When tensor decomposition meets compressed sensing

Pierre Comon

I3S, CNRS, University of Nice - Sophia Antipolis, France
Collaborator: Lek-Heng Lim

Sept. 27-30, 2010

Blind Source Separation & Sparse Representation

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

$\mathbf{H}$ is $K \times P$, *underdetermined*: $K < P$

- *Sparse representation*: Columns $\mathbf{h}_n \in \mathbb{D}$, known dictionary
- *BSS*: $\mathbf{H}$ unknown
  - $\mathbf{s}$ sparse
  - $\mathbf{s}$ not sparse

Blind Source Separation & Sparse Representation

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

$\mathbf{H}$ is $K \times P$, *underdetermined*: $K < P$

- *Sparse representation*: Columns $\mathbf{h}_n \in \mathbb{D}$, known dictionary
- *BSS*: $\mathbf{H}$ unknown
  - $\mathbf{s}$ sparse
  - $\mathbf{s}$ not sparse

Blind Source Separation & Sparse Representation

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

$\mathbf{H}$ is $K \times P$, *underdetermined*: $K < P$

- *Sparse representation*: Columns $\mathbf{h}_n \in \mathbb{D}$, known dictionary
- *BSS*: $\mathbf{H}$ unknown
  - $\mathbf{s}$ sparse
  - $\mathbf{s}$ *not sparse*

Blind Source Separation & Sparse Representation

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

$\mathbf{H}$ is $K \times P$, *underdetermined*: $K < P$

- *Sparse representation*: Columns $\mathbf{h}_n \in \mathbb{D}$, known dictionary
- *BSS*: $\mathbf{H}$ unknown
  - $\mathbf{s}$ sparse
  - $\mathbf{s}$ *not sparse*

Blind Source Separation & Sparse Representation

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

$\mathbf{H}$ is $K \times P$, *underdetermined*: $K < P$

- *Sparse representation*: Columns $\mathbf{h}_n \in \mathbb{D}$, known dictionary
- *BSS*: $\mathbf{H}$ unknown
  - $\mathbf{s}$ sparse
  - $\mathbf{s}$ *not sparse*

Blind identification of linear mixtures

- **Linear mixtures:**

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

If s_ℓ *statistically independent*, we have the *core equation*:

$$\Psi(\mathbf{u}) = \sum_{\ell=1}^{P} \varphi_\ell(\mathbf{u}^{\mathsf{T}}\mathbf{H})$$

- Take 3rd derivatives at point $\mathbf{u}$:

$$T_{ijk}(\mathbf{u}) = \sum_{\ell=1}^{P} H_{i\ell}\, H_{j\ell}\, H_{k\ell}\;\; C_{\ell\ell\ell}(\mathbf{u}) \tag{1}$$

Blind identification of linear mixtures

- **Linear mixtures:**

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

If s_ℓ *statistically independent*, we have the *core equation*:

$$\Psi(\mathbf{u}) = \sum_{\ell=1}^{P} \varphi_\ell(\mathbf{u}^{\mathsf{T}}\mathbf{H})$$

- Take 3rd derivatives at point $\mathbf{u}$:

$$T_{ijk}(\mathbf{u}) = \sum_{\ell=1}^{P} H_{i\ell}\, H_{j\ell}\, H_{k\ell}\;\; C_{\ell\ell\ell}(\mathbf{u}) \tag{1}$$

At $\mathbf{u} = \mathbf{0}$ ➠ symmetric decomposition of T_{ijk} *[HOS]*
At $\mathbf{u} \neq \mathbf{0}$ ➠ *[Taleb, Comon-Rajih, Yeredor]*

Blind identification of linear mixtures

- **Linear mixtures:**

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

If s_ℓ *statistically independent*, we have the *core equation*:

$$\Psi(\mathbf{u}) = \sum_{\ell=1}^{P} \varphi_\ell(\mathbf{u}^\mathsf{T}\mathbf{H})$$

- Take 3rd derivatives at point $\mathbf{u}$:

$$T_{ijk}(\mathbf{u}) = \sum_{\ell=1}^{P} H_{i\ell}\, H_{j\ell}\, H_{k\ell} \;\; C_{\ell\ell\ell}(\mathbf{u}) \tag{1}$$

At $\mathbf{u} = \mathbf{0}$ ➽ symmetric decomposition of T_{ijk} *[HOS]*
At $\mathbf{u} \neq \mathbf{0}$ ➽ *[Taleb, Comon-Rajih, Yeredor]*

Blind identification of linear mixtures

- **Linear mixtures:**

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

If s_ℓ *statistically independent*, we have the *core equation*:

$$\Psi(\mathbf{u}) = \sum_{\ell=1}^{P} \varphi_\ell(\mathbf{u}^\mathsf{T}\mathbf{H})$$

- Take 3rd derivatives at point $\mathbf{u}$:

$$T_{ijk}(\mathbf{u}) = \sum_{\ell=1}^{P} H_{i\ell}\,H_{j\ell}\,H_{k\ell}\;\; C_{\ell\ell\ell}(\mathbf{u}) \tag{1}$$

At $\mathbf{u} = \mathbf{0}$ ➽ symmetric decomposition of T_{ijk} *[HOS]*

At $\mathbf{u} \neq \mathbf{0}$ ➽ *[Taleb, Comon-Rajih, Yeredor]*

Blind identification of linear mixtures

- **Linear mixtures:**

$$\mathbf{x} = \mathbf{H}\,\mathbf{s}$$

If s_ℓ *statistically independent*, we have the *core equation*:

$$\Psi(\mathbf{u}) = \sum_{\ell=1}^{P} \varphi_\ell(\mathbf{u}^\mathsf{T}\mathbf{H})$$

- Take 3rd derivatives at point $\mathbf{u}$:

$$T_{ijk}(\mathbf{u}) = \sum_{\ell=1}^{P} H_{i\ell}\, H_{j\ell}\, H_{k\ell} \;\; C_{\ell\ell\ell}(\mathbf{u}) \tag{1}$$

At $\mathbf{u} = \mathbf{0}$ ➽ symmetric decomposition of T_{ijk} *[HOS]*
At $\mathbf{u} \neq \mathbf{0}$ ➽ *[Taleb, Comon-Rajih, Yeredor]*

Tensors: General items

Arrays and Multi-linearity

- A tensor of *order d* is a multi-linear map:

$$\mathcal{S}_1^* \otimes \ldots \otimes \mathcal{S}_m^* \longrightarrow \mathcal{S}_{m+1} \otimes \ldots \otimes \mathcal{S}_d$$

- Once bases of spaces $\mathcal{S}_\ell$ are fixed, they can be represented by *d-way arrays* of coordinates

➠ bilinear form, or linear operator: represented by a *matrix*
➠ trilinear form, or bilinear operator: by a *3rd order tensor*.

Arrays and Multi-linearity

- A tensor of *order d* is a multi-linear map:

$$\mathcal{S}_1^* \otimes \ldots \otimes \mathcal{S}_m^* \longrightarrow \mathcal{S}_{m+1} \otimes \ldots \otimes \mathcal{S}_d$$

- Once bases of spaces $\mathcal{S}_\ell$ are fixed, they can be represented by *d-way arrays* of coordinates

➽ bilinear form, or linear operator: represented by a *matrix*
➽ trilinear form, or bilinear operator: by a *3rd order tensor*.

Multi-linearity

Compact notation

Linear change in contravariant spaces:

$$T'_{ijk} = \sum_{npq} A_{in} B_{jp} C_{kq} T_{npq}$$

Denoted compactly

$$\mathcal{T}' = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{T} \tag{2}$$

Example: covariance matrix

$$\mathbf{z} = \mathbf{A}\mathbf{x} \Rightarrow \mathbf{R}_z = (\mathbf{A}, \mathbf{A}) \cdot \mathbf{R}_x = \mathbf{A}\,\mathbf{R}_x \mathbf{A}^{\mathsf{T}}$$

Decomposable tensor

- A dth order "*decomposable*" tensor is the tensor product of d vectors:

$$\mathcal{T} = \mathbf{u} \otimes \mathbf{v} \otimes \ldots \otimes \mathbf{w}$$

and has coordinates $T_{ij\ldots k} = u_i \, v_j \ldots w_k$.

- may be seen as a discretization of multivariate *functions whose variables separate*:

$$t(x, y, \ldots, z) = u(x) \, v(y) \, \ldots \, w(z)$$

- Nothing else but *rank-1* tensors, with forthcoming definition

Decomposable tensor

- A dth order "*decomposable*" tensor is the tensor product of d vectors:

$$\mathcal{T} = \mathbf{u} \otimes \mathbf{v} \otimes \ldots \otimes \mathbf{w}$$

  and has coordinates $T_{ij\ldots k} = u_i \, v_j \ldots \, w_k$.

- may be seen as a discretization of multivariate *functions whose variables separate*:

$$t(x, y, \ldots, z) = u(x) \, v(y) \, \ldots \, w(z)$$

- Nothing else but *rank-1* tensors, with forthcoming definition

Decomposable tensor

- A dth order "*decomposable*" tensor is the tensor product of d vectors:

$$\mathcal{T} = \mathbf{u} \otimes \mathbf{v} \otimes \ldots \otimes \mathbf{w}$$

and has coordinates $T_{ij\ldots k} = u_i\, v_j \ldots\, w_k$.

- may be seen as a discretization of multivariate *functions whose variables separate*:

$$t(x, y, \ldots, z) = u(x)\, v(y)\, \ldots\, w(z)$$

- Nothing else but *rank-1* tensors, with forthcoming definition

Example

- Take $$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
- Then $$\mathbf{v}^{\otimes 3} \stackrel{\text{def}}{=} \mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v} = \left[\begin{array}{rr|rr} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{array} \right]$$
- This is a “decomposable” symmetric tensor → *rank-1*

blue bullets $= 1$, red bullets $= -1$.

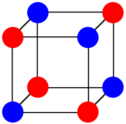
Example

- Take $$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
- Then $$\mathbf{v}^{\otimes 3} \stackrel{\text{def}}{=} \mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v} = \left[\begin{array}{rr|rr} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{array}\right]$$
- This is a "decomposable" symmetric tensor → *rank-1*

blue bullets $= 1$, red bullets $= -1$.

Example

- Take $$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
- Then $$\mathbf{v}^{\otimes 3} \stackrel{\text{def}}{=} \mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v} = \left[\begin{array}{cc|cc} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{array}\right]$$
- This is a "decomposable" symmetric tensor → *rank-1*



blue bullets $= 1$, red bullets $= -1$.

From SVD to tensor decompositions

Matrix SVD, $\mathbf{M} = (\mathbf{U}, \mathbf{V}) \cdot \mathbf{\Sigma}$, may be extended in at least two ways to tensors

- **Keep orthogonality:** Orthogonal Tucker, HOSVD

$$\mathcal{T} = (\mathbf{U}, \mathbf{V}, \mathbf{W}) \cdot \mathcal{C}$$

$\mathcal{C}$ is $R_1 \times R_2 \times R_3$: *multilinear rank* $= (R_1, R_2, R_3)$

- **Keep diagonality:** Canonical Polyadic decomposition (CP)

$$\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$$

$\mathcal{L}$ is $R \times R \times R$ diagonal, $\lambda_i \neq 0$: *rank* $= R$.

From SVD to tensor decompositions

Matrix SVD, $\mathbf{M} = (\mathbf{U}, \mathbf{V}) \cdot \mathbf{\Sigma}$, may be extended in at least two ways to tensors

- **Keep orthogonality:** Orthogonal Tucker, HOSVD

$$\mathcal{T} = (\mathbf{U}, \mathbf{V}, \mathbf{W}) \cdot \mathcal{C}$$

$\mathcal{C}$ is $R_1 \times R_2 \times R_3$: *multilinear rank* $= (R_1, R_2, R_3)$

- **Keep diagonality:** Canonical Polyadic decomposition (CP)

$$\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$$

$\mathcal{L}$ is $R \times R \times R$ diagonal, $\lambda_i \neq 0$: *rank* $= R$.

From SVD to tensor decompositions

Matrix SVD, $\mathbf{M} = (\mathbf{U}, \mathbf{V}) \cdot \mathbf{\Sigma}$, may be extended in at least two ways to tensors

- **Keep orthogonality:** Orthogonal Tucker, HOSVD

$$\mathcal{T} = (\mathbf{U}, \mathbf{V}, \mathbf{W}) \cdot \mathcal{C}$$

$\mathcal{C}$ is $R_1 \times R_2 \times R_3$: *multilinear rank* $= (R_1, R_2, R_3)$

- **Keep diagonality:** Canonical Polyadic decomposition (CP)

$$\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$$

$\mathcal{L}$ is $R \times R \times R$ diagonal, $\lambda_i \neq 0$: *rank* $= R$.

From SVD to tensor decompositions

Matrix SVD, $\mathbf{M} = (\mathbf{U}, \mathbf{V}) \cdot \mathbf{\Sigma}$, may be extended in at least two ways to tensors

- **Keep orthogonality:** Orthogonal Tucker, HOSVD

$$\mathcal{T} = (\mathbf{U}, \mathbf{V}, \mathbf{W}) \cdot \mathcal{C}$$

$\mathcal{C}$ is $R_1 \times R_2 \times R_3$: *multilinear rank* $= (R_1, R_2, R_3)$

- **Keep diagonality:** Canonical Polyadic decomposition (CP)

$$\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$$

$\mathcal{L}$ is $R \times R \times R$ diagonal, $\lambda_i \neq 0$: *rank* $= R$.

From SVD to tensor decompositions

Matrix SVD, $\mathbf{M} = (\mathbf{U}, \mathbf{V}) \cdot \mathbf{\Sigma}$, may be extended in at least two ways to tensors

- **Keep orthogonality:** Orthogonal Tucker, HOSVD

$$\mathcal{T} = (\mathbf{U}, \mathbf{V}, \mathbf{W}) \cdot \mathcal{C}$$

$\mathcal{C}$ is $R_1 \times R_2 \times R_3$: *multilinear rank* $= (R_1, R_2, R_3)$

- **Keep diagonality:** Canonical Polyadic decomposition (CP)

$$\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$$

$\mathcal{L}$ is $R \times R \times R$ diagonal, $\lambda_i \neq 0$: *rank* $= R$.

Exact Canonical Polyadic (CP) decomposition

Canonical Polyadic (CP) decomposition

- Any $I \times J \times \cdots \times K$ tensor $\mathcal{T}$ can be decomposed as

$$\mathcal{T} = \sum_q \lambda_q \, \mathbf{u}^{(q)} \otimes \mathbf{v}^{(q)} \otimes \ldots \otimes \mathbf{w}^{(q)}$$

➽ "Polyadic form" *[Hitchcock'27]*

- The *tensor rank* of $\mathcal{T}$ is the minimal number $R(\mathcal{T})$ of "decomposable" terms such that equality holds.
- May impose unit norm vectors $\mathbf{u}^{(q)}$, $\mathbf{v}^{(q)}$, $\ldots$ $\mathbf{w}^{(q)}$

Canonical Polyadic (CP) decomposition

- Any $I \times J \times \cdots \times K$ tensor $\mathcal{T}$ can be decomposed as

$$\mathcal{T} = \sum_{q}^{R(\mathcal{T})} \lambda_q \, \mathbf{u}^{(q)} \otimes \mathbf{v}^{(q)} \otimes \ldots \otimes \mathbf{w}^{(q)}$$

➜ "Polyadic form" *[Hitchcock'27]*

- The *tensor rank* of $\mathcal{T}$ is the minimal number $R(\mathcal{T})$ of "decomposable" terms such that equality holds.
- May impose unit norm vectors $\mathbf{u}^{(q)}, \mathbf{v}^{(q)}, \ldots \mathbf{w}^{(q)}$

Canonical Polyadic (CP) decomposition

- Any $I \times J \times \cdots \times K$ tensor $\mathcal{T}$ can be decomposed as

$$\mathcal{T} = \sum_{q}^{R(\mathcal{T})} \lambda_q \, \mathbf{u}^{(q)} \otimes \mathbf{v}^{(q)} \otimes \ldots \otimes \mathbf{w}^{(q)}$$

  ➽ "Polyadic form" *[Hitchcock'27]*

- The *tensor rank* of $\mathcal{T}$ is the minimal number $R(\mathcal{T})$ of "decomposable" terms such that equality holds.
- May impose unit norm vectors $\mathbf{u}^{(q)}$, $\mathbf{v}^{(q)}$, … $\mathbf{w}^{(q)}$

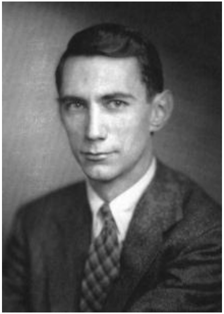
Hitchcock

Frank Lauren Hitchcock
(1875-1957)
[Courtesy of L-H.Lim]

Hitchcock

Frank Lauren Hitchcock (1875-1957)

[Courtesy of L-H.Lim]



Claude Elwood Shannon (1916-2001)

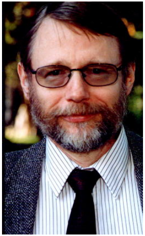
Towards a unique terminology

- Minimal *Polyadic Form [Hitchcock'27]*
- *Canonical decomposition [Weinstein'84, Carroll'70, Chiantini-Ciliberto'06, Comon'00, Khoromskij, Tyrtyshnikov]*
- *Parafac [Harshman'70, Sidiropoulos'00]*
- *Optimal computation [Strassen'83]*
- Minimum-length *additive decomposition* (AD) *[Iarrobino'96]*

Suggestion:

- *Canonical Polyadic decomposition (CP) [Comon'08, Grasedyk, Espig...]*
- *CP* does also already stand for Candecomp/Parafac *[Bro'97, Kiers'98, tenBerge'04...]*

Psychometrics



Richard A. Harshman
(1970)

J. Douglas Carroll
(1970)

Uniqueness: Kruskal 1/2

The *Kruskal rank* of a matrix $\mathbf{A}$ is the maximum number k_A, such that *any* subset of k_A columns are linearly independent.

Uniqueness: Kruskal 2/2

Sufficient condition for uniqueness of CP
[Kruskal'77, Sidiropoulos-Bro'00, Landsberg'09]:

Essential uniqueness is ensured if tensor rank R is below the so-called *Kruskal's bound*:

$$2R + 2 \leq k_A + k_B + k_C \tag{3}$$

or *generically*, for a tensor of order d and dimensions N_ℓ:

$$2R \leq \sum_{\ell=1}^{d} \min(N_\ell, R) - d + 1$$

- Bound much smaller than expected rank:
  $\exists$ *a much better bound*, in almost sure sense

Uniqueness: Kruskal 2/2

Sufficient condition for uniqueness of CP
[Kruskal'77, Sidiropoulos-Bro'00, Landsberg'09]:

Essential uniqueness is ensured if tensor rank R is below the so-called *Kruskal's bound*:

$$2R + 2 \leq k_A + k_B + k_C \tag{3}$$

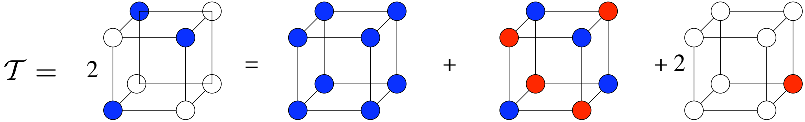
or *generically*, for a tensor of order d and dimensions N_ℓ:

$$2R \leq \sum_{\ell=1}^{d} \min(N_\ell, R) - d + 1$$

➽ Bound much smaller than expected rank:
$\exists$ *a much better bound*, in almost sure sense

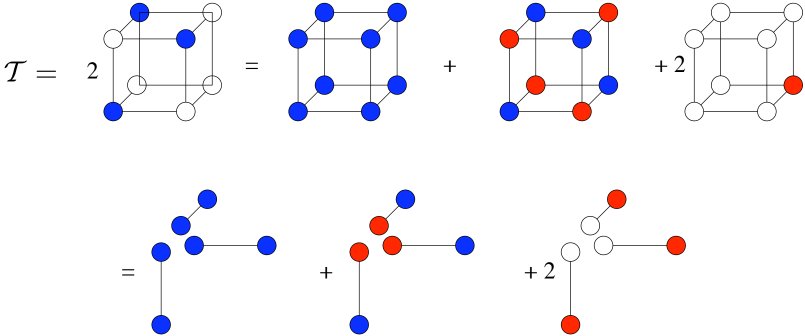
Rank-3 example 1/2

Example



Rank-3 example 1/2

Example



blue bullets $= 1$, red bullets $= -1$.

Rank-3 example 2/2

Conclusion: the $2 \times 2 \times 2$ tensor

$$\mathcal{T} = 2 \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right]$$

admits the CP

$$\mathcal{T} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{\otimes 3} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}^{\otimes 3} + 2 \begin{pmatrix} 0 \\ -1 \end{pmatrix}^{\otimes 3}$$

and has rank 3, hence *larger* than dimension

Approximate Canonical Polyadic (CP) decomposition

Why need for approximation?

- Additive noise in measurements
- Noise has a continuous probability distribution
- Then tensor rank is *generic*
- Hence there are often *infinitely many* CP decompositions

➠ Approximations aim at getting rid of noise, *and* at restoring uniqueness:

$$Arg \inf_{\mathbf{a}(p),\mathbf{b}(p),\mathbf{c}(p)} ||\mathcal{T} - \sum_{p=1}^{r} \mathbf{a}(p) \otimes \mathbf{b}(p) \ldots \otimes \mathbf{c}(p)||^2 \quad (4)$$

But infimum may be reached for tensors of rank $> r$!

Why need for approximation?

- Additive noise in measurements
- Noise has a continuous probability distribution
- Then tensor rank is *generic*
- Hence there are often *infinitely many* CP decompositions

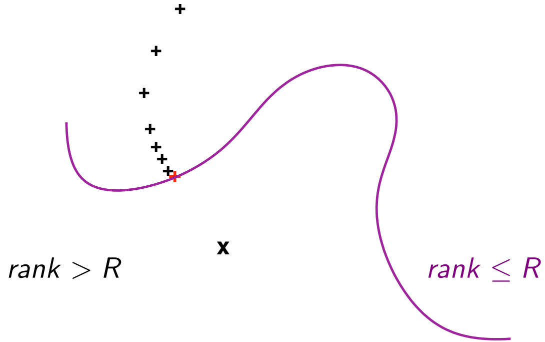
➽ Approximations aim at getting rid of noise, *and* at restoring uniqueness:

$$Arg \inf_{\mathbf{a}(p),\mathbf{b}(p),\mathbf{c}(p)} ||\mathcal{T} - \sum_{p=1}^{r} \mathbf{a}(p) \otimes \mathbf{b}(p) \dots \otimes \mathbf{c}(p)||^2 \tag{4}$$

But infimum may be reached for tensors of rank $> r$!

Border rank

$\mathcal{T}$ has *border rank* R iff it is the limit of tensors of rank R, and not the limit of tensors of lower rank.
[Bini'79, Schönhage'81, Strassen'83, Likteig'85]



Example

Let $\mathbf{u}$ and $\mathbf{v}$ be non collinear vectors. Define $\mathcal{T}_0$ *[Comon et al.'08]*:

$$\mathcal{T}_0 = \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

And define sequence $\mathcal{T}_\varepsilon = \frac{1}{\varepsilon} \left[(\mathbf{u} + \varepsilon\, \mathbf{v})^{\otimes 4} - \mathbf{u}^{\otimes 4} \right]$.

Then $\mathcal{T}_\varepsilon \to \mathcal{T}_0$ as $\varepsilon \to 0$

Hence $\text{rank}\{\mathcal{T}_0\} = 4$, but $\underline{\text{rank}}\{\mathcal{T}_0\} = 2$

Example

Let $\mathbf{u}$ and $\mathbf{v}$ be non collinear vectors. Define $\mathcal{T}_0$ *[Comon et al.'08]*:

$$\mathcal{T}_0 = \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

And define sequence $\mathcal{T}_\varepsilon = \frac{1}{\varepsilon} \left[(\mathbf{u} + \varepsilon\, \mathbf{v})^{\otimes 4} - \mathbf{u}^{\otimes 4} \right]$.

Then $\mathcal{T}_\varepsilon \rightarrow \mathcal{T}_0$ as $\varepsilon \rightarrow 0$

➽ Hence $\text{rank}\{\mathcal{T}_0\} = 4$, but $\underline{\text{rank}}\{\mathcal{T}_0\} = 2$

Example

Let **u** and **v** be non collinear vectors. Define $\mathcal{T}_0$ *[Comon et al.'08]*:

$$\mathcal{T}_0 = \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

And define sequence $\mathcal{T}_\varepsilon = \frac{1}{\varepsilon} \left[(\mathbf{u} + \varepsilon\, \mathbf{v})^{\otimes 4} - \mathbf{u}^{\otimes 4} \right]$.

Then $\mathcal{T}_\varepsilon \to \mathcal{T}_0$ as $\varepsilon \to 0$

➾ Hence $\text{rank}\{\mathcal{T}_0\} = 4$, but $\underline{\text{rank}}\{\mathcal{T}_0\} = 2$

Example

Let **u** and **v** be non collinear vectors. Define $\mathcal{T}_0$ *[Comon et al.'08]*:

$$\mathcal{T}_0 = \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

And define sequence $\mathcal{T}_\varepsilon = \frac{1}{\varepsilon} \left[(\mathbf{u} + \varepsilon\, \mathbf{v})^{\otimes 4} - \mathbf{u}^{\otimes 4} \right]$.

Then $\mathcal{T}_\varepsilon \to \mathcal{T}_0$ as $\varepsilon \to 0$

➽ Hence $\text{rank}\{\mathcal{T}_0\} = 4$, but $\underline{\text{rank}}\{\mathcal{T}_0\} = 2$

Ill-posedness

Tensors for which $\underline{\text{rank}}\{\mathcal{T}\} < \text{rank}\{\mathcal{T}\}$ are such that the approximating sequence contains several decomposable tensors which

- tend to infinity *and*
- cancel each other, viz, some columns become close to collinear

Ideas towards a well-posed problem:

- Prevent collinearity *or* bound columns.

Ill-posedness

Tensors for which $\underline{\text{rank}}\{\mathcal{T}\} < \text{rank}\{\mathcal{T}\}$ are such that the approximating sequence contains several decomposable tensors which

- tend to infinity *and*
- cancel each other, viz, some columns become close to collinear

Ideas towards a well-posed problem:

- Prevent collinearity *or* bound columns.

Ill-posedness

Tensors for which $\underline{\text{rank}}\{\mathcal{T}\} < \text{rank}\{\mathcal{T}\}$ are such that the approximating sequence contains several decomposable tensors which

- tend to infinity *and*
- cancel each other, viz, some columns become close to collinear

Ideas towards a well-posed problem:

- Prevent collinearity *or* bound columns.

Ill-posedness

Tensors for which $\underline{\text{rank}}\{\mathcal{T}\} < \text{rank}\{\mathcal{T}\}$ are such that the approximating sequence contains several decomposable tensors which

- tend to infinity *and*
- cancel each other, viz, some columns become close to collinear

Ideas towards a well-posed problem:

- Prevent collinearity *or* bound columns.

Remedies

1. Impose *orthogonality* of columns within factor matrices *[Comon'92]*
2. Impose *orthogonality* between decomposable tensors *[Kolda'01]*
3. Prevent tendency to infinity by *norm constraint* on factor matrices *[Paatero'00]*
4. Nonnegative tensors: impose decomposable tensors to be *nonnegative* *[Lim-Comon'09]* $\to$ "nonnegative rank"
5. Impose *minimal angle* between columns of factor matrices *[Lim-Comon'10]*

Remedies

1. Impose *orthogonality* of columns within factor matrices *[Comon'92]*
2. Impose *orthogonality* between decomposable tensors *[Kolda'01]*
3. Prevent tendency to infinity by *norm constraint* on factor matrices *[Paatero'00]*
4. Nonnegative tensors: impose decomposable tensors to be *nonnegative* *[Lim-Comon'09]* $\rightarrow$ "nonnegative rank"
5. Impose *minimal angle* between columns of factor matrices *[Lim-Comon'10]*

Nonnegativity constraint: example

Fluorescence Spectroscopy 1/3

An optical excitation produces several effects

- Rayleigh diffusion
- Raman diffusion
- Fluorescence

  At low concentrations, Beer-Lambert law can be linearized *[Luciani'09]*

$$I(\lambda_f, \lambda_e, k) = I_o \sum_{\ell} \gamma_\ell(\lambda_f)\, \epsilon_\ell(\lambda_e)\, c_{k,\ell}$$

➽ Hence 3rd array decomposition with real *nonnegative factors [Bro'97].*

Fluorescence Spectroscopy 1/3

An optical excitation produces several effects

- Rayleigh diffusion
- Raman diffusion
- Fluorescence

At low concentrations, Beer-Lambert law can be linearized *[Luciani'09]*

$$I(\lambda_f, \lambda_e, k) = I_o \sum_{\ell} \gamma_\ell(\lambda_f)\, \epsilon_\ell(\lambda_e)\, c_{k,\ell}$$

➠ Hence 3rd array decomposition with real *nonnegative factors [Bro'97]*.

Fluorescence Spectroscopy 1/3

An optical excitation produces several effects

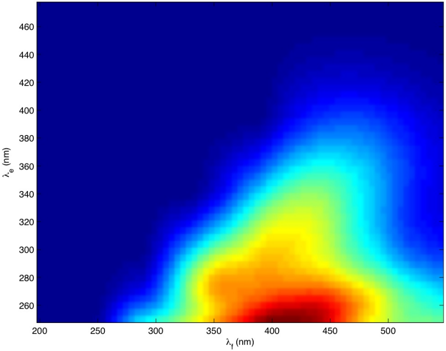
- Rayleigh diffusion
- Raman diffusion
- Fluorescence
  At low concentrations, Beer-Lambert law can be linearized *[Luciani'09]*

$$I(\lambda_f, \lambda_e, k) = I_o \sum_{\ell} \gamma_\ell(\lambda_f)\, \epsilon_\ell(\lambda_e)\, c_{k,\ell}$$

➽ Hence 3rd array decomposition with real *nonnegative factors [Bro'97]*.

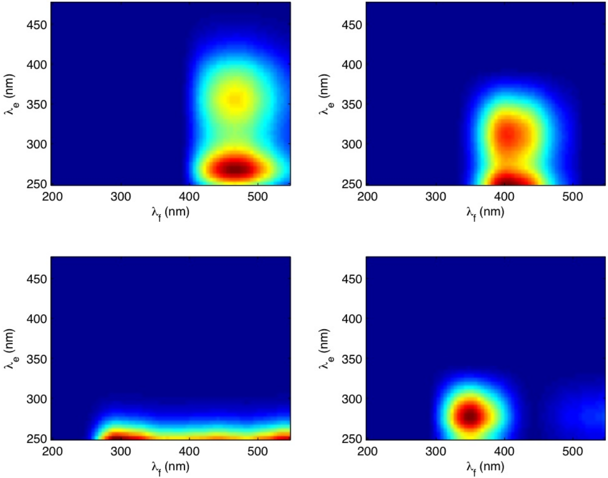
Fluorescence Spectroscopy 2/3

Mixture of 4 solutes (one concentration shown)



Fluorescence Spectroscopy 3/3

Obtained results with $R = 4$



Approximate CP decomposition: Coherence constraint

Coherence

- **Definition:** *[Donoho'03, Gribonval'03, Candès'07]*
  let $\mathbf{A}$ a matrix with unit-norm columns, $\mathbf{a}_p$.

$$\mu_A = \max_{p \neq q} |\langle \mathbf{a}_p, \mathbf{a}_q \rangle| \tag{5}$$

- this *"coherence of* $\mathbf{A}$*"* is used in Sparse Representation theory

Existence

Best rank-R approximate under angular constraint

- **Proposition:** *[Lim-Comon'2010]*
  Let $\mathcal{L}$ diagonal, and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ have R unit norm columns.
  If $R < [\mu_A \mu_B \mu_C]^{-1}$, then:

$$\inf_{\mathbf{A},\mathbf{B},\mathbf{C}} ||\mathcal{T} - (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}||$$

is attained.

Uniqueness: Back to Kruskal

- **Lemma:** (spark) *[Gribonval'03]*

$$k_A \geq \frac{1}{\mu_A}$$

Uniqueness

- **Proposition:** *[Lim-Comon'10]*
  If $\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$, with $\lambda_p \neq 0$ for $1 \leq p \leq R$, **A**, **B**, **C** matrices with unit norm columns, and:

$$2R + 2 \leq \frac{1}{\mu_A} + \frac{1}{\mu_B} + \frac{1}{\mu_C} \tag{6}$$

  then $\mathcal{T}$ has *a unique CP* decomposition of rank R, up to unit modulus scalar factors (ρ_A, ρ_B, ρ_C), $\rho_A \rho_B \rho_C = 1$.

- Hence it suffices that *one* μ is small, not *every*

Uniqueness

- **Proposition:** *[Lim-Comon'10]*
  If $\mathcal{T} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}$, with $\lambda_p \neq 0$ for $1 \leq p \leq R$, **A**, **B**, **C** matrices with unit norm columns, and:

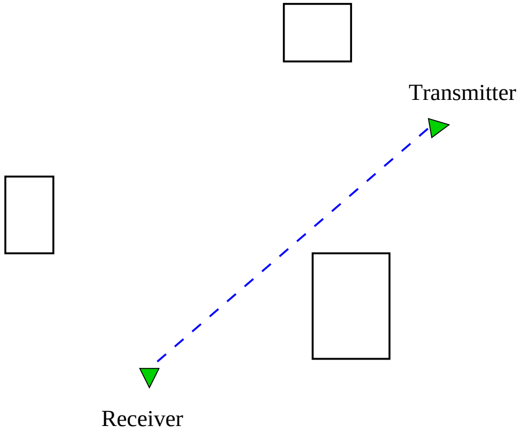
$$2R + 2 \leq \frac{1}{\mu_A} + \frac{1}{\mu_B} + \frac{1}{\mu_C} \tag{6}$$

  then $\mathcal{T}$ has *a unique CP* decomposition of rank R, up to unit modulus scalar factors (ρ_A, ρ_B, ρ_C), $\rho_A \rho_B \rho_C = 1$.

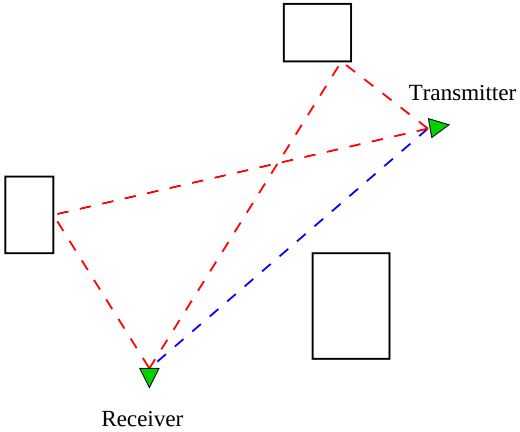
- Hence it suffices that *one* μ is small, not *every*

Angular constraint: example

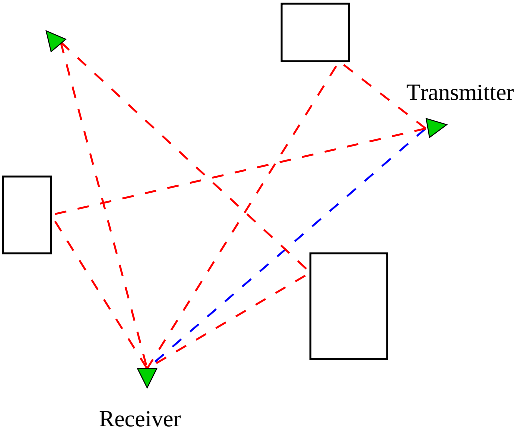
Antenna Array Processing



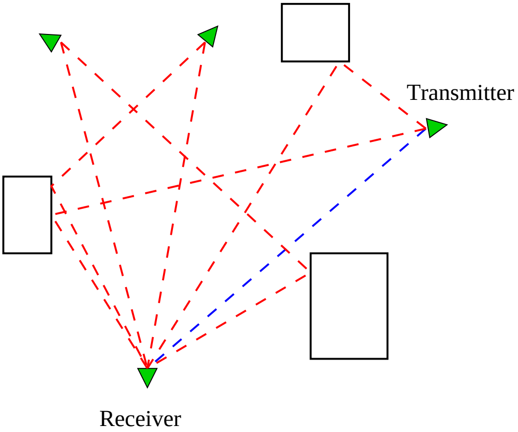
Antenna Array Processing



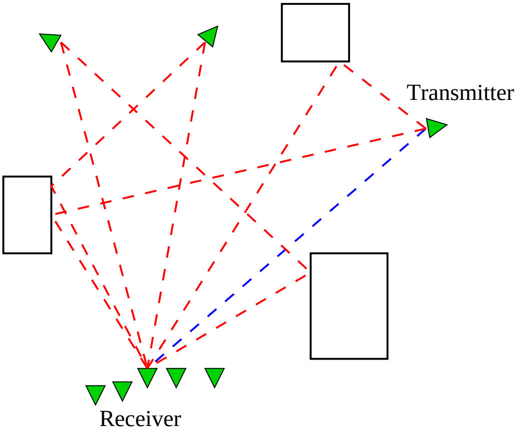
Antenna Array Processing



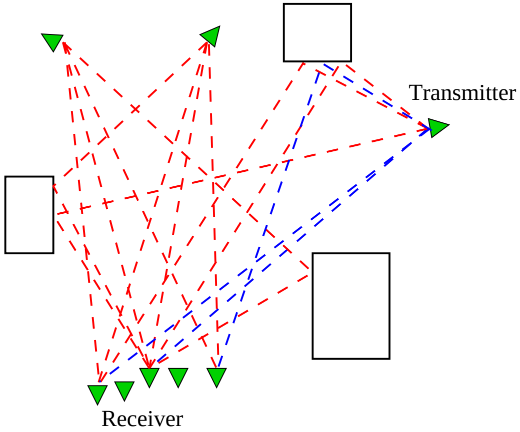
Antenna Array Processing



Antenna Array Processing



Antenna Array Processing



Narrow band model in the far field

Modeling the signals received on an array of antennas generally leads to a *matrix decomposition*:

$$T_{ij} = \sum_q a_{iq}\, s_{jq}$$

i: space	q: path, source	**A**: antenna gains
j: time		**S**: transmitted signals

But in the presence of additional *diversity*, a tensor can be constructed, thanks to new index p

Narrow band model in the far field

Modeling the signals received on an array of antennas generally leads to a *matrix decomposition*:

$$T_{ijp} = \sum_q a_{iq}\, s_{jq}\, h_{pq}$$

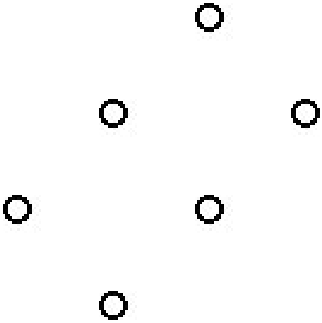
i: space	q: path, source	**A**: antenna gains
j: time		**S**: transmitted signals

But in the presence of additional *diversity*, a tensor can be constructed, thanks to new index p

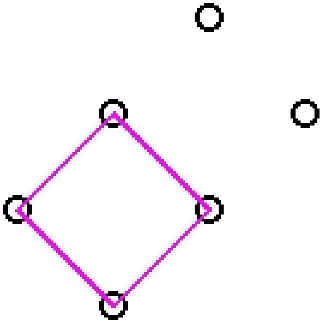
Possible diversities in Signal Processing

- space
- time
- space translation (array geometry)
- time translation (chip)
- frequency/wavenumber (nonstationarity)
- excess bandwidth (oversampling)
- cyclostationarity
- polarization
- finite alphabet
- ...

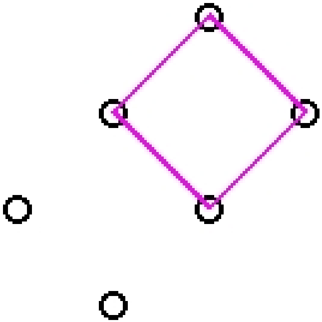
Space translation diversity (1/2)



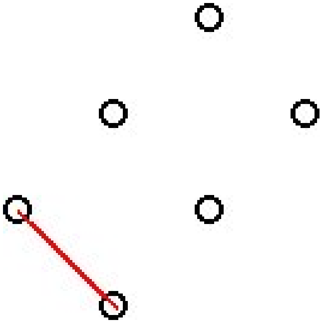
Space translation diversity (1/2)



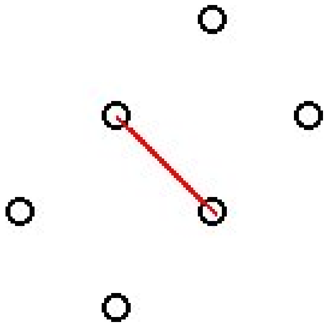
Space translation diversity (1/2)



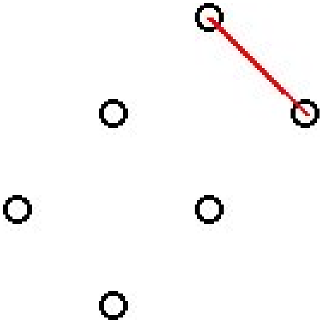
Space translation diversity (1/2)



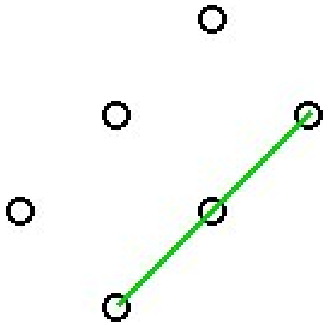
Space translation diversity (1/2)



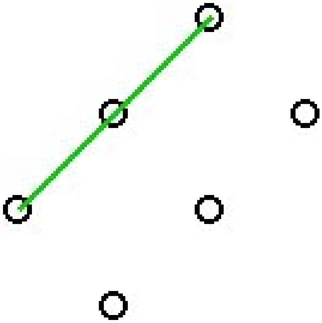
Space translation diversity (1/2)



Space translation diversity (1/2)



Space translation diversity (1/2)



Space translation diversity (2/2)

A_{iq}: gain between sensor i and source q
H_{pq}: transfer between reference and subarray p
S_{jq}: sample j of source q
β_q: path loss, $\mathbf{d}_q$: DOA, $\mathbf{b}_i$: sensor location

Tensor model (NB far field) *[Sidiropoulos'00]*

- Reference subarray: $A_{iq} = \beta_q \exp(\jmath \frac{\omega}{C} \mathbf{b}_i^\mathsf{T} \mathbf{d}_q)$
- Space translation (from reference subarray):
  $\beta_q \exp(\frac{\omega}{C} [\mathbf{b}_i + \mathbf{\Delta}_p]^\mathsf{T} \mathbf{d}_q) \stackrel{\text{def}}{=} A_{iq} H_{pq}$
- Trilinear model:

$$T_{ijp} = \sum_q A_{iq} S_{jq} H_{pq}$$

p: index of subarray

Space translation diversity (2/2)

A_{iq}: gain between sensor i and source q
H_{pq}: transfer between reference and subarray p
S_{jq}: sample j of source q
β_q: path loss, $\mathbf{d}_q$: DOA, $\mathbf{b}_i$: sensor location

Tensor model (NB far field) *[Sidiropoulos'00]*

- Reference subarray: $A_{iq} = \beta_q \exp(\jmath \frac{\omega}{C} \mathbf{b}_i^\mathsf{T} \mathbf{d}_q)$
- Space translation (from reference subarray):
  $\beta_q \exp(\frac{\omega}{C} [\mathbf{b}_i + \mathbf{\Delta}_p]^\mathsf{T} \mathbf{d}_q) \stackrel{\text{def}}{=} A_{iq} H_{pq}$
- Trilinear model:

$$T_{ijp} = \sum_q A_{iq} S_{jq} H_{pq}$$

p: index of subarray

Space translation diversity (2/2)

A_{iq}: gain between sensor i and source q
H_{pq}: transfer between reference and subarray p
S_{jq}: sample j of source q
β_q: path loss, $\mathbf{d}_q$: DOA, $\mathbf{b}_i$: sensor location

Tensor model (NB far field) *[Sidiropoulos'00]*

- Reference subarray: $A_{iq} = \beta_q \exp(\jmath \frac{\omega}{C} \mathbf{b}_i^\mathsf{T} \mathbf{d}_q)$
- Space translation (from reference subarray):
  $\beta_q \exp(\frac{\omega}{C} [\mathbf{b}_i + \boldsymbol{\Delta}_p]^\mathsf{T} \mathbf{d}_q) \stackrel{\text{def}}{=} A_{iq} H_{pq}$
- Trilinear model:

$$T_{ijp} = \sum_q A_{iq} S_{jq} H_{pq}$$

p: index of subarray

Meaning of coherence

1. Unconstrained joint source estimation & localization *[Sidiropoulos'2000]* (ill-posed if approximate)
2. Coherence-constrained joint source estimation & localization *[Lim-Comon'2010]*

   - *time diversity:* $\mu_A \rightarrow$ *cross correlation*
   - *space diversity:* μ_B, $\mu_C \rightarrow$ *angular separation*

Unadressed topics

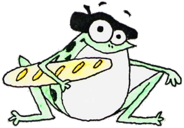
- spread spectrum communications
- brain inverse problems
- medical imaging (MRI)
- NL filtering
- noise reduction
- compression (Tensor trains...)
- probability
- hyperspectral imaging
- structured tensors
- convolutive mixtures
- nonnegative factors
- ...
- Algorithms

QUIZ: Why use tensors?

Main reason: essential uniqueness
- Identifiability recovery, *up to scale-permutation*

Sometimes: powerful *deterministic* approaches

Secondary reason: more sources with fewer sensors
- Matrices **A**, **B**, **C** may have more columns than rows



QUIZ: Why use tensors?

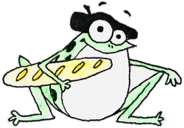
Main reason: essential uniqueness

- Identifiability recovery, *up to scale-permutation*

Sometimes: powerful *deterministic* approaches

Secondary reason: more sources with fewer sensors

- Matrices **A**, **B**, **C** may have more columns than rows



QUIZ: Why use tensors?

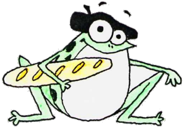
Main reason: essential uniqueness

- Identifiability recovery, *up to scale-permutation*

Sometimes: powerful *deterministic* approaches

Secondary reason: more sources with fewer sensors

- Matrices **A**, **B**, **C** may have more columns than rows



QUIZ: Why use tensors?

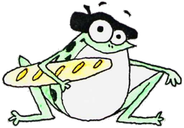
Main reason: essential uniqueness

- Identifiability recovery, *up to scale-permutation*

Sometimes: powerful *deterministic* approaches

Secondary reason: more sources with fewer sensors

- Matrices **A**, **B**, **C** may have more columns than rows

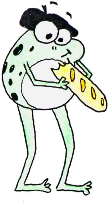


Other perspectives

Ignorance is the necessary condition for human being happiness.
Anatole France (1844-1924)

Other perspectives

Ignorance is the necessary condition for human being happiness.
Anatole France (1844-1924)



Only when the last tree has died, the last river has been poisoned and the last fish has been caught will we realize that we cannot eat money.
Cree proverb