Outline

- Motivation
- Background and notation
- New t-product and associated algebraic framework
- Implementation considerations
- The t-SVD and optimality
  - Application in Facial Recognition
- Proper Orthogonal Decomposition
  - Dynamic Model Reduction
- A tensor Nuclear Norm from the t-SVD
  - Applications in video completion
Tensor Applications:

- **Machine vision**: understanding the world in 3D, enable understanding phenomena such as perspective, occlusions, illumination

- **Latent semantic tensor indexing**: common terms vs. entries vs. parts, co-occurrence of terms

Tensor subspace Analysis for Viewpoint Recognition, T. Ivanov, L. Mathies, M.A.O. Vasilescu, ICCV, 2nd IEEE International Workshop on Subspace Methods, September, 2009
Tensor Applications:

- **Medical imaging**: naturally involves 3D (spatio) and 4D (spatio-temporal) correlations

- **Video surveillance and Motion signature**: 2D images + 3rd dimension of time, 3D/4D motion trajectory

Multi-target Tracking with Motion Context in Tenor Power Iteration X. Shi, H. Ling, W. Hu, C. Yuan, and J. Xing IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Columbus OH, 2014
1927 F.L. Hitchcock: “The expression of a tensor or a polyadic as a sum of products” (Journal of Mathematics and Physics)

1944 R.B. Cattell introduced a multiway model: “Parallel proportional profiles and other principles for determining the choice of factors by rotation” (Psychometrika)

1960 L.R. Tucker: “Some mathematical notes on three-mode factor analysis” (Psychometrika)

1981 tensor decomposition was first used in chemometrics

Past decade, computer vision, image processing, data mining, graph analysis, etc.
The Power of Proper Representation

- What is that?

Representation matters! Some correlations can only be realized in appropriate representation.
The Power of Proper Representation

- What is that?

- Let’s observe the same data but in a different (matrix rather than vector) representation
The Power of Proper Representation

- What is that?

- Let’s observe the same data but in a different (matrix rather than vector) representation

- Representation matters! Some correlations can only be realized in appropriate representation.
Motivation

- Much real-world data is inherently multidimensional
  - color video data – 4 way
  - 3D medical image, evolving in time (4 way); multiple patients (5 way)
- Many operators and models are also multi-way
- Traditional matrix-based methods based on data vectorization (e.g. matrix PCA) generally agnostic to possible high dimensional correlations

Can we uncover hidden patterns in tensor data by computing an appropriate tensor decomposition/approximation?

Need to decide on the tensor decomposition – application dependent!

What do we mean by ‘decompose’?
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor
- **Examples**

0th order tensor - scalar
1st order tensor - vector
2nd order tensor - matrix
3rd order tensor ...
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor
- **Examples**
  - $0^{th}$ order tensor
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor
- **Examples**
  - $0^{th}$ order tensor - scalar
**Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor

**Examples**
- $0^{th}$ order tensor - scalar
- $1^{st}$ order tensor
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor

- **Examples**
  - $0^{th}$ order tensor - scalar
  - $1^{st}$ order tensor - vector
**Notation**: \( A^{n_1 \times n_2 \ldots \times n_j} \) - \( j^{th} \) order tensor

**Examples**

- \( 0^{th} \) order tensor - **scalar**
- \( 1^{st} \) order tensor - **vector**
- \( 2^{nd} \) order tensor
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor

- **Examples**
  - $0^{th}$ order tensor - **scalar**
  - $1^{st}$ order tensor - **vector**
  - $2^{nd}$ order tensor - **matrix**
**Tensors: Background and Notation**

- **Notation**: $\mathcal{A}^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor

- **Examples**
  - $0^{th}$ order tensor - scalar
  - $1^{st}$ order tensor - vector
  - $2^{nd}$ order tensor - matrix
  - $3^{rd}$ order tensor ...

CVPR 2017 New Tensor Algebra

Lior Horesh & Misha Kilmer
Tensors: Background and Notation

- **Notation**: $A^{n_1 \times n_2 \ldots \times n_j}$ - $j^{th}$ order tensor

- **Examples**
  - $0^{th}$ order tensor - scalar
  - $1^{st}$ order tensor - vector
  - $2^{nd}$ order tensor - matrix
  - $3^{rd}$ order tensor ...

CVPR 2017 New Tensor Algebra

Lior Horesh & Misha Kilmer 18
\( A_{i,j,k} = \) element of \( A \) in row \( i \), column \( j \), tube \( k \)
\( \mathcal{A}_{i,j,k} = \text{element of } \mathcal{A} \text{ in row } i, \text{ column } j, \text{ tube } k \)
\( A_{i,j,k} = \text{element of } A \text{ in row } i, \text{ column } j, \text{ tube } k \)
\[ A_{i,j,k} = \text{element of } A \text{ in row } i, \text{ column } j, \text{ tube } k \]
Tensors: Background and Notation

- **Fiber** - a vector defined by fixing all but one index while varying the rest

![Image of a stack of vectors]

- **Slice** - a matrix defined by fixing all but two indices while varying the rest

![Image of a stack of matrices]
Definition: The $k$-mode multiplication of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \ldots \times n_d}$ with a matrix $U \in \mathbb{R}^{J \times n_k}$ is denoted by $\mathcal{X} \times_k U$ and is of size $n_1 \times \cdots \times n_{k-1} \times J \times n_{k+1} \times \cdots \times n_d$.

Element-wise

$$(\mathcal{X} \times_k U)_{i_1 \cdots i_{k-1} j_{k+1} \cdots i_d} = \sum_{i_k=1}^{n_d} x_{i_1 i_2 \cdots i_d} u_{j_{k}}$$

1-mode multiplication
Tensor Holy Grail and the Matrix Analogy

Find a way to express a tensor that leads to the possibility for compressed representation (near redundancy removed) that maintains important features of the original tensor.

\[
\min \|A - B\|_F \quad \text{s.t. } B \text{ has rank } p \leq r
\]

\[
B = \sum_{i=1}^{p} \sigma_i (v u^{(i)} \circ v^{(i)}) \quad \text{where } A = \sum_{i=1}^{r} \sigma_i (u^{(i)} \circ v^{(i)})
\]
Tensor Decompositions - CP

- **CP (CANDECOMP-PARAFAC) Decomposition**

\[
X \approx \sum_{i=1}^{r} a_i \circ b_i \circ c_i
\]

- **Outer product**

\[
T = u \circ v \circ w \Rightarrow T_{ijk} = u_i v_j w_k
\]

- Columns of $A = [a_1, \ldots, a_r], B = [b_1, \ldots, b_r], C = [c_1, \ldots, c_r]$ are not orthogonal
- If $r$ is minimal, then $r$ is called the *rank* of the tensor
- No perfect procedure for fitting CP for a given number of components

---

2. V. de Silva, L. Lim, *Tensor Rank and the Ill-Posedness of the Best Low-Rank Approximation Problem*, 2008
Tensor Decompositions - Tucker

- **Tucker Decomposition**:

\[ \mathcal{X} \approx \mathcal{C} \times_1 G \times_2 T \times_3 S = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \sum_{k=1}^{r_3} c_{ijk} g_i \circ t_j \circ s_k \]

- \( \mathcal{C} \) is the *core* tensor
- \( G, T, S \) are the *components* of factors
- Can either have diagonal core or orthogonal columns in components [DeLathauwer et al.]
- Truncated Tucker decomposition is *not optimal* in approximating the norm of the difference

\[
\| \mathcal{X} - \mathcal{C} \times_1 G \times_2 T \times_3 S \| 
\]
**t-product**: Let $A$ be $n_1 \times n_2 \times n_3$ and $B$ be $n_2 \times \ell \times n_3$. Then the *t-product* $A * B$ is the $n_1 \times \ell \times n_3$ tensor

$$A * B = \text{fold}(\text{circ}(A) \cdot \text{vec}(B))$$

**fold(vec($B$)) = $B$**

$A_i, \ B_i, \ i = 1, \ldots, n_3$ are frontal slices of $A$ and $B$

A block circulant can be block-diagonalized by a (normalized) DFT in the $2^{nd}$ dimension:

$$(\mathbf{F} \otimes \mathbf{I}) \circ \mathcal{A} (\mathbf{F}^* \otimes \mathbf{I}) = \begin{bmatrix}
\hat{\mathbf{A}}_1 & 0 & \cdots & 0 \\
0 & \hat{\mathbf{A}}_2 & 0 & \cdots \\
0 & \cdots & \ddots & 0 \\
0 & \cdots & 0 & \hat{\mathbf{A}}_n
\end{bmatrix}$$

- Here $\otimes$ is a Kronecker product of matrices.
- If $\mathbf{F}$ is $n \times n$, and $\mathbf{I}$ is $m \times m$, $(\mathbf{F} \otimes \mathbf{I})$ is the $mn \times mn$ block matrix, of $n$ block rows and columns, each block is $m \times m$, where the $ij^{th}$ block is $f_{i,j} \mathbf{I}$.
- But we never implement it this way because an FFT along tube fibers of $\mathcal{A}$ yields a tensor, $\hat{\mathcal{A}}$ whose frontal slices are the $\hat{\mathbf{A}}_i$. 
Definition: The $n \times n \times \ell$ identity tensor $I_{n \times n \times \ell}$ is the tensor whose frontal face is the $n \times n$ identity matrix, and whose other faces are all zeros.
Definition: The $n \times n \times \ell$ identity tensor $\mathcal{I}_{nn\ell}$ is the tensor whose frontal face is the $n \times n$ identity matrix, and whose other faces are all zeros.

Class Exercise: Let $\mathcal{A}$ be $n_1 \times n \times n_3$, show that

$$\mathcal{A} \ast \mathcal{I} = \mathcal{A} \quad \text{and} \quad \mathcal{I} \ast \mathcal{A} = \mathcal{A}$$
Definition: The $n \times n \times \ell$ identity tensor $I_{n n \ell}$ is the tensor whose frontal face is the $n \times n$ identity matrix, and whose other faces are all zeros.

Class Exercise: Let $A$ be $n_1 \times n \times n_3$, show that

$$A \ast I = A \quad \text{and} \quad I \ast A = A$$

$A \ast I$
Definition: The \( n \times n \times \ell \) identity tensor \( I_{nn\ell} \) is the tensor whose frontal face is the \( n \times n \) identity matrix, and whose other faces are all zeros.

---

Class Exercise: Let \( A \) be \( n_1 \times n \times n_3 \), show that

\[
A \ast I = A \quad \text{and} \quad I \ast A = A
\]

\[
A \ast I = \text{fold} ( \text{circ} (A) \cdot \text{vec} (I) )
\]
Definition: The $n \times n \times \ell$ identity tensor $I_{nn\ell}$ is the tensor whose frontal face is the $n \times n$ identity matrix, and whose other faces are all zeros.

Class Exercise: Let $\mathcal{A}$ be $n_1 \times n \times n_3$, show that

$$\mathcal{A} \ast I = \mathcal{A} \quad \text{and} \quad I \ast \mathcal{A} = \mathcal{A}$$

$$\mathcal{A} \ast I = \text{fold} \left( \text{circ} \left( \mathcal{A} \right) \cdot \text{vec} \left( I \right) \right) = \\
\begin{pmatrix}
A_1 & A_{n_3} & A_{n_3-1} & \cdots & A_2 \\
A_2 & A_1 & A_{n_3} & \cdots & A_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_{n_3-1} & A_{n_3-2} & A_{n_3-3} & \cdots & A_{n_3} \\
A_{n_3} & A_{n_3-1} & A_{n_3-2} & \cdots & A_1 \\
\end{pmatrix}
\begin{pmatrix}
I \\
0 \\
0 \\
\vdots \\
0 \\
\end{pmatrix}$$
**Definition:** The \( n \times n \times \ell \) identity tensor \( I_{nn\ell} \) is the tensor whose frontal face is the \( n \times n \) identity matrix, and whose other faces are all zeros.

**Class Exercise:** Let \( \mathcal{A} \) be \( n_1 \times n_\times n_3 \), show that

\[
\mathcal{A} \star I = \mathcal{A} \quad \text{and} \quad I \star \mathcal{A} = \mathcal{A}
\]

\[
\mathcal{A} \star I = \text{fold} \left( \text{circ} \left( \mathcal{A} \right) \cdot \text{vec} \left( I \right) \right) = \begin{pmatrix}
\mathcal{A}_1 & \mathcal{A}_{n_3} & \mathcal{A}_{n_3-1} & \cdots & \mathcal{A}_2 \\
\mathcal{A}_2 & \mathcal{A}_1 & \mathcal{A}_{n_3} & \cdots & \mathcal{A}_3 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\mathcal{A}_{n_3-1} & \mathcal{A}_{n_3-2} & \mathcal{A}_{n_3-3} & \cdots & \mathcal{A}_{n_3} \\
\mathcal{A}_{n_3} & \mathcal{A}_{n_3-1} & \mathcal{A}_{n_3-2} & \cdots & \mathcal{A}_1
\end{pmatrix}
\begin{pmatrix}
I \\
0 \\
0 \\
\vdots \\
0
\end{pmatrix} = \mathcal{A}
\]
**Definition**: If \( A \) is \( n_1 \times n_2 \times n_3 \), then \( A^\top \) is the \( n_2 \times n_1 \times n_3 \) tensor obtained by transposing each of the frontal faces and then reversing the order of transposed faces 2 through \( n_3 \).

**Example**: If \( A \in \mathbb{R}^{n_1 \times n_2 \times 4} \) and its frontal faces are given by the \( n_1 \times n_2 \) matrices \( A_1, A_2, A_3, A_4 \), then

\[
A^\top = \text{fold} \left( \begin{bmatrix}
A_1^\top \\
A_4^\top \\
A_3^\top \\
A_2^\top 
\end{bmatrix} \right)
\]

**Mimetic property**: when \( n = 1 \), the \( \ast \) operator collapses to traditional matrix multiplication between two matrices and transpose becomes matrix transposition.
**Definition:** An $n \times n \times l$ real-valued tensor $Q$ is **orthogonal** if

$$Q^T \ast Q = Q \ast Q^T = I$$

Note that this means that

$$Q(:,i,:)^T \ast Q(:,j,:) = \begin{cases} e_1 & i = j \\ 0 & i \neq j \end{cases}$$
**Theorem:** Let the $T$-SVD of $A \in \mathbb{R}^{\ell \times m \times n}$ be given by $A = U \ast S \ast V^\top$, with $\ell \times \ell \times n$ orthogonal tensor $U$, $m \times m \times n$ orthogonal tensor $V$, and $\ell \times m \times n$ f-diagonal tensor $S$

For $k < \min(l, m)$, define

$$A_k = U(:, 1 : k, :) \ast S(1 : k, 1 : k, :) \ast V^\top(:, 1 : k, :) = \sum_{i=1}^{k} U(:, i, :) \ast S(i, i, :) \ast V(:, i, :)^\top$$

Then

$$A_k = \arg\min_{\hat{A} \in M} \|A - \hat{A}\|$$

where $M = \{C = X \ast Y \mid X \in \mathbb{R}^{\ell \times k \times n}, Y \in \mathbb{R}^{k \times m \times n}\}$
Let $\mathcal{A} \in \mathbb{R}^{m \times p \times n}$, for $k < \min(m, p)$, define

$$\mathcal{A}_k = \sum_{i=1}^{k} U(:, i, :) \ast S(i, i, :) \ast V(:, i, :)^\top$$

Then

$$\mathcal{A}_k = \arg \min_{\tilde{\mathcal{A}} \in M} \| \mathcal{A} - \tilde{\mathcal{A}} \|$$

where $M = \{ \mathcal{C} = \mathcal{X} \ast \mathcal{Y} | \mathcal{X} \in \mathbb{R}^{m \times k \times n}, \mathcal{Y} \in \mathbb{R}^{k \times p \times n} \}$
Let $\mathcal{A}$ be $2 \times 2 \times 2$

$$(F \otimes I)\text{circ}(\mathcal{A})(F^* \otimes I) = \begin{bmatrix} \hat{A}_1 & 0 \\ 0 & \hat{A}_2 \end{bmatrix} \in \mathbb{C}^{4 \times 4}$$

$$\begin{bmatrix} \hat{A}_1 & 0 \\ 0 & \hat{A}_2 \end{bmatrix} = \begin{bmatrix} \hat{U}_1 & 0 \\ 0 & \hat{U}_2 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^{(1)} & 0 \\ 0 & \hat{\sigma}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^{(2)} & 0 \\ 0 & \hat{\sigma}_2^{(2)} \end{bmatrix} \begin{bmatrix} \hat{V}_1^* & 0 \\ 0 & \hat{V}_2^* \end{bmatrix}$$

The $U, S, V^T$ are formed by putting the hat matrices as frontal slices, then ifft along tubes

e.g. $S_{(1,1,:)}$ obtained from ifft of vector $\begin{bmatrix} \hat{\sigma}_1^{(1)} \\ \hat{\sigma}_1^{(2)} \end{bmatrix}$ oriented into screen
$T$-SVD and Multiway PCA

- $X_j, j = 1, 2, \ldots, m$ are the training images
- $M$ is the mean image
- $A(:, j, :) = X_j - M$ stores the mean-subtracted images
- $K = A \ast A^\top = U \ast S \ast S^\top \ast U^\top$ is the covariance tensor
- Left orthogonal $U$ contains the principal components with respect to $K$

$$A(:, j, :) \approx U(:, 1 : k, :) \ast U(:, 1 : k, :)^\top \ast A(:, j, :) = \sum_{t=1}^{k} U(:, t, :) \ast C(t, j, :)$$
\textbf{Theorem:} Let $\mathcal{A}$ be an $\ell \times m \times n$ real-valued tensor, then $\mathcal{A}$ can be factored as

$$\mathcal{A} \ast \mathcal{P} = \mathcal{Q} \ast \mathcal{R}$$

where $\mathcal{Q}$ is \textbf{orthogonal} $\ell \times \ell \times n$, $\mathcal{R}$ is $\ell \times m \times n$ \textbf{f-upper triangular}, and $\mathcal{P}$ is a \textbf{permutation} tensor.

\textbf{Cheaper for} \textbf{updating} and \textbf{downdating}
Face Recognition Task

- Experiment 1: randomly selected 15 images of each person as training set and test all remaining images
- Experiment 2: randomly selected 5 images of each person as the training set and test all remaining images
- Preprocessing: decimated the images by a factor of 3 to $64 \times 56$ pixels
- 20 trials for each experiment

The Extended Yale Face Database B, http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html
### \( T \)-SVD vs. PCA

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>Storage for ( T )-SVD</th>
<th>Storage for PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.8095</td>
<td>34762</td>
<td>98654</td>
</tr>
<tr>
<td>median</td>
<td>0.83</td>
<td>34580</td>
<td>91274</td>
</tr>
<tr>
<td>maximum</td>
<td>0.93</td>
<td>37492</td>
<td>132056</td>
</tr>
<tr>
<td>minimum</td>
<td>0.61</td>
<td>31668</td>
<td>77680</td>
</tr>
</tbody>
</table>

**Table:** Comparison between Tensor SVD and PCA in Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>Storage for ( T )-SVD</th>
<th>Storage for PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.6845</td>
<td>16203</td>
<td>94310</td>
</tr>
<tr>
<td>median</td>
<td>0.7</td>
<td>16318</td>
<td>92100</td>
</tr>
<tr>
<td>maximum</td>
<td>0.79</td>
<td>21812</td>
<td>117888</td>
</tr>
<tr>
<td>minimum</td>
<td>0.5467</td>
<td>12464</td>
<td>73680</td>
</tr>
</tbody>
</table>

**Table:** Comparison between Tensor SVD and PCA in Experiment 2.

---

$\tau$-SVD vs. PCA

### Table: Comparison between Tensor PQR and PCA in Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>Storage for T-PQR</th>
<th>Storage for PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.849</td>
<td>78788</td>
<td>127978</td>
</tr>
<tr>
<td>median</td>
<td>0.86</td>
<td>78624</td>
<td>133998</td>
</tr>
<tr>
<td>maximum</td>
<td>0.95</td>
<td>85540</td>
<td>147592</td>
</tr>
<tr>
<td>minimum</td>
<td>0.72</td>
<td>71708</td>
<td>100984</td>
</tr>
</tbody>
</table>

### Table: Comparison between Tensor PQR and PCA in Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>Storage for T-PQR</th>
<th>Storage for PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.731</td>
<td>40164</td>
<td>121940</td>
</tr>
<tr>
<td>median</td>
<td>0.745</td>
<td>39852</td>
<td>116046</td>
</tr>
<tr>
<td>maximum</td>
<td>0.85</td>
<td>49036</td>
<td>154728</td>
</tr>
<tr>
<td>minimum</td>
<td>0.5667</td>
<td>31160</td>
<td>84732</td>
</tr>
</tbody>
</table>

Given a nonnegative third-order tensor $T \in \mathbb{R}^{\ell \times m \times n}$ and a positive integer $k < \min(l, m, n)$

Find nonnegative $G \in \mathbb{R}^{\ell \times k \times n}$, $H \in \mathbb{R}^{k \times m \times n}$ such that

$$\min_{G, H} \|T - G \ast H\|_F^2$$
Non-Negative Tensor Decompositions - t-product

- Given a nonnegative third-order tensor $\mathbf{T} \in \mathbb{R}^{\ell \times m \times n}$ and a positive integer $k < \min(l, m, n)$
- Find nonnegative $\mathbf{G} \in \mathbb{R}^{\ell \times k \times n}, \mathbf{H} \in \mathbb{R}^{k \times m \times n}$ such that

$$\min_{\mathbf{G}, \mathbf{H}} \| \mathbf{T} - \mathbf{G} \ast \mathbf{H} \|^2_F$$

Facial Recognition Example:
- Dataset: The Center for Biological and Computational Learning (CBCL) Database
- Training images: 200
- $k = 10$
Reconstructed Images Based on NMF, NTF-CP and NTF-GH

\[ A(:,j) \approx WH(:,j) \]
\[ \mathcal{T}(:,j,:) \approx \sum_{i=1}^{K} b^{(i)}(j)(a^{(i)} \circ c^{(i)}) \]
\[ \mathcal{T}(:,j,:) \approx \mathcal{G} \ast \mathcal{H}(:,j,:) \]

If $A$ is an $\ell \times m, \ell \geq m$ matrix with singular values $\sigma_i$, the nuclear norm $\|A\|_\oplus = \sum_{i=1}^{m} \sigma_i$.

However, in the t-SVD, we have singular tubes (the entries of which need not be positive), which sum up to a singular tube!

The entries in the $j$th singular tube are the inverse Fourier coefficients of the length-$n$ vector of the $j$th singular values of $\hat{A}_{::,i, i} = 1..n$.

**Definition**

For $\mathcal{A} \in \mathbb{R}^{\ell \times m \times n}$, our tensor nuclear norm is $\|\mathcal{A}\|_\oplus = \sum_{i=1}^{\min(\ell,m)} \|\sqrt{n}F_{vs_i}\|_1 = \sum_{i=1}^{\min(\ell,m)} \sum_{j=1}^{n} \hat{S}_{i,j}$. (Same as the matrix nuclear norm of $\text{circ}(\mathcal{A})$).
Theorem (Semerci, Hao, Kilmer, Miller)

The tensor nuclear norm is a valid norm.

Since the t-SVD extends to higher-order tensors [Martin et al, 2012], the norm does, as well.
Tensor Completion

Given unknown tensor $T_M$ of size $n_1 \times n_2 \times n_3$, given a subset of entries $\{T_{M_{ijk}} : (i, j, k) \in \Omega\}$ where $\Omega$ is an indicator tensor of size $n_1 \times n_2 \times n_3$. Recover the entire $T_M$:

$$\min \|T_X\|_\otimes$$
subject to $P_\Omega(T_X) = P_\Omega(T_M)$

The $(i, j, k)_{th}$ component of $P_\Omega(T_X)$ is equal to $T_{M_{ijk}}$ if $(i, j, k) \in \Omega$ and zero otherwise.

Similar to the previous problem, this can be solved by ADMM, with 3 update steps, one which decouples, one that is a shrinkage / thresholding step.
Numerical Results

- MERL\(^3\) video, Basketball video

\(^3\text{with thanks to A. Agrawal}\)
Numerical Results

![Graph 1: Error in SNR vs Sample Rate](image1)

![Graph 2: Error in SNR vs Sample Rate](image2)
Toolbox and References


