Tutorial: Tensor Approximation in Visualization and Computer Graphics

Tensor Decomposition Models

Renato Pajarola, Susanne K. Suter, and Roland Ruiters

University of Zurich UZH

Visualization and Multimedia Lab

universität bonn

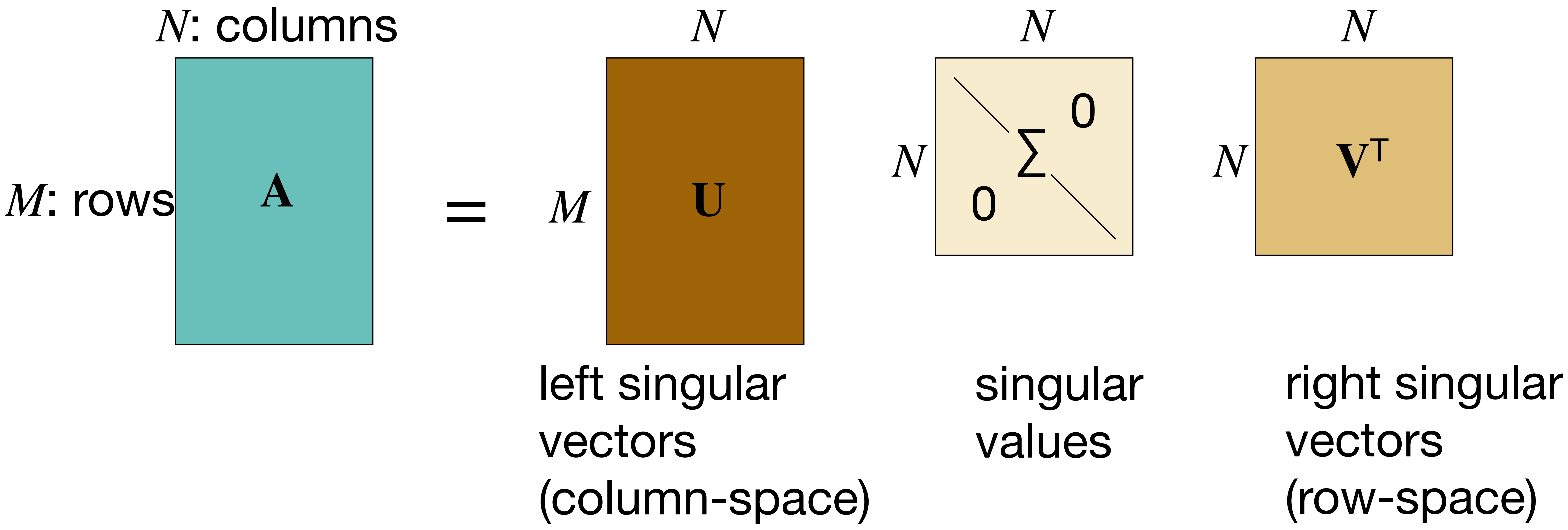
Institute of Computer Science II Computer Graphics

Data Reduction and Approximation

- A fundamental concept of data reduction is to remove redundant and irrelevant information while preserving the relevant features
  - e.g. through frequency analysis by projection onto pre-defined bases, or extraction of data intrinsic principal components
    - identify spatio-temporal and frequency redundancies
  - maintain strongest and most significant signal components
- Data reduction linked to concepts and techniques of data compression, noise reduction as well as feature extraction and recognition/extraction

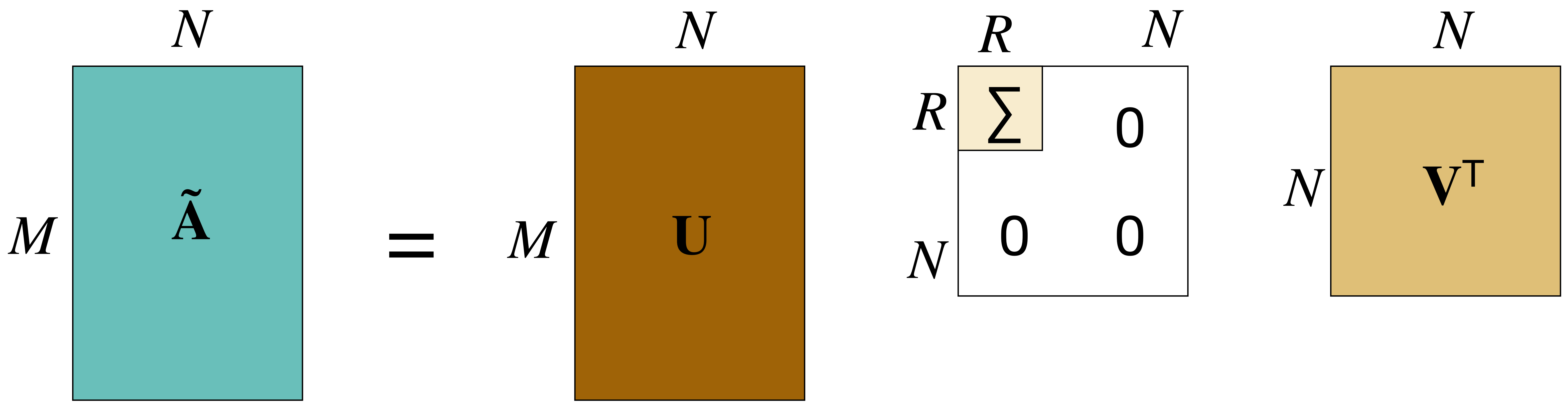
Data Approximation using SVD

- Singular Value Decomposition (SVD) standard tool for matrices, i.e., 2D input datasets
  - see also principal component analysis (PCA)



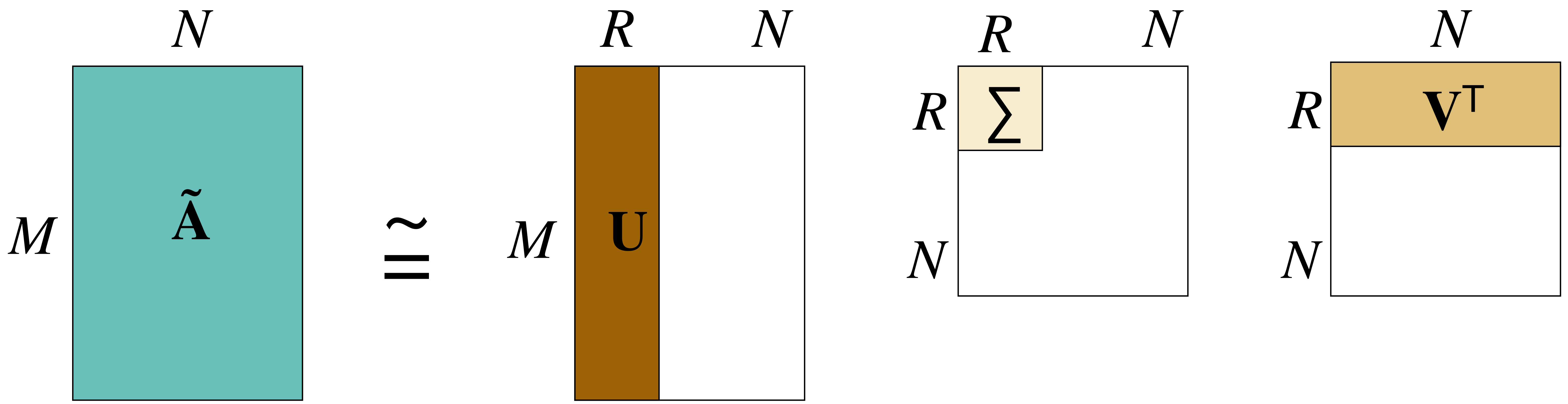
Low-rank Approximation

- Exploit ordered singular values: $s_1 \geq s_2 \geq \ldots \geq s_N$
- Select first r singular values (rank reduction)



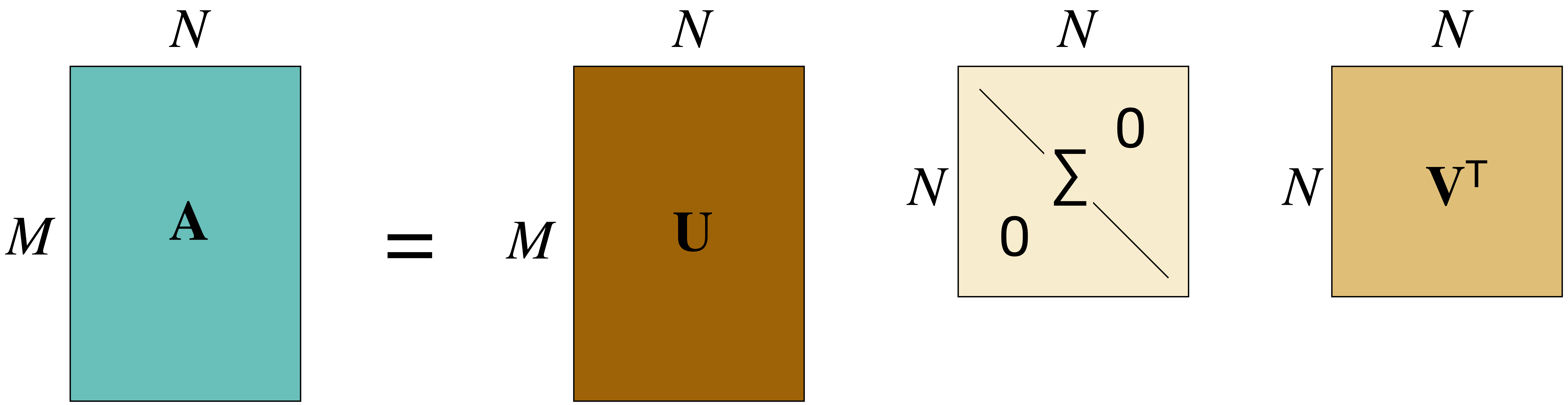
Low-rank Approximation

- Exploit ordered singular values: $s_1 \geq s_2 \geq \ldots \geq s_N$
- Select first r singular values (rank reduction)
  - use only bases (singular vectors) of corresponding subspace

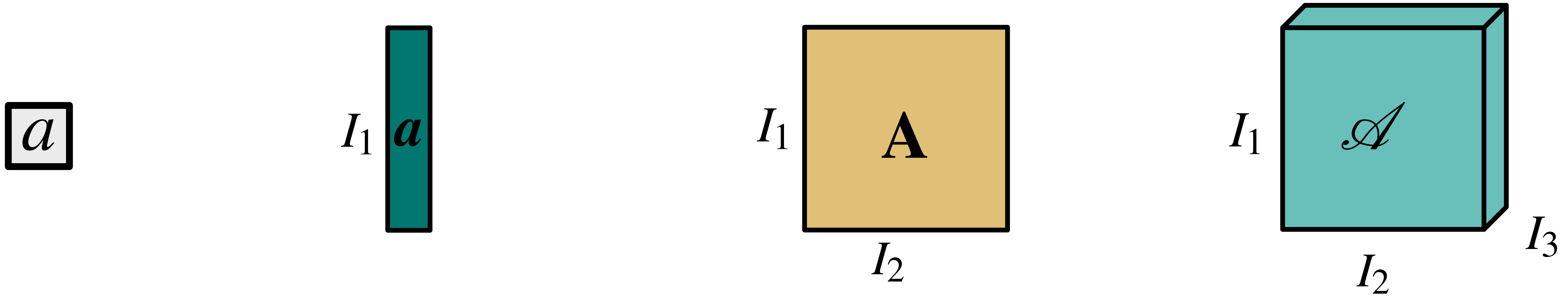


Matrix SVD Properties

- Matrix SVD
  - rank reducibility
  - orthonormal row/column matrices



What is a Tensor?



$$\mathscr{A}^{N-\text{th order}} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$$

$i_1 = 1, \ldots, I_1$ | $i_2 = 1, \ldots, I_2$ | $i_3 = 1, \ldots, I_3$

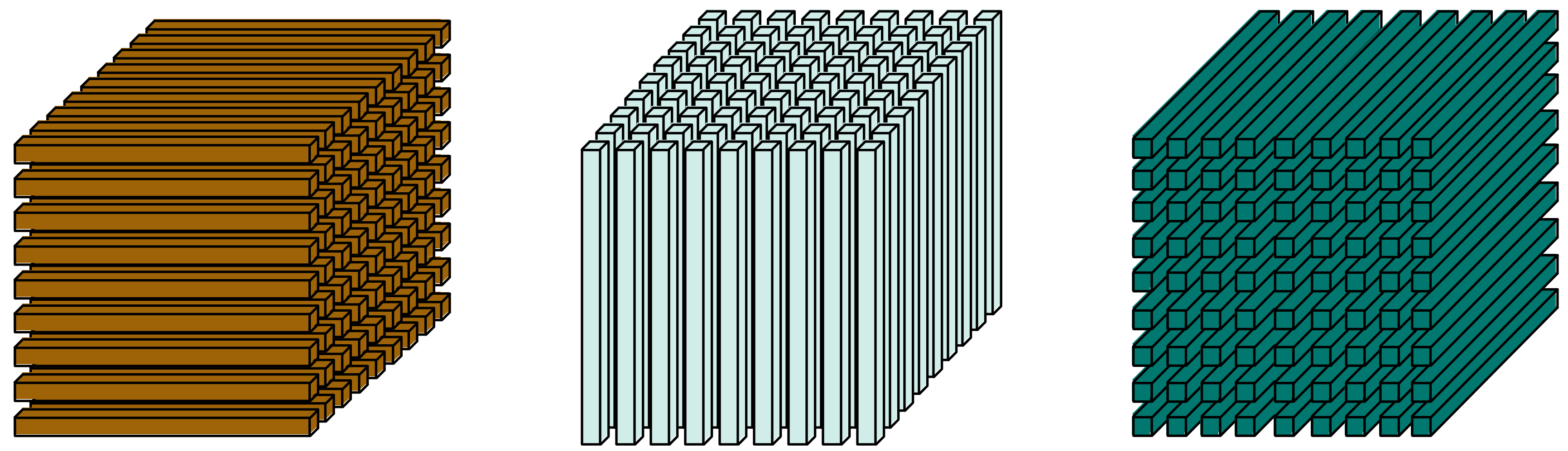
0-order tensor | 1st-order tensor | 2nd-order tensor | 3rd-order tensor | ...

- Data sets are often multidimensional arrays (tensors)
  - images, image collections, video, volume data etc.

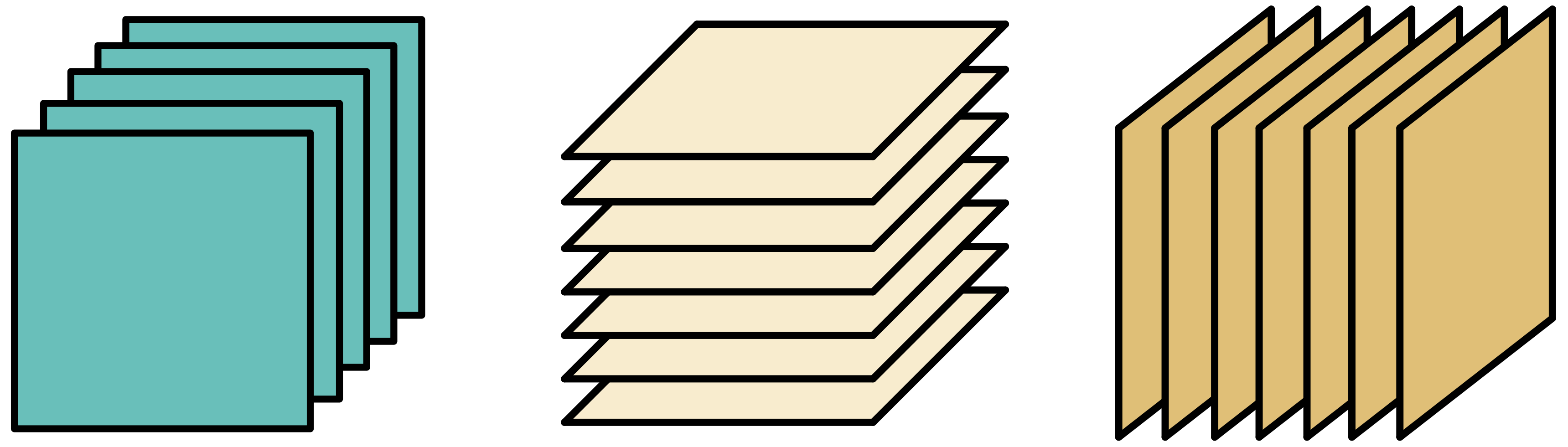
Fibers and Slices

- Individual elements of a vector $\boldsymbol{a}$ are given by a_{i1}, from a matrix $\mathbf{A}$ by $a_{i1,i2}$ and from a tensor $\mathscr{A}$ by $a_{i1,i2,i3}$

- The generalization of rows, columns (and tubes) is a *fiber* in a particular mode

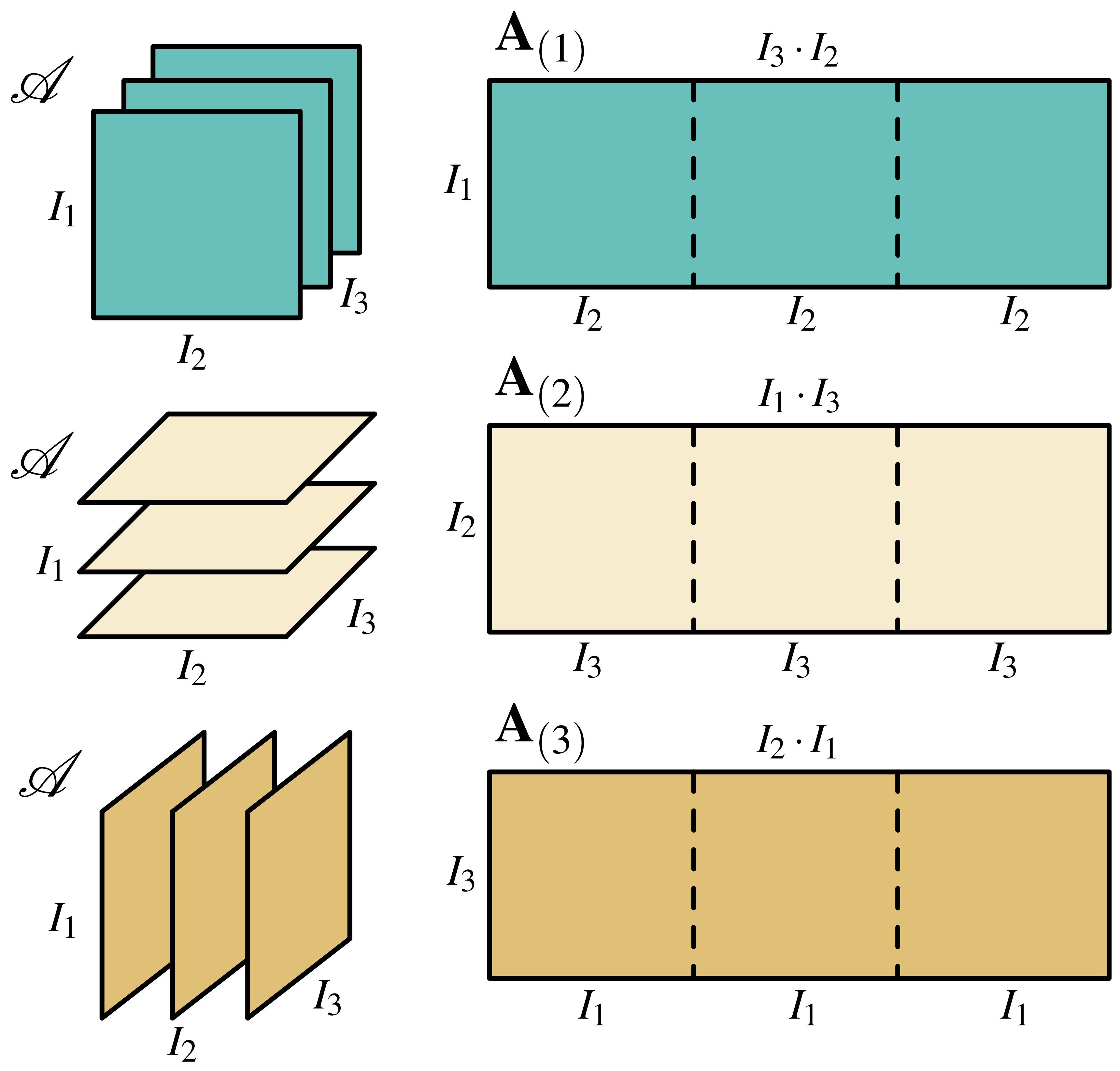


- Two dimensional sections of a tensor are called slices
  - frontal, horizontal and lateral for $\mathscr{A} \in \mathbb{R}^3$



Unfolding and Ranks

- Operations with tensors often performed as matrix operations using unfolded tensor representations
  - different tensor unfolding strategies possible
- Forward cyclic unfolding $\mathbf{A}_{(n)}$ of a 3rd order tensor $\mathscr{A}$ (or 3D volume)
- The n-rank of a tensor is typically defined on an unfolding
  - n-rank $R_n = \mathrm{rank}_n(\mathscr{A}) \; = \mathrm{rank}(\mathbf{A}_{(n)})$
  - multilinear rank-$(R_1, R_2, \ldots, R_N)$ of $\mathscr{A}$

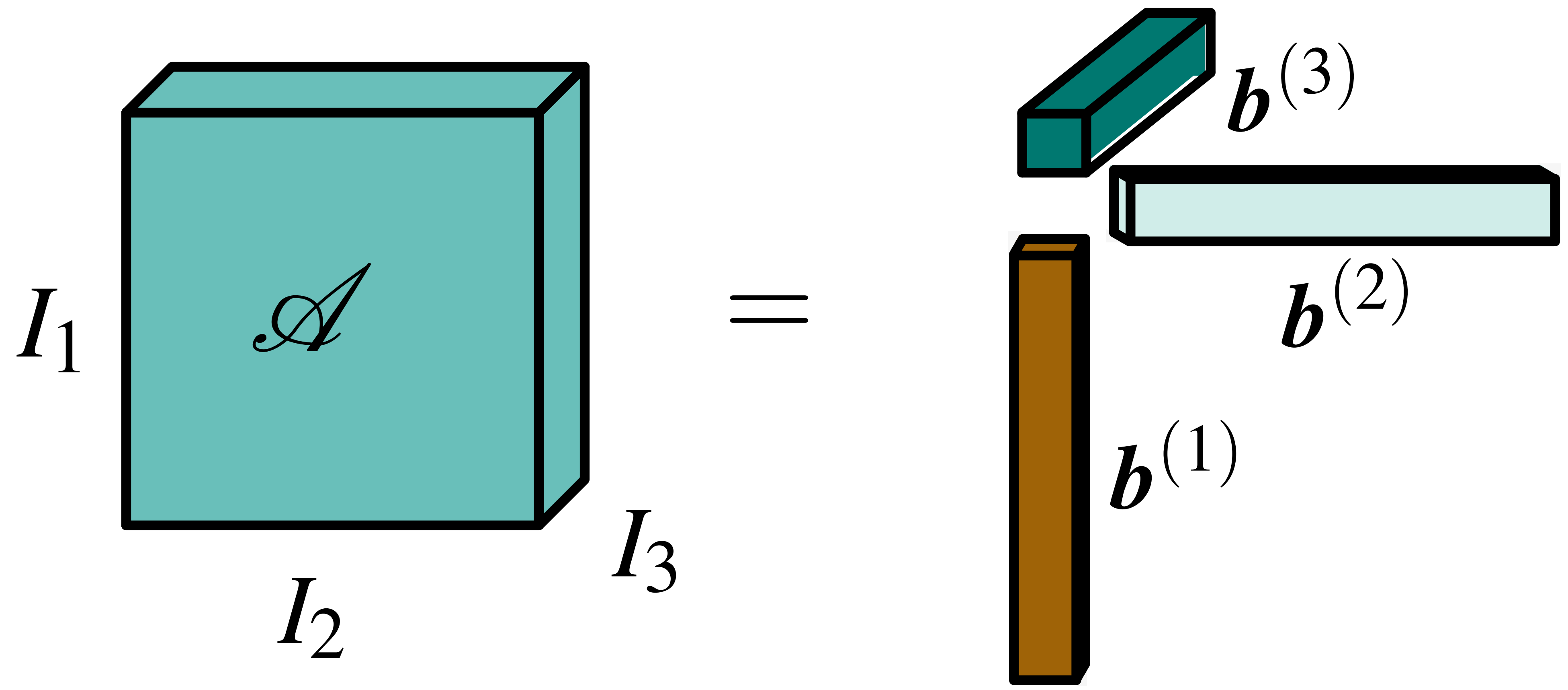


Rank-one Tensor

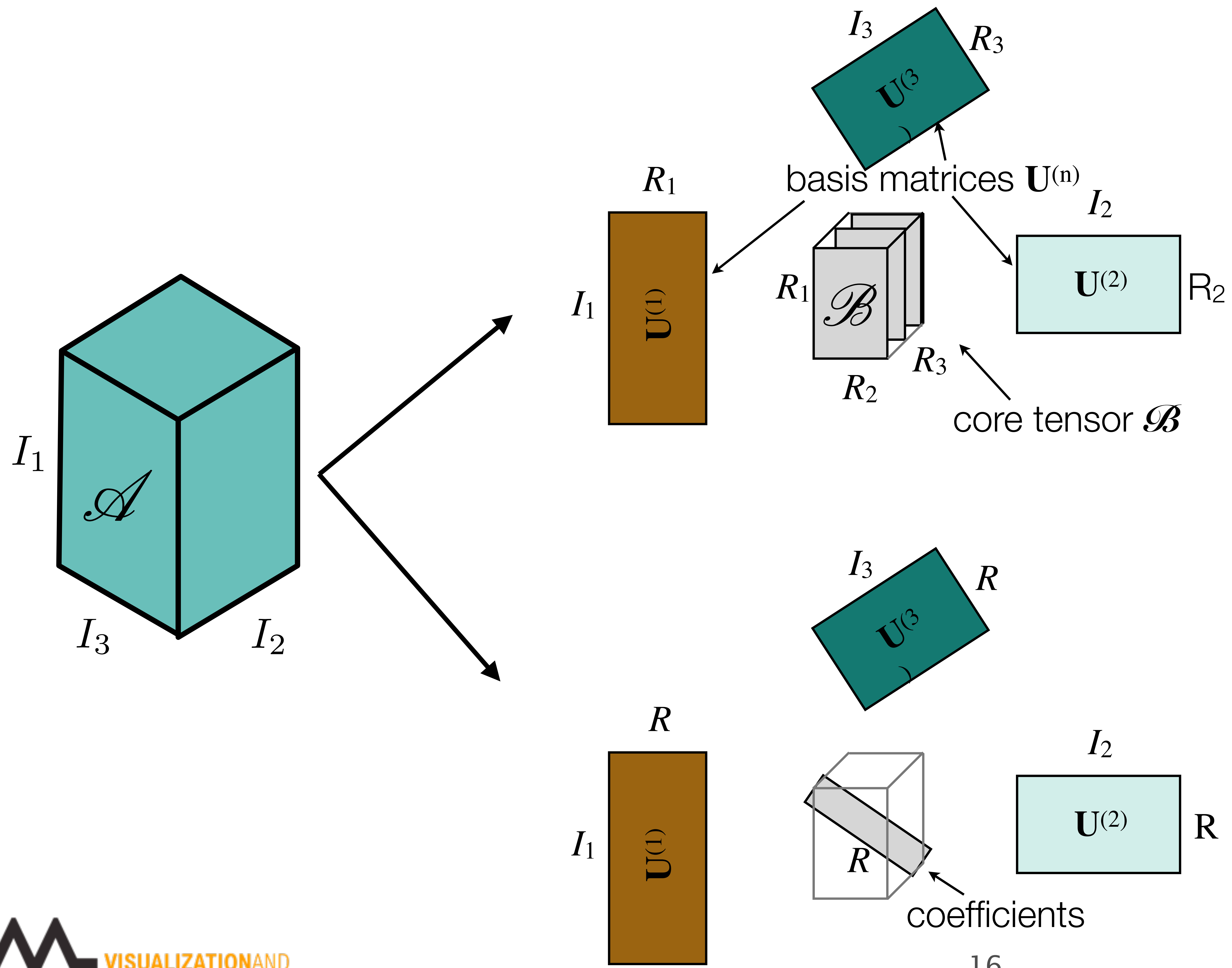
- N-mode tensor $\mathscr{A} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ that can be expressed as the outer product of N vectors
  - Kruskal tensor

$$\mathscr{A} = \boldsymbol{b}^{(1)} \circ \boldsymbol{b}^{(2)} \circ \ldots \circ \boldsymbol{b}^{(N)}$$

- Useful to understand principles of rank-reduced tensor reconstruction
  - linear combination of rank-one tensors



Tensor Decomposition Models



Tucker

- Three-mode factor analysis (**3MFA/Tucker3**) [Tucker, 1964+1966]
- Higher-order SVD (**HOSVD**) [De Lathauwer et al., 2000a]

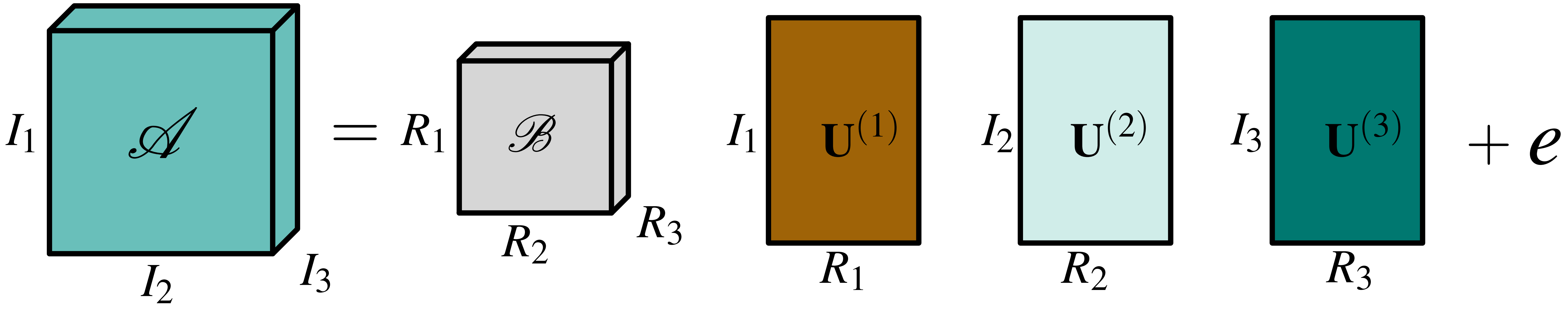
CP

- **PARAFAC** (parallel factors) [Harshman, 1970]
- **CANDECOMP** (CAND) (canonical decomposition) [Caroll & Chang, 1970]

Tucker Model

- Higher order tensor $\mathscr{A} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ represented as a product of a core tensor $\mathscr{B} \in \mathbb{R}^{R_1 \times \ldots \times R_N}$ and N factor matrices $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}$
  - using n-mode products $\times_n$

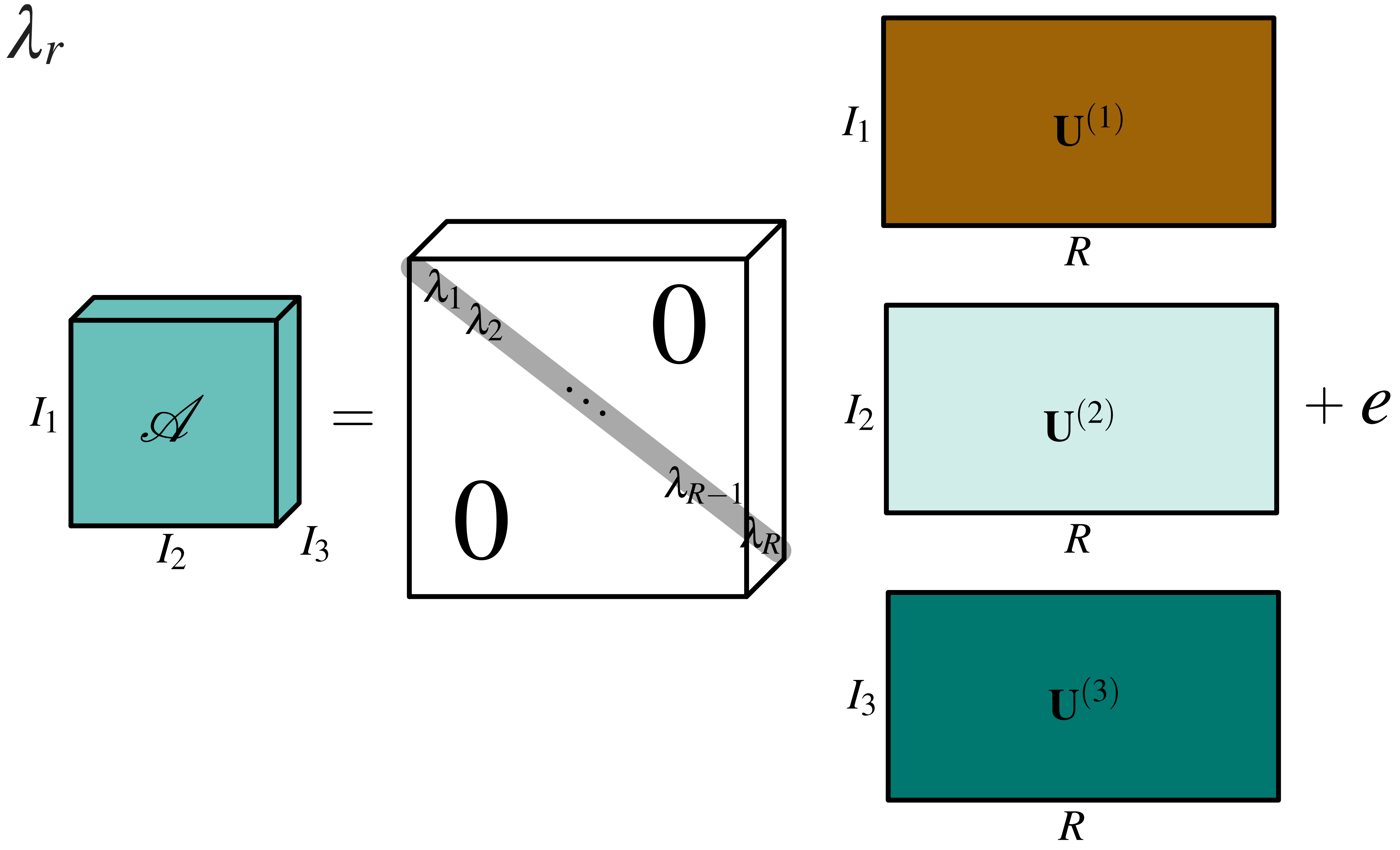
$$\mathscr{A} = \mathscr{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \cdots \times_N \mathbf{U}^{(N)} + \varepsilon$$



CANDECOMP-PARAFAC Model

- *Canonical decomposition* or *parallel factor analysis* model (CP)
- Higher order tensor $\mathscr{A}$ factorized into a sum of rank-one tensors
  - normalized column vectors $\boldsymbol{u}_r^{(n)}$ define factor matrices $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R}$ and weighting factors λ_r

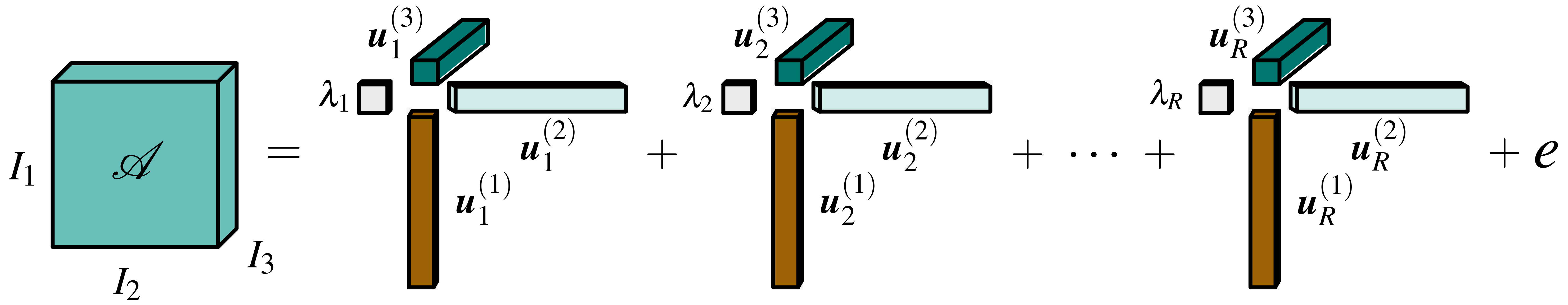
$$\mathscr{A} = \sum_{r=1}^{R} \lambda_r \cdot \boldsymbol{u}_r^{(1)} \circ \boldsymbol{u}_r^{(2)} \circ \ldots \boldsymbol{u}_r^{(N)} + \varepsilon$$



Linear Combination of Rank-one Tensors

- The CP model is defined as a linear combination of rank-one tensors

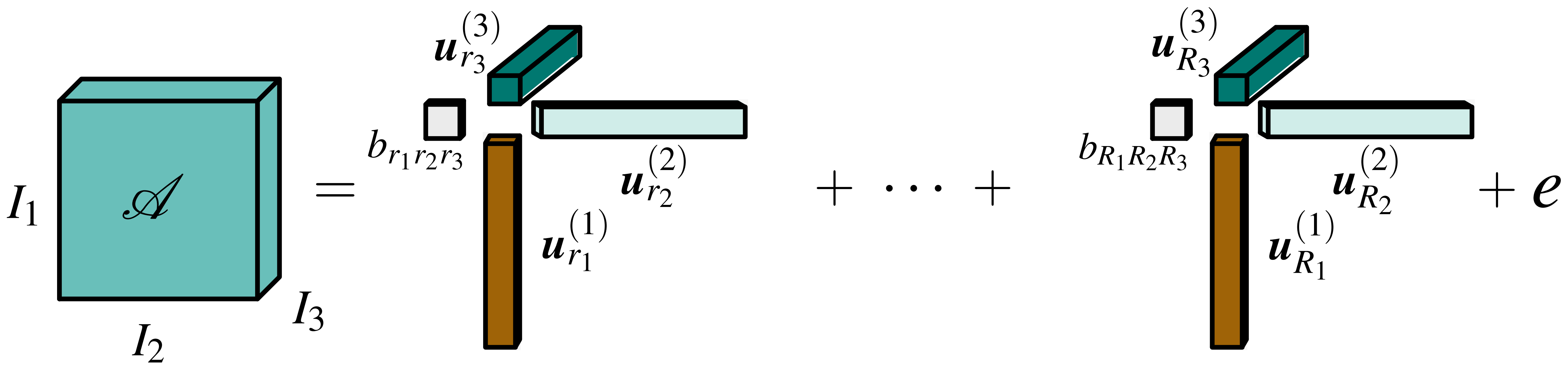
$$\mathscr{A} = \sum_{r=1}^{R} \lambda_r \cdot \boldsymbol{u}_r^{(1)} \circ \boldsymbol{u}_r^{(2)} \circ \ldots \boldsymbol{u}_r^{(N)} + \varepsilon$$



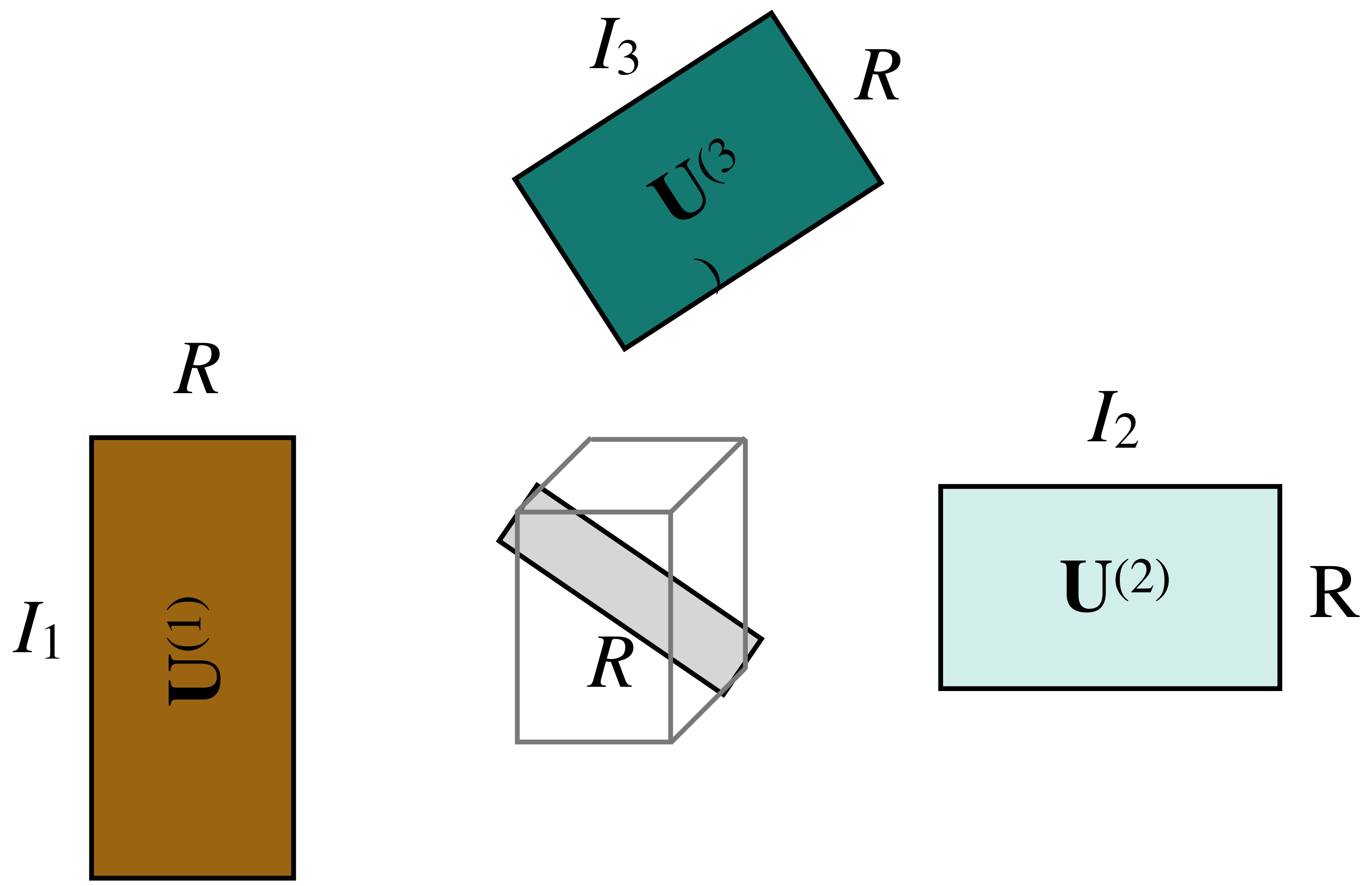
Linear Combination of Rank-one Tensors

- The CP model is defined as a linear combination of rank-one tensors
- The Tucker model can be interpreted as linear combination of rank-one tensors

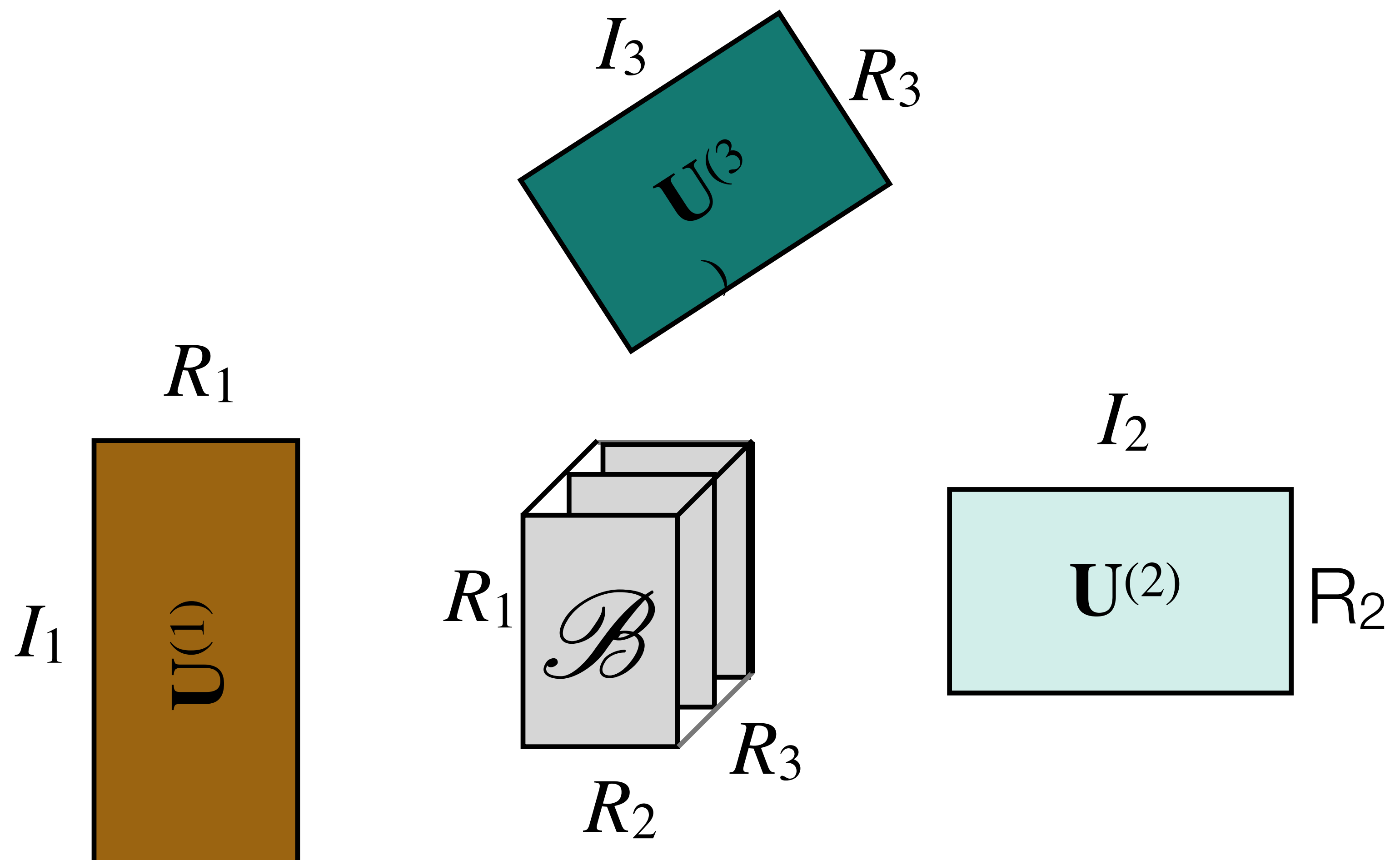
$$\mathscr{A} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} b_{r_1 r_2 \ldots r_N} \cdot \boldsymbol{u}_{r_1}^{(1)} \circ \boldsymbol{u}_{r_2}^{(2)} \circ \ldots \boldsymbol{u}_{r_N}^{(N)} + \varepsilon$$



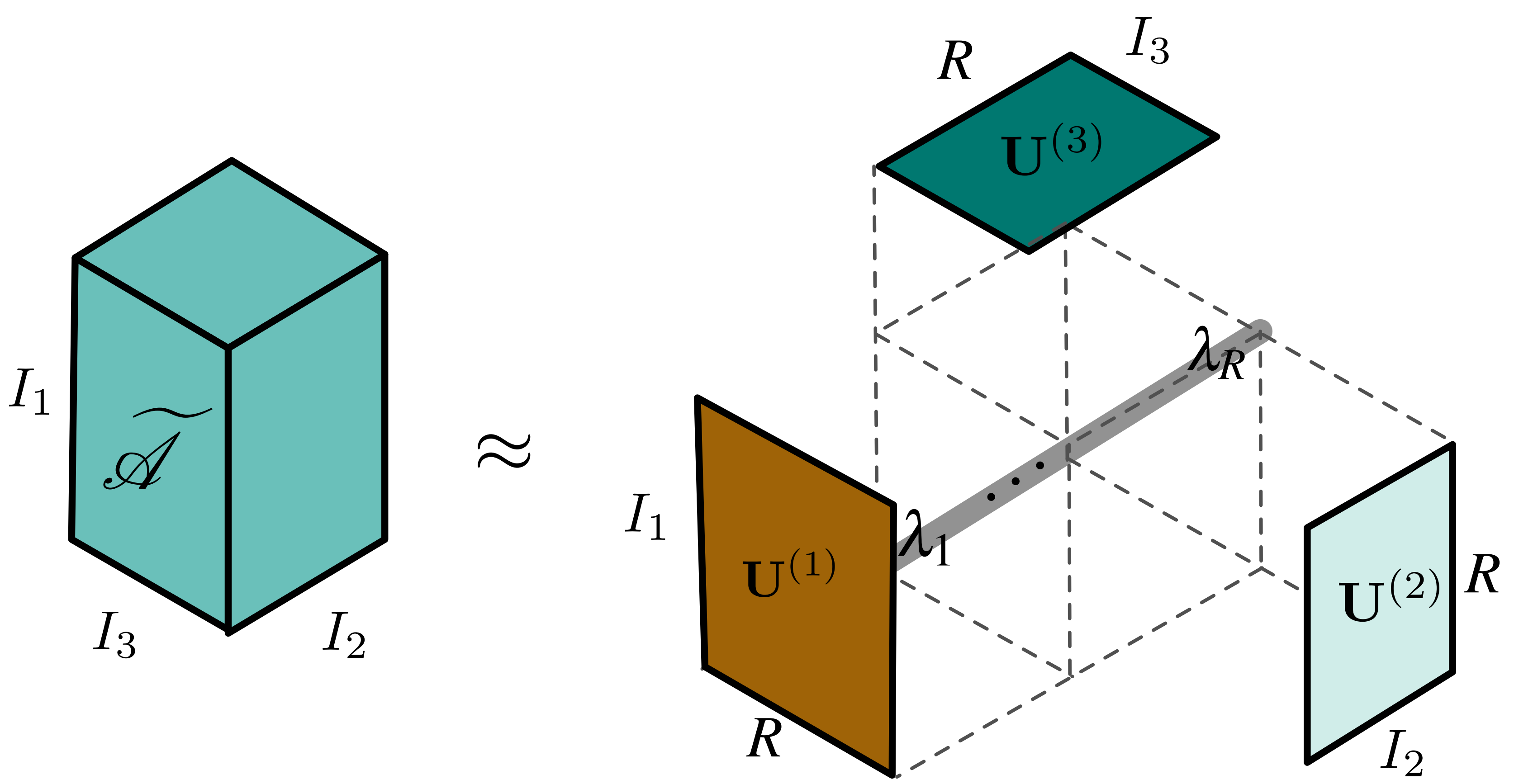
CP a Special Case of Tucker

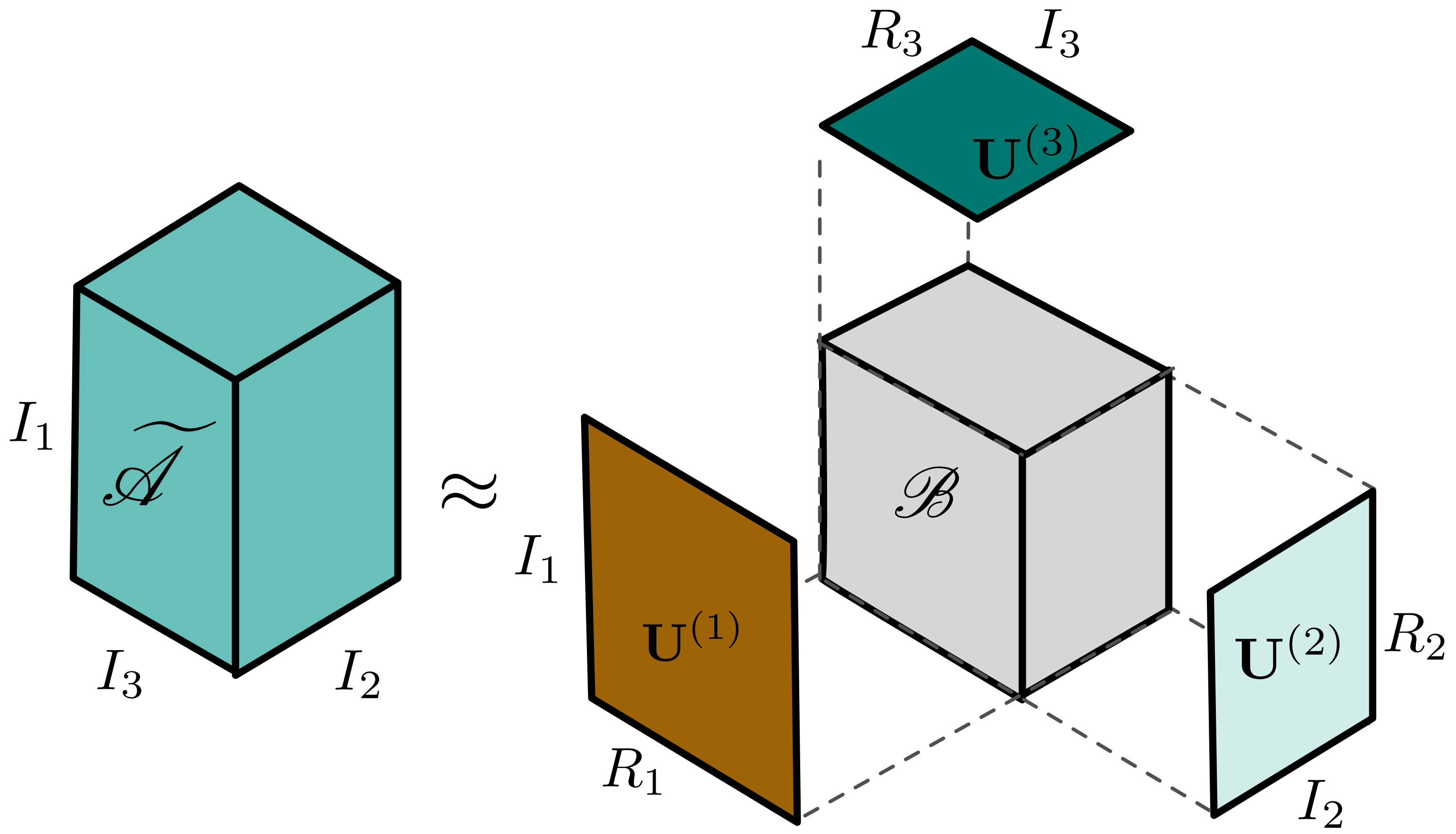


CP



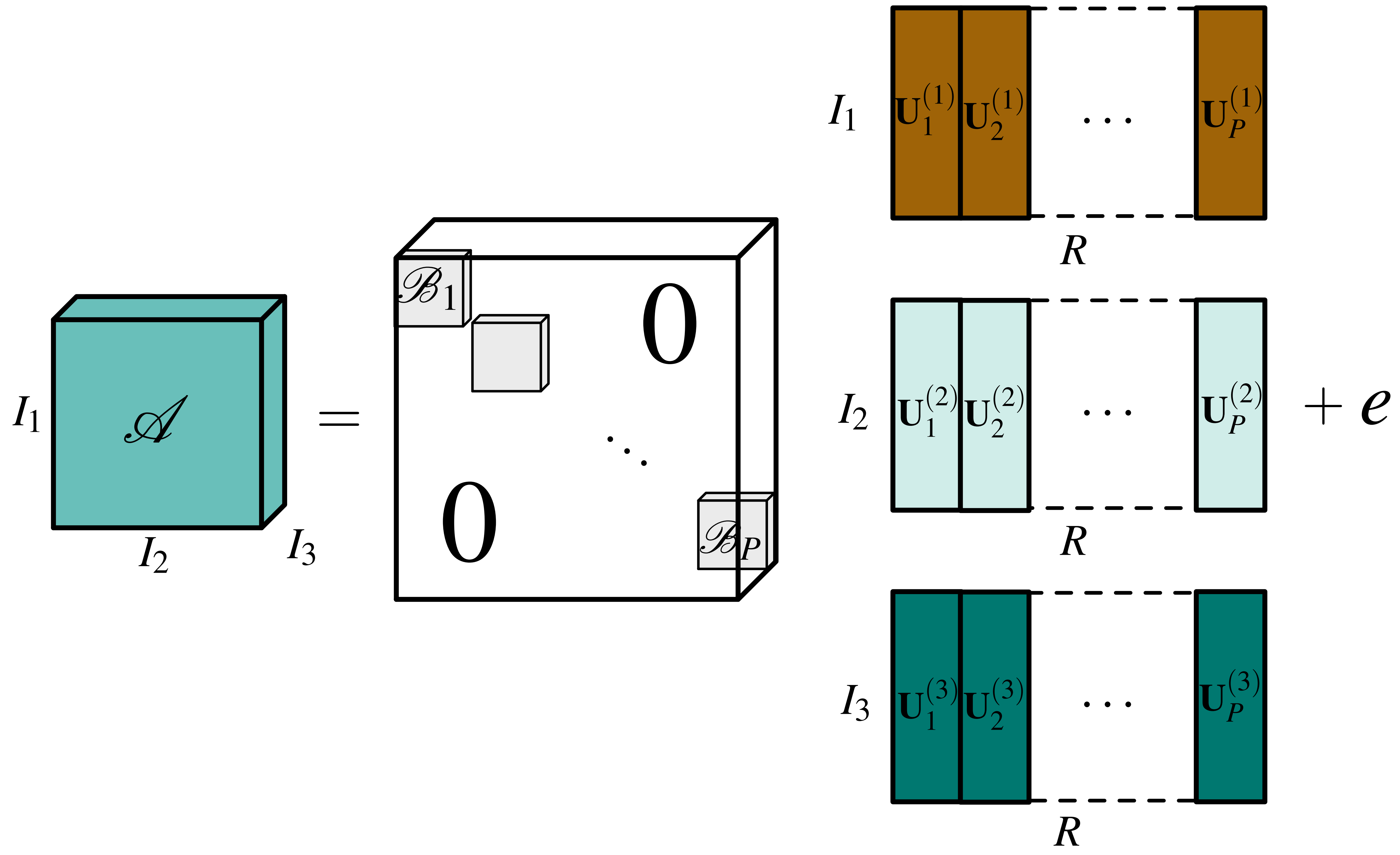
Tucker





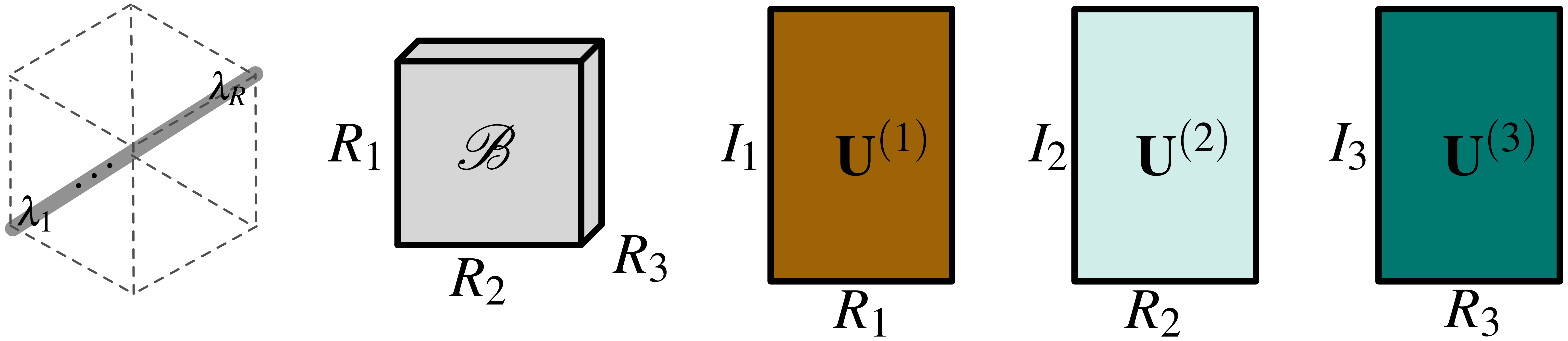
Generalizations

- Any special form of core and corresponding factor matrices
  - e.g. blocks along diagonal



Reduced Rank Approximation

- Full reconstruction using a Tucker or CP model may require excessively many coefficients and wide factor matrices
  - large rank values R (CP), or $R_1, R_2 \ldots R_N$ (Tucker)
- Quality of approximation increases with the rank, and number of column vectors of the factor matrices
  - best possible fit of these bases matrices discussed later



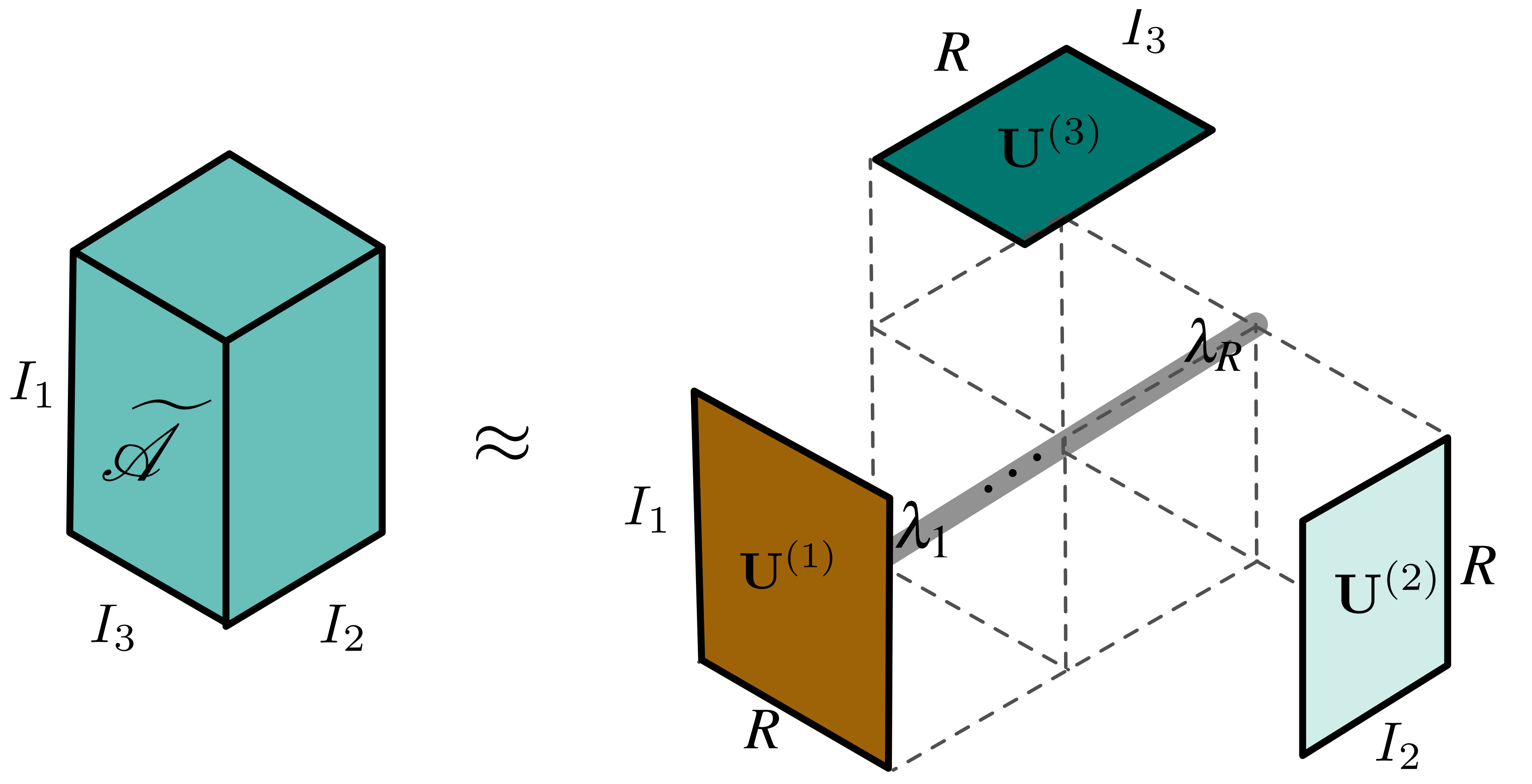
VISUALIZATION AND MULTIMEDIA LAB

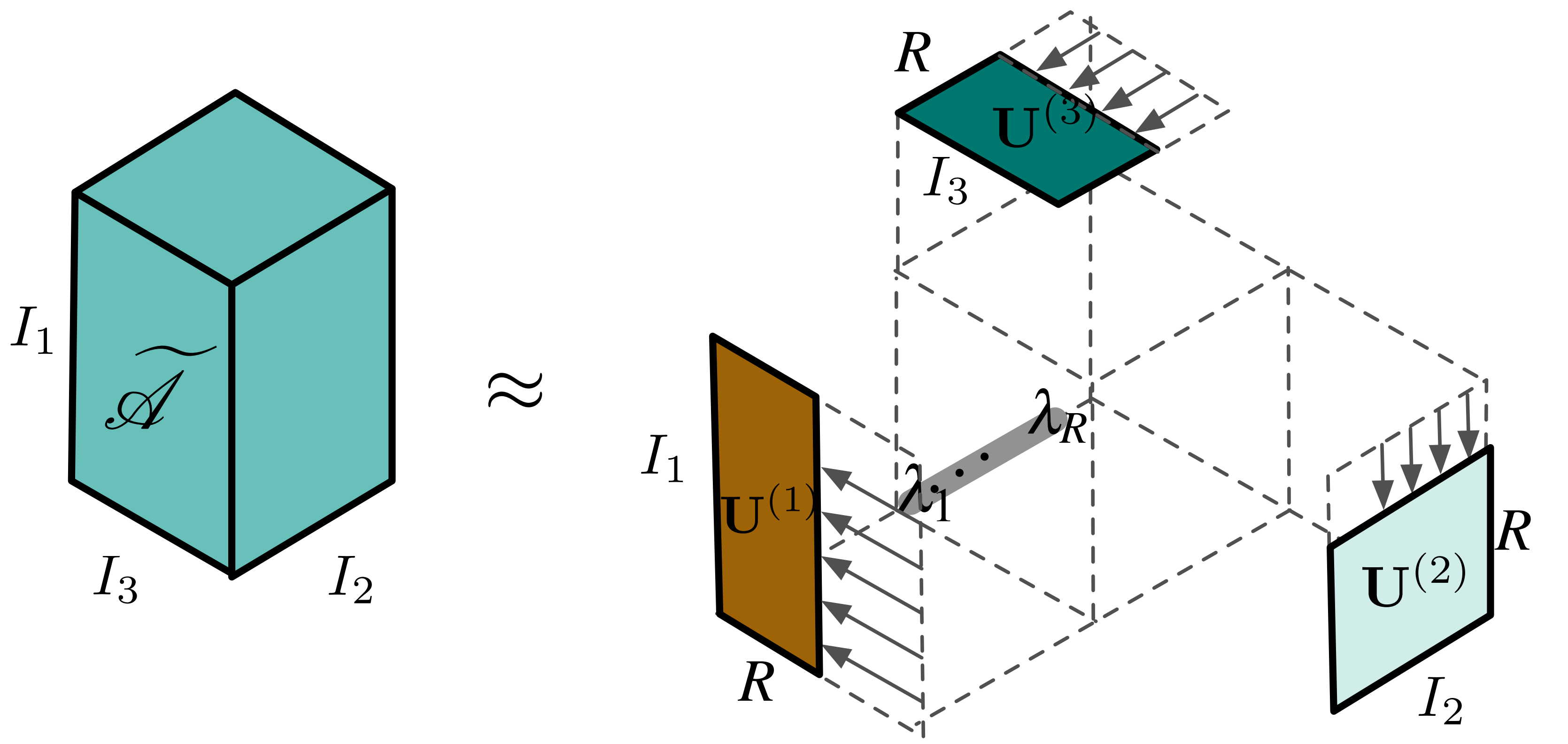
universität bonn

Rank-R Approximation

- Approximation of a tensor as a linear combination of ranke-one tensors using a limited number R of terms
  - CP model of limited rank R

$$\widetilde{\mathscr{A}} = \sum_{r=1}^{R} \lambda_r \cdot \boldsymbol{u}_r^{(1)} \circ \boldsymbol{u}_r^{(2)} \circ \ldots \boldsymbol{u}_r^{(N)}$$

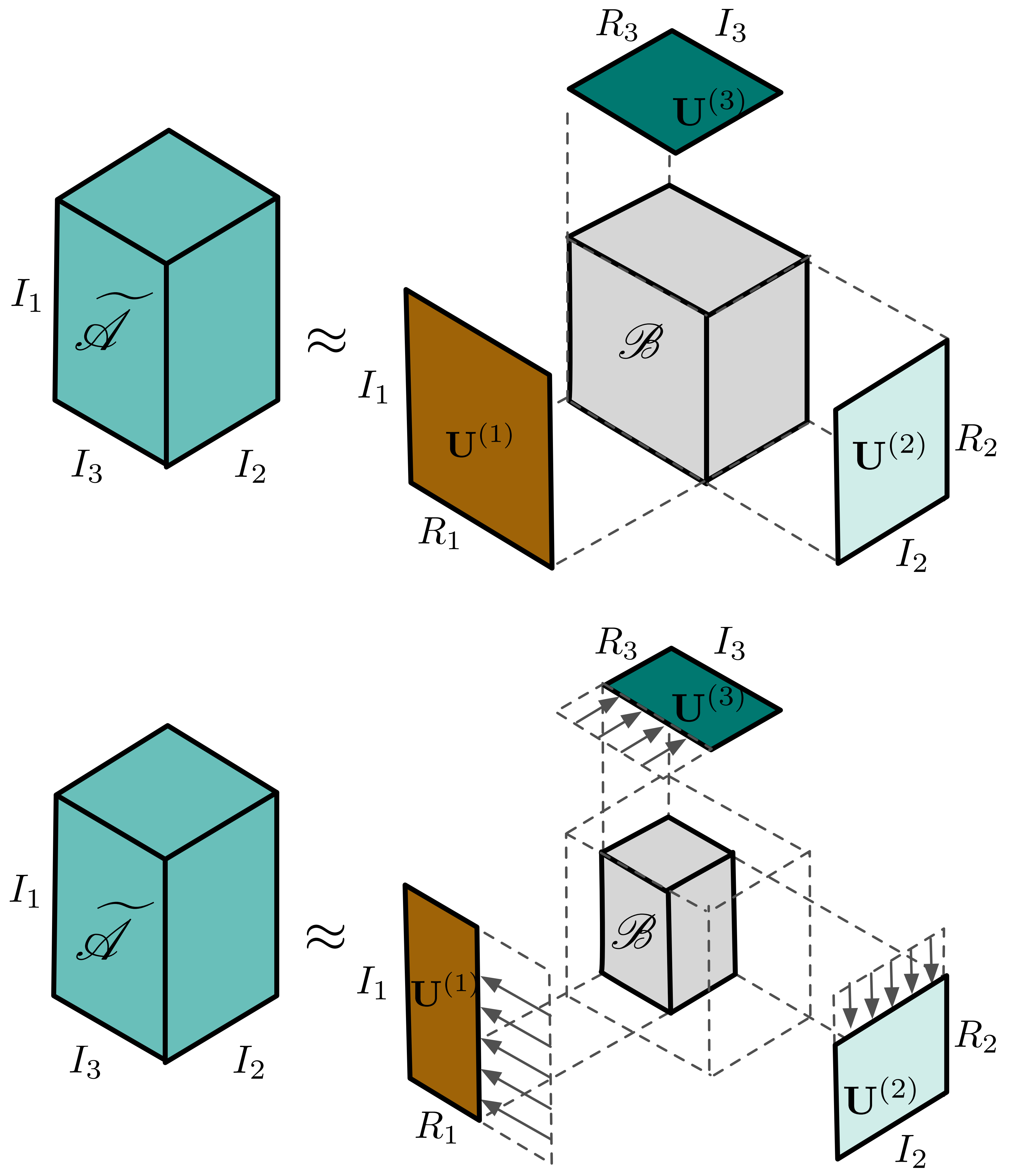




Rank-$(R_1, R_2, \ldots, R_N)$ Approximation

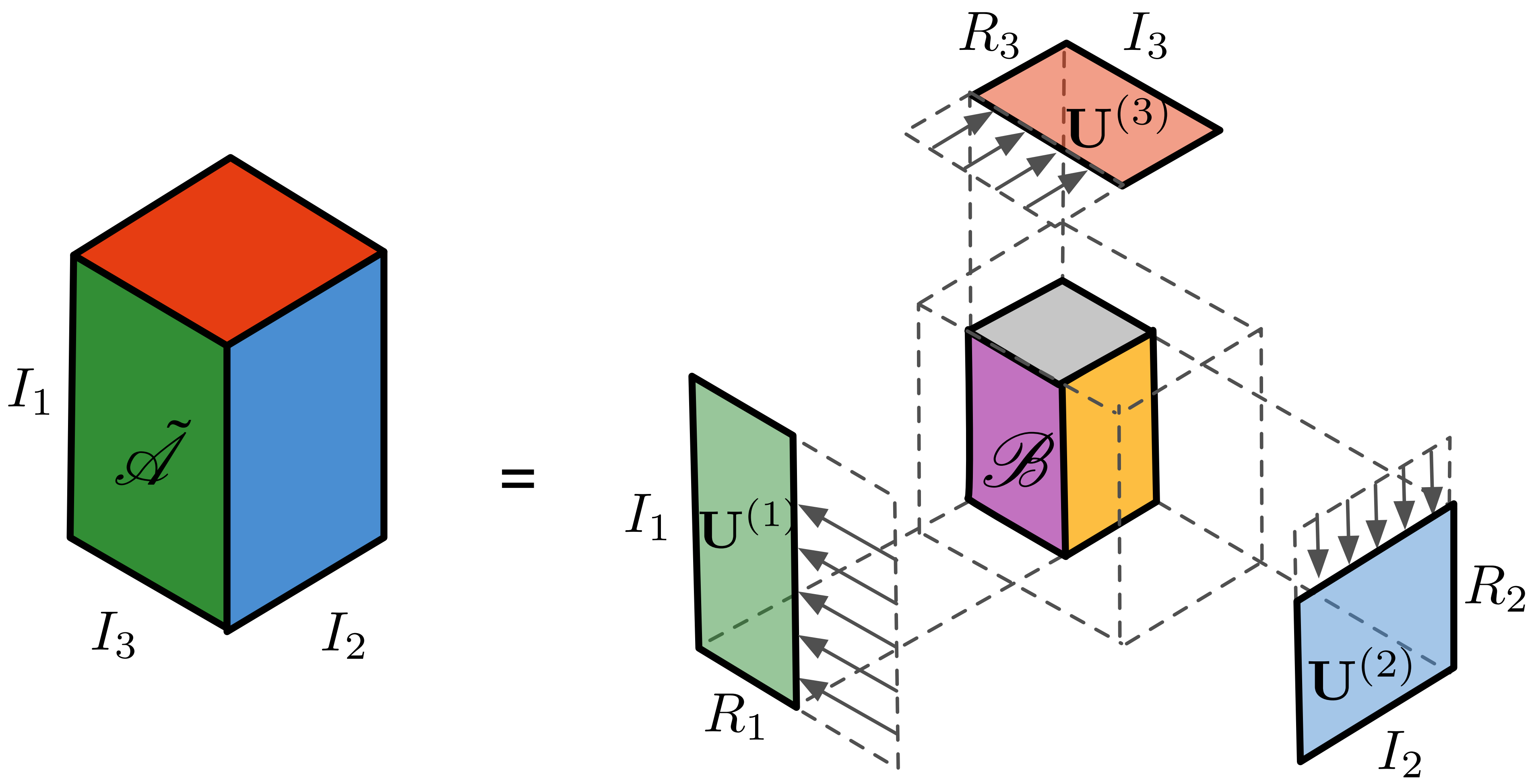
- Decomposition into a tensor with reduced, lower multilinear $\operatorname{rank}(R_1, R_2, \ldots, R_N)$
  - $\operatorname{rank}_n(\widetilde{\mathscr{A}}) = R_n \leq \operatorname{rank}_n(\mathscr{A}) = \operatorname{rank}(\mathbf{A}_{(n)})$
- ***n***-mode products of factor matrices and core tensor in a given reduced rank space
  - Tucker model with limited ranks R_i

$$\widetilde{\mathscr{A}} = \mathscr{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \cdots \times_N \mathbf{U}^{(N)}$$



Best Rank Approximation

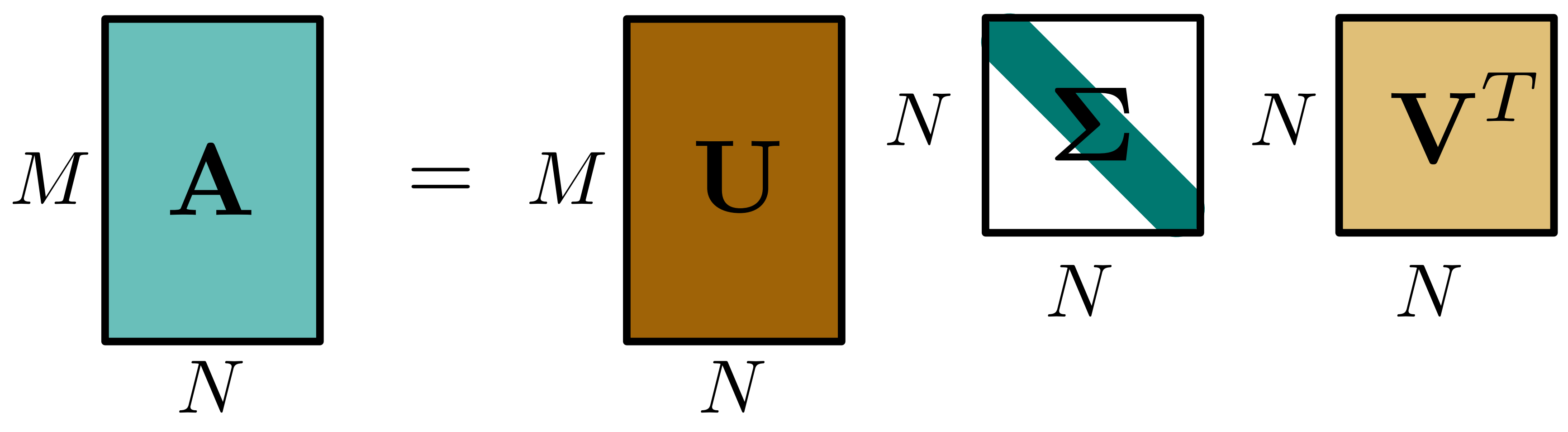
- Rank reduced approximation that minimizes least-squares cost
  - $\widetilde{\mathscr{A}} = \arg\min(\widetilde{\mathscr{A}}) \left\| \mathscr{A} - \widetilde{\mathscr{A}} \right\|^2$
- Alternating least squares (ALS) iterative algorithm that converges to a minimum approximation error based on the Frobenius norm $\|\ldots\|_F$
  - rotation of components in basis matrices



$$\widetilde{\mathscr{A}} = \mathscr{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$$

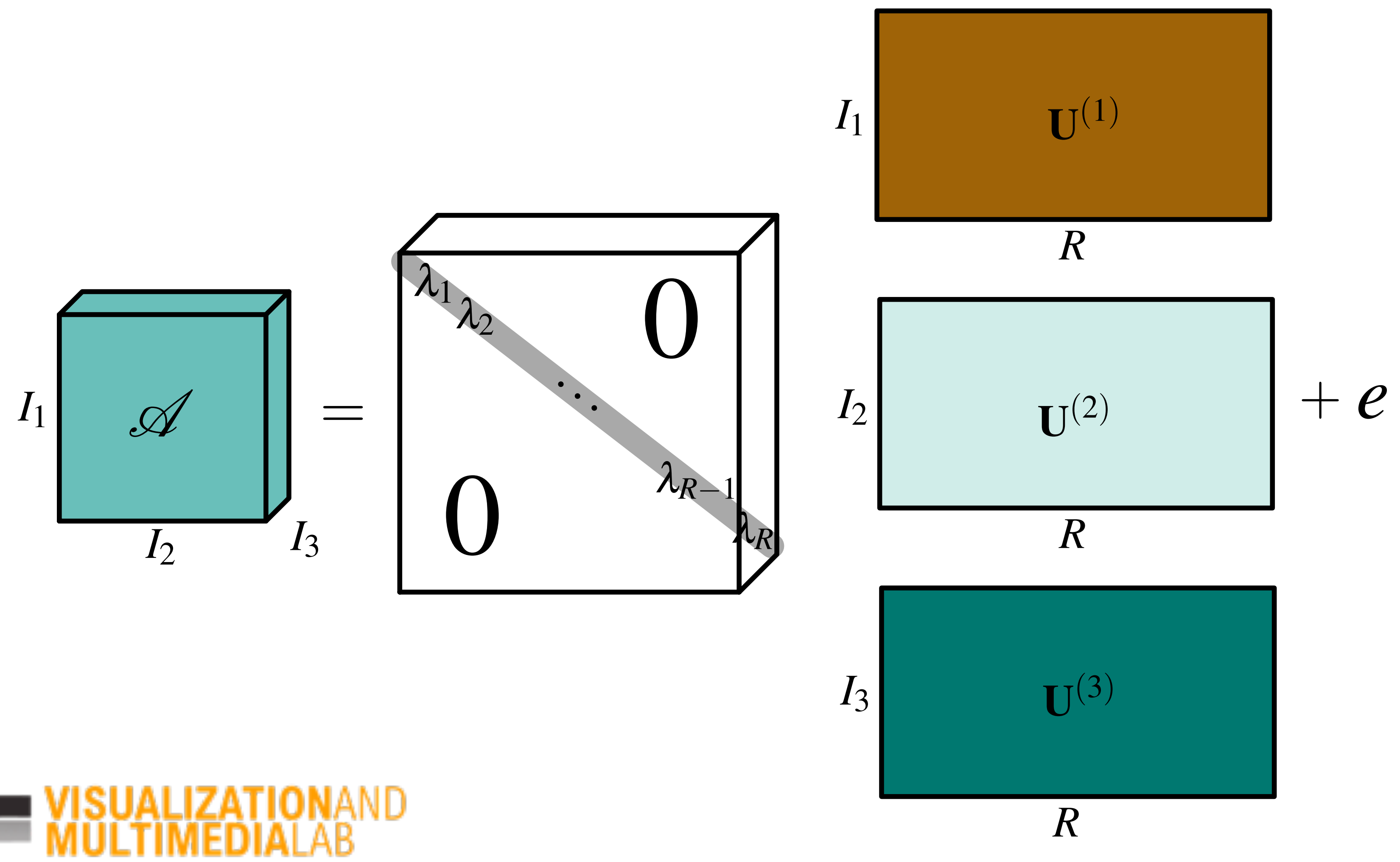
typical high-quality data reduction:
$R_k \leq I_k / 2$

Generalization of the Matrix SVD



higher orders

CP model



Tucker model