$
- Lagrange multipliers: $x^T Ax - \lambda (||x||^2 - 1)$
- Best rank-1 approximation: $\min_{||x||=1} ||A - \lambda xx^T||$

A general:

- Pseudospectrum: $\sigma_\varepsilon(A) = \{\lambda \mid ||(A-\lambda I)^{-1}|| > \varepsilon^{-1} \}$
- Numerical range: $W(A) = \{x^* Ax \mid x^T x = 1\}$

How to generalize to tensors?

# Numerical Range

Rayleigh Q. $A(x_1,\ldots,x_d) = A\cdot\left(x_1,\ldots,x_d\right) = \sum_{i_1,\ldots,i_d} A_{i_1,\ldots,i_d} x_{1,i_1}\cdots x_{d,i_d}$

Range: $W(A) = A(x,\ldots,x) = A\cdot\left(x,\ldots,x\right) = \sum_{i_1,\ldots,i_d} A_{i_1,\ldots,i_d} x_{i_1}\cdots x_{i_d}$

CP or Tucker as generalization of SVD:
-CP loosing the orthogonality!
- Tucker loosing the diagonal form of the core tensor!

# Eigenvalues

As critical points of the Rayleigh quotient

$$\frac{A(x,\ldots,x)}{\|x\|_d^d} = \frac{A(x,\ldots,x)}{I(x,\ldots,x)}$$

By Lagrangian $L(x,\lambda) := A(x,\ldots,x) - \lambda\left(\|x\|_d^d - 1\right)$

Characteristic Polynomial p(λ) is the resultant of the two polynomials $Ax^{m-1}$ and $\lambda x^{m-1}$ (searching for common zeros).

Eigenvalues are roots of p. #eigenvalues: $n(m-1)^{n-1}$
Product of eigenvalues = det(A) = resultant of $Ax^{m-1}$ and 0
Sum of all eigenvalues is equal $(m-1)^{n-1}$trace(A)

Liqun Qi, Lek-Heng Lim

# Software

Kolda: Data structures, CP, Tucker
Oseledets: TT
Kressner, Tobler: H-Tucker
ALPS: Quantum simulation

http://www.sandia.gov/~tgkolda/TensorToolbox/index-2.5.html
http://spring.inm.ras.ru/osel/?page_id=24
http://www.sam.math.ethz.ch/NLAgroup/htucker_toolbox.html
https://www.rdb.ethz.ch/projects/project.php?proj_id=8486