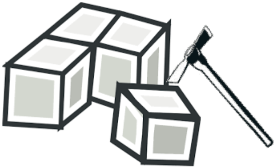
Environmental Multiway Data Mining



Jérémy E. Cohen

GIPSA-Lab, CNRS

05 September 2016

PhD Supervisor : Pierre Comon

Acknowledgements

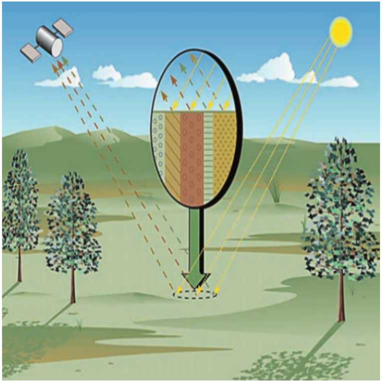


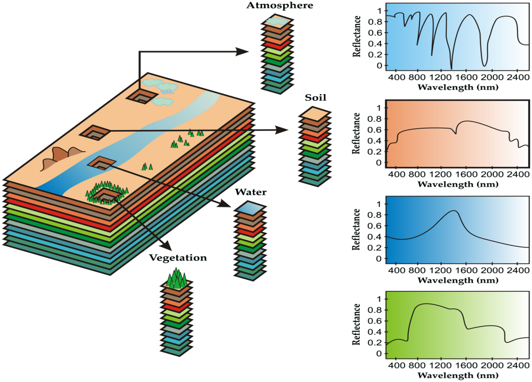




1. Introduction
   - Some environmental multiway data
   - Canonical Polyadic Decomposition
   - Challenges in environmental data mining
2. Compressed Constrained CPD
3. Multiway Data Fusion
4. Current Works
5. Conclusion

Hyperspectral imaging principle

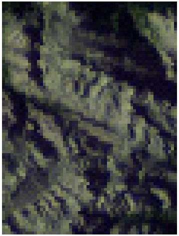




- Each image is a **mixture** of various materials.
- Each material has a unique spectral response.

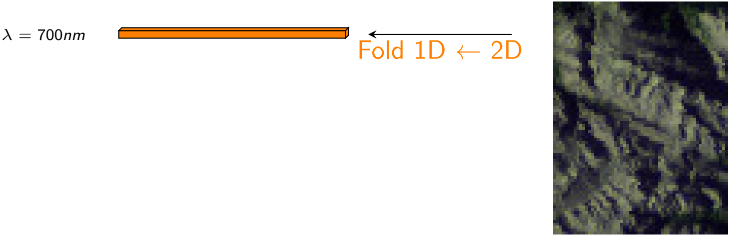
Credits for illustrations : Veganzones(left) and Bioucas(right)

Hyperspectral Data



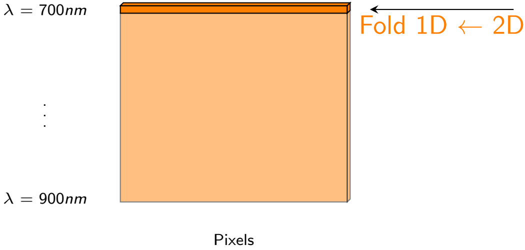
Snow in the Alps [Veganzones,2015]

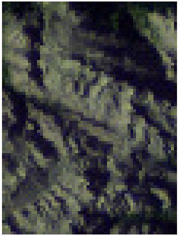
Hyperspectral Data



Snow in the Alps [Veganzones,2015]

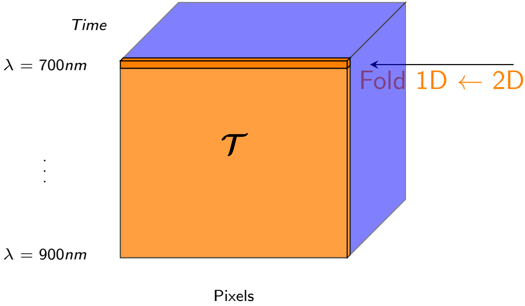
Hyperspectral Data

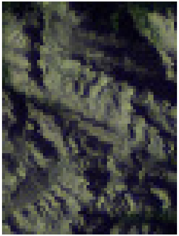




Snow in the Alps [Veganzones,2015]

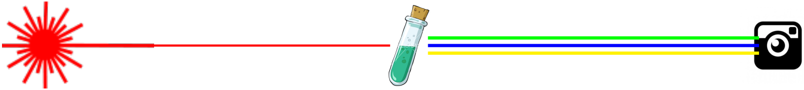
Hyperspectral Data

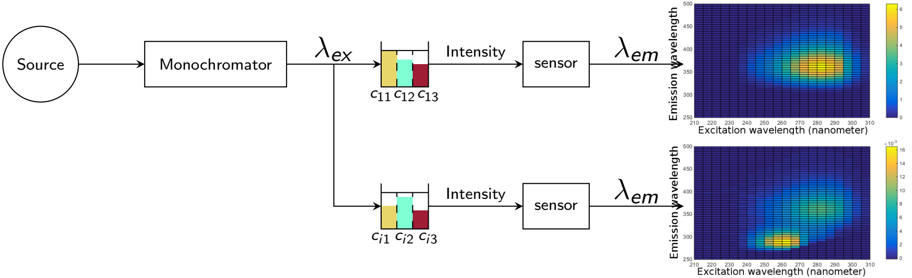




Snow in the Alps [Veganzones,2015]

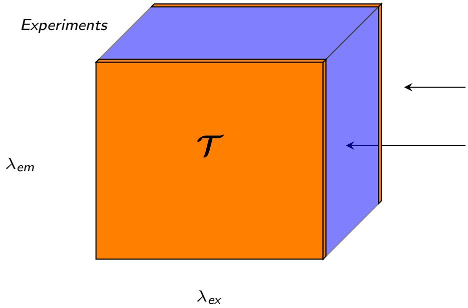
Another example : Fluorescence Spectroscopy

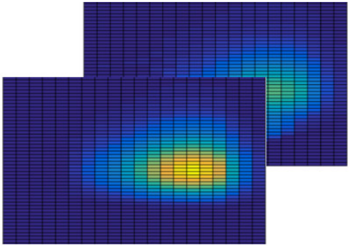




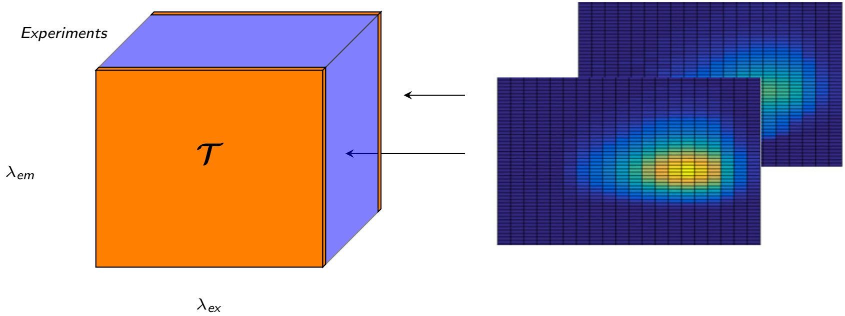
[Acar,2013]

Fluorescence Spectroscopy Data





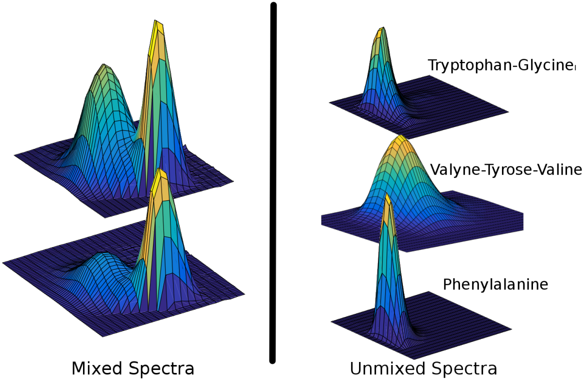
Fluorescence Spectroscopy Data



- Multiway arrays (tensors) appear naturally in data processing.
- Data are **mixtures** of signals unique to each material or chemical.

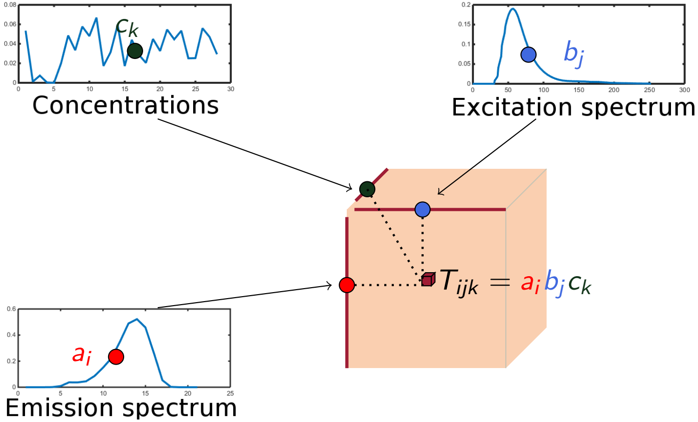
Main Issue : unmixing the data

How to extract meaningful information from multiway data ?



Unmixed data = rank 1 tensors

In some cases, meaningful information is contained in **simpler** tensors i.e. rank 1 tensors :



Main tool : Canonical Polyadic Decomposition

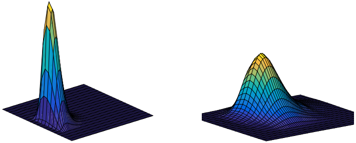
Canonical Polyadic Decomposition [Hitchcock,1927] aims at extracting all R components.

Tensor = first component $+ \cdots +$ Rth component

- Unmixing does not require additional knowledge
- Not applicable for 2-way arrays

Challenges in environmental and biomedical data mining

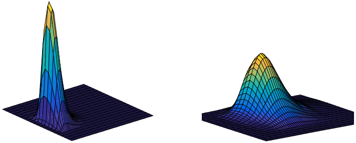
- Constrained Decompositions - Compressed Decompositions



$\rightarrow$ Nonnegative Large tensor

Challenges in environmental and biomedical data mining

- Constrained Decompositions - Compressed Decompositions

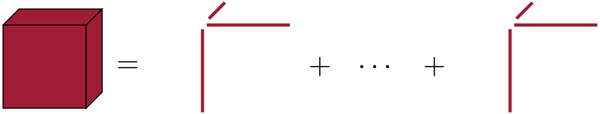


$\rightarrow$ Nonnegative Large tensor

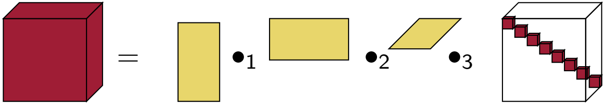
- Data Fusion

1. Introduction
2. Compressed Constrained CPD
   - Some definitions and properties
   - Compressed-based CPD
   - Nonnegative CPD
3. Multiway Data Fusion
4. Current Works
5. Conclusion

Some notations



$$\mathcal{T} = \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1 + \cdots + \mathbf{a}_R \otimes \mathbf{b}_R \otimes \mathbf{c}_R$$



$$\mathcal{T} = (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})\ \mathcal{I}_R$$

$\mathcal{T}$ has sizes $K \times L \times M$ $\qquad$ $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R]$ has sizes $K \times R$

$\otimes$ is the tensor product $\qquad$ $\bullet_i$ is the contraction on mode i

R is the rank of $\mathcal{T}$, *i.e.* smallest number of rank-one tensors spanning $\mathcal{T}$.

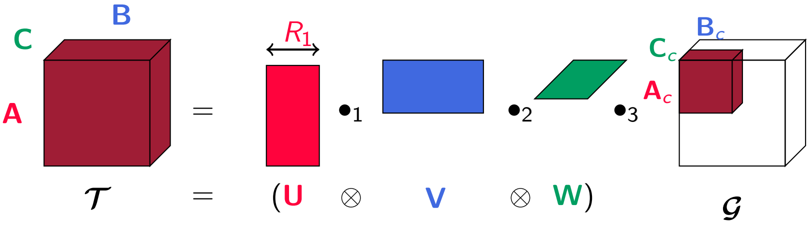
Some definitions and properties : multilinear transformation

Definition (Orthogonal Tucker Decomposition)

A tensor $\mathcal{T} \in \mathbb{R}^K \otimes \mathbb{R}^L \otimes \mathbb{R}^M$ can be expressed in an orthonormal basis $\mathbf{U} \otimes \mathbf{V} \otimes \mathbf{W}$ so that

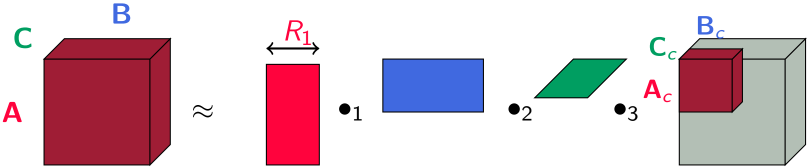
$$\mathcal{T} = (\mathbf{U} \otimes \mathbf{V} \otimes \mathbf{W})\,\mathcal{G}$$

where $\mathbf{U} \in K \times R_1$, $\mathbf{V} \in L \times R_2$, $\mathbf{W} \in M \times R_3$ and $R_1, R_2, R_3 \leq \text{rank}(\mathcal{T})$.



Tensor compression in the noisy case

In the noisy case : **(approximate) truncated HOSVD** **[DeLathauwer,2000]**



$$\mathcal{T} \approx (\mathbf{U} \otimes \mathbf{V} \otimes \mathbf{W}) \mathcal{G}$$

where $\hat{\mathbf{U}}\mathbf{N}_1 = TSVD(\mathbf{T}_{(1)})$ $\hat{\mathbf{V}}\mathbf{N}_2 = TSVD(\mathbf{T}_{(2)})$ $\hat{\mathbf{W}}\mathbf{N}_3 = TSVD(\mathbf{T}_{(3)})$

$$\widehat{\mathcal{T}} \approx (\widehat{\mathbf{U}} \otimes \widehat{\mathbf{V}} \otimes \widehat{\mathbf{W}})\widehat{\mathcal{G}} \text{ and } \widehat{\mathcal{G}} = (\mathbf{A}_c \otimes \mathbf{B}_c \otimes \mathbf{C}_c)\mathcal{I}_R + \mathcal{E}_c$$

Tensor CPD using compression

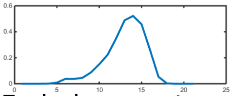
1. Compress $\mathcal{T}$ using any fast HOSVD
2. Decompose $\widehat{\mathcal{G}}$ to get $\widehat{\mathbf{A}}_c$, $\widehat{\mathbf{B}}_c$, $\widehat{\mathbf{C}}_c$
3. Decompress : $\widehat{\mathbf{A}} = \widehat{\mathbf{U}}\widehat{\mathbf{A}}_c$, $\widehat{\mathbf{B}} = \widehat{\mathbf{V}}\widehat{\mathbf{B}}_c$, $\widehat{\mathbf{C}} = \widehat{\mathbf{W}}\widehat{\mathbf{C}}_c$

Time in seconds for CP dec. $N \times N \times N$ - rank 5 random tensor :

N	10	50	100	200	300
Alternating Least Squares (ALS)	0.36	0.70	1.92	7.13	26.43
Compressed ALS	0.33	0.39	0.45	0.93	2.14

Nonnegativity constraints

In many applications : $\mathbf{A}, \mathbf{B}, \mathbf{C} \geqslant 0$



Emission spectrum

Compressed Nonnegative CPD :

$$\min_{\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c} \quad \|\widehat{\mathcal{G}} - (\mathbf{A}_c \otimes \mathbf{B}_c \otimes \mathbf{C}_c)\mathcal{I}_R\|_F^2$$
$$\text{sub. to} \quad \widehat{\mathbf{U}}\mathbf{A}_c, \widehat{\mathbf{V}}\mathbf{B}_c, \widehat{\mathbf{W}}\mathbf{C}_c \geqslant 0$$

Issue : Difficult exact projection on $\widehat{\mathbf{U}}\mathbf{A}_c \geq 0$

The uncompressed projective algorithm : ANLS

The cost function is minimized with respect to each factor alternatively :

$$\begin{array}{ll} \underset{\mathbf{A}}{\text{min.}} & \|\mathcal{T} - (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})\mathcal{I}_R\|_F^2 \\ \text{sub. to} & \mathbf{A} \geq 0 \end{array}$$

First, the unconstrained least squares update is computed :

$$\widehat{\mathbf{A}} = \mathbf{T}_1 \left(\mathbf{B} \odot \mathbf{C}\right)^{\dagger}$$

Then the least squares estimate is **projected** on the constraint space :

$$\widehat{\mathbf{A}} = \left[\widehat{\mathbf{A}}\right]^{+}$$

Approximate projection and PROCO-ALS

Approximate projection $\mathbf{\Pi}$:

Given Least Squares update $\widehat{\mathbf{A}}_c$

1. Decompression : $\widehat{\mathbf{A}} := \widehat{\mathbf{U}}\widehat{\mathbf{A}}_c$
2. Projection : $\widehat{\mathbf{A}} := \left[\widehat{\mathbf{A}}\right]^+$
3. Compression : $\widehat{\mathbf{A}}_c := \widehat{\mathbf{U}}^\mathsf{T}\widehat{\mathbf{A}}$

$$\mathbf{\Pi}\left[\widehat{\mathbf{A}}\right] = \mathbf{U}^\mathsf{T}\left[\mathbf{U}\mathbf{A}_c\right]^+$$

Projected and compressed framework (PROCO) [Cohen,2014]

Other possible algorithms and related problems

- PROCO-ALS [Cohen,2014], Compressed-AOADMM [Cohen,2016]

$$\begin{aligned} &\text{minimize } \|\widehat{\mathcal{G}} - (\mathbf{A}_c \otimes \mathbf{B}_c \otimes \mathbf{C}_c)\mathcal{I}_R\|_F^2 \\ &\text{w.r.t. } \mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c \\ &\text{s.t. } \widehat{\mathbf{U}}\mathbf{A}_c \succeq 0 \end{aligned}$$

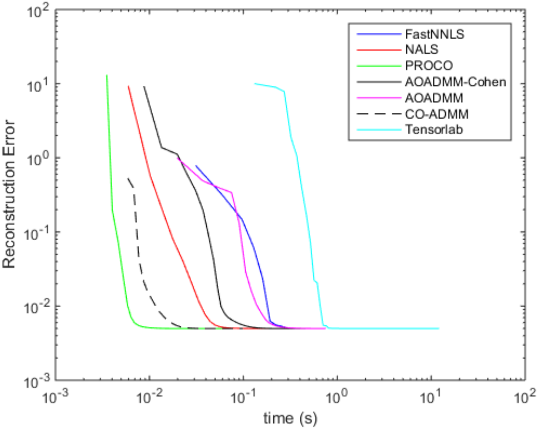
- Tensorlab 3.0 [Vervliet,2016]

$$\begin{aligned} &\text{minimize } \| \left(\widehat{\mathbf{U}} \otimes \widehat{\mathbf{V}} \otimes \widehat{\mathbf{W}}\right) \widehat{\mathcal{G}} - (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})\mathcal{I}_R\|_F^2 \\ &\text{w.r.t. } \mathbf{A}, \mathbf{B}, \mathbf{C} \\ &\text{s.t. } \mathbf{A} \succeq 0 \end{aligned}$$

- AOADMM [Huang,2015], FastNNLS [Bro,1997], ANLS

$$\begin{aligned} &\text{minimize } \|\mathcal{T} - (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})\mathcal{I}_R\|_F^2 \\ &\text{w.r.t. } \mathbf{A}, \mathbf{B}, \mathbf{C} \\ &\text{s.t. } \mathbf{A} \succeq 0 \end{aligned}$$

Simulated Data



Size : $100 \times 100 \times 100$
Rank : 5
Gaussian factors
R_i : $5 \times 5 \times 5$
SNR : 30dB
Gaussian i.i.d. noise

Experimental Data : Fluorescence Spectroscopy

Fluorescence spectroscopy data : excitation spectra, emission spectra, mixtures

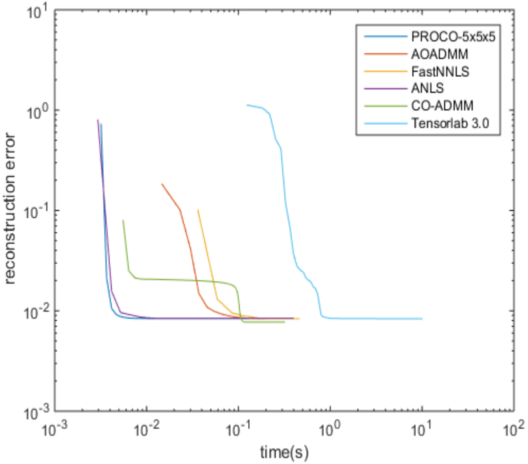
multimodal chemometrics data set from [Acar,2013]

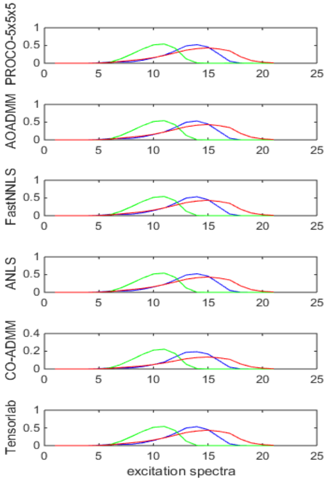
Description

5 compounds : Valine-Tyrosine-Valine (Val), Tryptophan- Glycine (Gly), Phenylalanine (Phe), Maltoheptaose (Mal) and Propanol (Pro)

Nb. of excitation wave lengths	21 (**A**)
Nb. of emission wave lengths	251 (**B**)
Nb. of Mixtures	28 (**C**)
Missing values	30% (replaced by zeros)

Experimental data : Results





1. Introduction

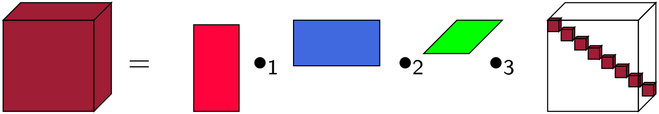
2. Compressed Constrained CPD

3. Multiway Data Fusion
   - Problem statement
   - Flexible data fusion
   - Experiments

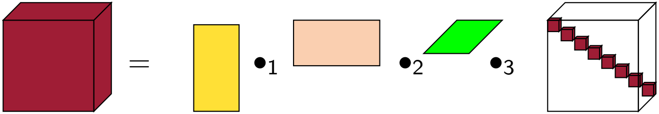
4. Current Works

5. Conclusion

Direct coupling [Harshman,1984]



$$\mathcal{T}_1 = (\mathbf{A}_1 \otimes \mathbf{B}_1 \otimes \mathbf{C})\ \mathcal{I}_R$$



$$\mathcal{T}_2 = (\mathbf{A}_2 \otimes \mathbf{B}_2 \otimes \mathbf{C})\ \mathcal{I}_R$$

Example : Fluorescence spectroscopy data and Nuclear Magnetic Resonance data

Direct coupling (2)

$$\forall\ i \in [1, N], \quad \begin{cases} \mathcal{T}_i & = (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R + \mathcal{E}_i \\ \mathbf{C}_i & = \mathbf{C}^* \end{cases}$$

If the noises $\mathcal{E}_i$ are Gaussian with i.i.d. entries, then the Maximum Likelihood Estimator (MLE) of the factors is

$$\underset{\mathbf{A}_i,\mathbf{B}_i,\mathbf{C}}{\text{argmin}} \sum_{i=1}^{N} \|\mathcal{T}_i - (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}^*)\mathcal{I}_R\|_F^2$$

For computation,

- CMTF by Acar et al.
- ALS by Cabral Farias, Cohen et al. (Tensor Package)
- Tensorlab 3.0 by Vervliet et al.

Other coupling models

- Parafac 2 [Harshman,1972] :

$$\forall i \in [1, N], \begin{cases} \mathcal{T}_i = (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}_i) \mathbf{\Sigma}_{\mathbf{i}R} + \mathcal{E}_i \\ \mathbf{C}_i = \mathbf{P}_i \mathbf{C}^* \\ \mathbf{P}_i^\top \mathbf{P}_i = \mathbf{I} \end{cases}$$

- Shift Parafac [Harshman,2003] :

$$\forall i \in [1, N], \begin{cases} \mathbf{M}_i = (\mathbf{A} \otimes \mathbf{B}_i) \mathbf{\Sigma}_{\mathbf{i}R} + \mathbf{E}_i \\ \mathbf{b}_r^{(i)} = \tau^{\delta_{ir}}(\mathbf{b}_r^*) \end{cases}$$

Other coupling models

- Parafac 2 [Harshman,1972] :

$$\forall i \in [1, N], \begin{cases} \mathcal{T}_i = (\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}_i)\, \boldsymbol{\Sigma}_{\mathbf{i}R} + \boldsymbol{\mathcal{E}}_i \\ \mathbf{C}_i = \mathbf{P}_i \mathbf{C}^* \\ \mathbf{P}_i^{\top} \mathbf{P}_i = \mathbf{I} \end{cases}$$

- Shift Parafac [Harshman,2003] :

$$\forall i \in [1, N], \begin{cases} \mathbf{M}_i = (\mathbf{A} \otimes \mathbf{B}_i)\, \boldsymbol{\Sigma}_{\mathbf{i}R} + \mathbf{E}_i \\ \mathbf{b}_r^{(i)} = \tau^{\delta_{ir}}(\mathbf{b}_r^*) \end{cases}$$

A more general framework ?

General Framework using a Bayesian approach [Cabral Farias, Cohen,2015]

- Parameters $\boldsymbol{\theta}_i = \begin{bmatrix} \text{vec}(\mathbf{A}_i) \\ \text{vec}(\mathbf{B}_i) \\ \text{vec}(\mathbf{C}_i) \end{bmatrix}$ are random

- Known prior distribution $p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)$ and likelihoods $p(\mathcal{Y}_i|\boldsymbol{\theta}_i)$

MAP estimation under conditionnal independance

$$\underset{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N}{\arg\max}\; p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N | \mathcal{Y}_1, \ldots, \mathcal{Y}_N) = \underset{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N}{\arg\min}\; \Upsilon(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)$$

$$\begin{aligned} \Upsilon(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N) &= -\sum_{i=1}^{N} \log p(\mathcal{Y}_{\mathbf{i}}|\boldsymbol{\theta}_i) - \log p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N) \\ &= \text{ data fitting terms } + \text{ coupling} \end{aligned}$$

Examples of flexible coupling models

Noisy exact coupling on $\mathbf{C}_i$

$$\forall\ i \in [1, N], \quad \begin{cases} \mathcal{T}_i & = (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R + \mathcal{E}_i \\ \mathbf{C}_i & = \mathbf{C}^* + \boldsymbol{\Gamma}_i \\ \boldsymbol{\Gamma}_i & \sim \mathcal{AN}\left(\mathbf{0}, \frac{1}{\sigma_{c,i}^2}\mathbf{I} \otimes \mathbf{I}\right) \end{cases}$$

$$\Upsilon(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N, \mathbf{C}^*) = -\sum_{i=1}^{N} \frac{1}{\sigma_1^2} \|\mathcal{T}_i - (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R\|_F^2 - \sum_{i=1}^{N} \frac{1}{\sigma_{ci}^2} \|\mathbf{C}_i - \mathbf{C}^*\|_F^2$$

Examples of flexible coupling models

Noisy exact coupling on $\mathbf{C}_i$

$$\forall\ i \in [1, N], \quad \begin{cases} \mathcal{T}_i & = (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R + \mathcal{E}_i \\ \mathbf{C}_i & = \mathbf{C}^* + \mathbf{\Gamma}_i \\ \mathbf{\Gamma}_i & \sim \mathcal{AN}\left(\mathbf{0}, \frac{1}{\sigma_{c,i}^2}\mathbf{I} \otimes \mathbf{I}\right) \end{cases}$$

$$\Upsilon(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N, \mathbf{C}^*) = -\sum_{i=1}^{N} \frac{1}{\sigma_1^2} \|\mathcal{T}_i - (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R\|_F^2 - \sum_{i=1}^{N} \frac{1}{\sigma_{ci}^2} \|\mathbf{C}_i - \mathbf{C}^*\|_F^2$$

Linear coupling on $\mathbf{C}_i$

$$\forall\ i \in [1, N], \quad \begin{cases} \mathcal{T}_i & = (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_R + \mathcal{E}_i \\ \mathbf{H}_i\mathbf{C}_i & = \mathbf{H}_j\mathbf{C}_j + \mathbf{\Gamma}_{ij} \\ \mathbf{\Gamma}_{ij} & \sim \mathcal{AN}\left(\mathbf{0}, \frac{1}{\sigma_{ij}^2}\mathbf{I} \otimes \mathbf{I}\right) \end{cases}$$

Optimization Strategies

Alternating Least Squares (ALS) based methods

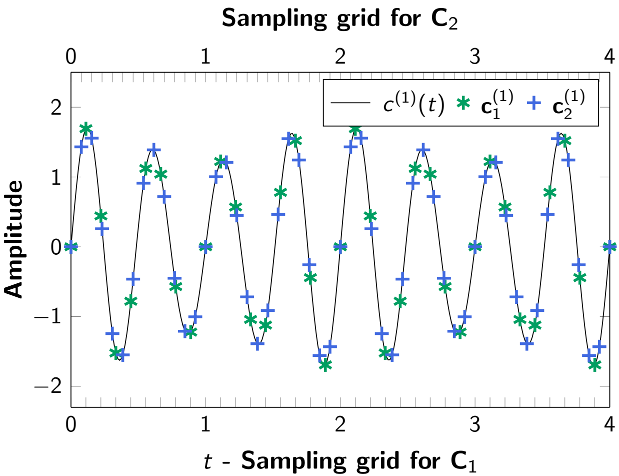
- Fast
- Simple to implement
- Tackle various coupling models
- Require warm initialization
- Unadapted for non-Gaussian distributions

Second-order methods (Tensorlab 3.0)

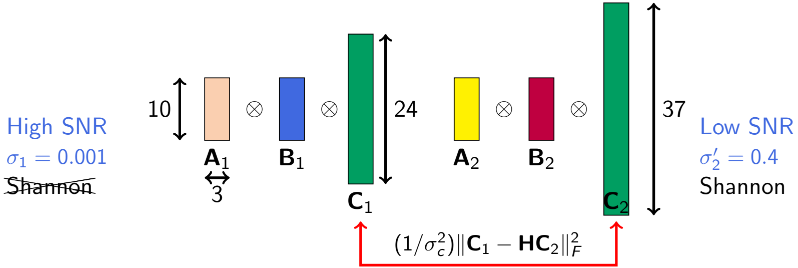
- Tackle a wide variety of models and noise distributions
- Not very sensitive to initialization
- Slow

Simulation : Resampling Bandlimited Signals

Bandlimited periodic signal



Simulation : Resampling Bandlimited Signals



Total MSE on the continuous functions (numerical integration)

	$\mathbf{C}_1$ ~~Shannon~~	$\mathbf{C}_2$ noisy
Uncoupled	33.4968	2.6581
Coupled	33.4968	1.0375

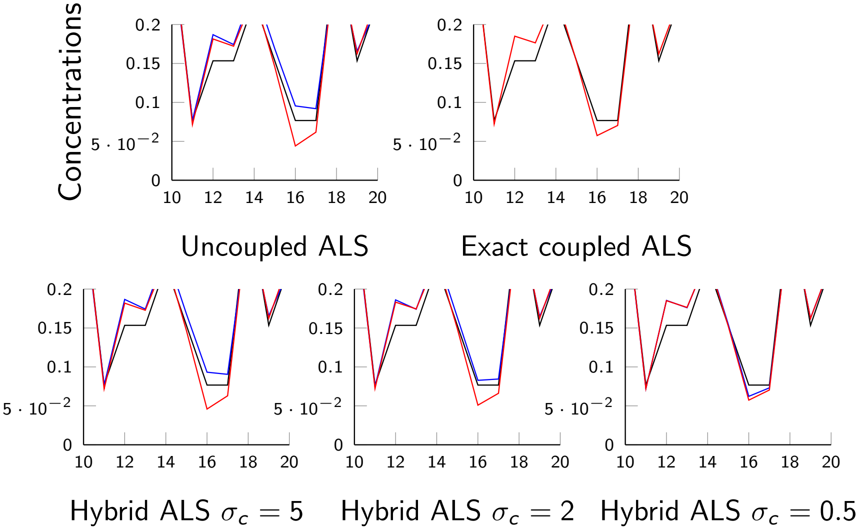
Experimental Data : NMR

Nuclear magnetic resonance data : chemical shifts, gradient levels, mixtures

same sample presented previously : **coupling through C**

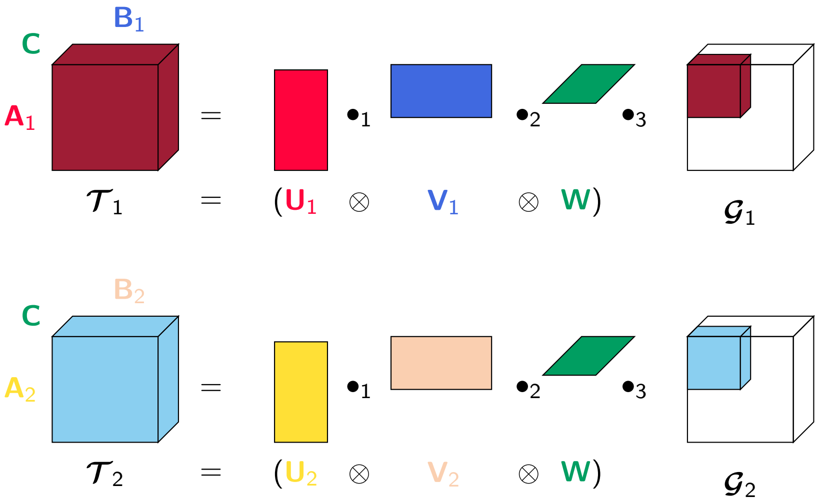
Nb. of chemical shifts	13324 (**A′**)
Nb. of gradient levels	8 (**B′**)
Nb. of Mixtures	28 (**C**)
Missing values	0% (replaced by zeros)

Results : Relative Concentrations of Phenylalanine

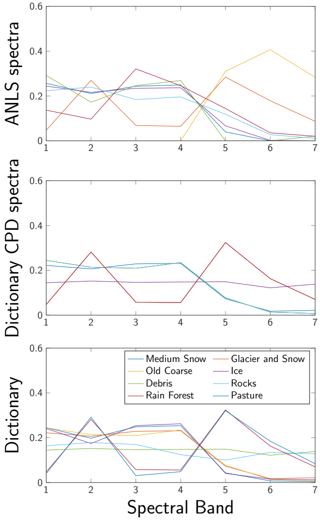


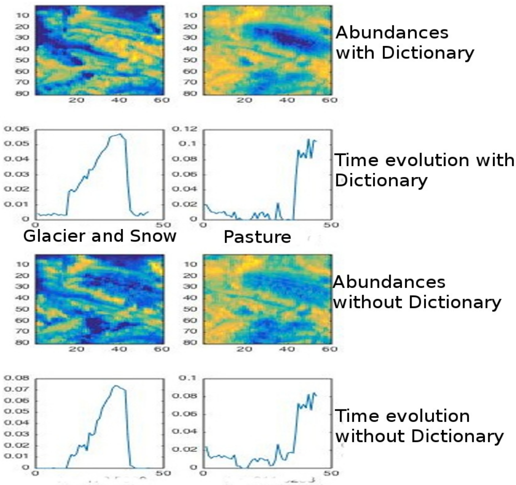
1. Introduction
2. Compressed Constrained CPD
3. Multiway Data Fusion
4. Current Works
5. Conclusion

Joint Compression



Dictionary-based CPD : Linear and Sparsity constraints





1. Introduction
2. Compressed Constrained CPD
3. Multiway Data Fusion
4. Current Works
5. Conclusion

Conclusions

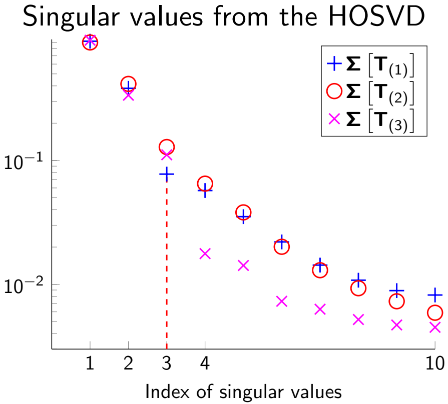
- Fast algorithm PROCO-ALS for constrained decompositions when projection is available. Need more theoretical results.
- Data fusion framework, first step towards non-trivial coupling models. More experimental data required.
- Among other research topics covered during the PhD : non-linear fluorescence tensor decomposition, notations, data fusion for Gaze-EEG data.

Selected list of publications

- J. E. Cohen, P. Comon, and X. Luciani, "Correcting inner filter effects, a non multilinear tensor decomposition method," *Chemometrics and Intelligent Laboratory Systems*, vol. 150, pp. 29–40, 2016.
- M. A. Veganzones, J. E. Cohen, R. Cabral Farias, J. Chanussot, and P. Comon, "Nonnegative tensor CP decomposition of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, Nov. 2015.
- J. E. Cohen, R. C. Farias, and P. Comon, "Fast decomposition of large nonnegative tensors," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 862–866, 2015.
- R. Cabral Farias, J. E. Cohen, C. Jutten, and P. Comon, "Joint decompositions with flexible couplings," in *12th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA)*, (Liberec, Czech Republic), Aug. 2015.

Thank you for your attention !

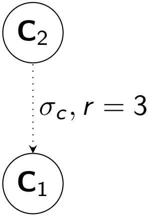
Rank and minimum compressed dimensions



Exproratory : it suggests a **Rank 2 or 3** : Mal and Pro are not fluorescent

Modeling

$$\begin{cases} \mathcal{T}_{EEM} = (\mathbf{A}_{EEM} \otimes \mathbf{B}_{EEM}\mathbf{C}_{EEM})\mathcal{I}_3 + \mathcal{E}_{EEM} \\ \mathcal{T}_{NMR} = (\mathbf{A}_{NMR} \otimes \mathbf{B}_{NMR}\mathbf{C}_{NMR})\mathcal{I}_5 + \mathcal{E}_{NMR} \\ \mathbf{C}_{EEM} = \mathbf{C}_{NMR}(r = 1, 2, 3) + \mathbf{\Gamma}_c \\ \|\mathbf{c}_i^{EEM}\|_1 = 1 \ \forall i \leq 3 \\ \mathcal{E}_{EEM} \sim \mathcal{AN}\left(\mathbf{0}, \mathbf{I}_{21} \otimes \mathbf{I}_{251} \otimes \mathbf{I}_{28}\right) \\ \mathcal{E}_{NMR} \sim \mathcal{AN}\left(\mathbf{0}, \frac{1}{\sigma_{NMR}^2}\mathbf{I}_8 \otimes \mathbf{I}_{13324} \otimes \mathbf{I}_{28}\right) \\ \mathbf{\Gamma}_c \sim \mathcal{AN}\left(\mathbf{0}, \frac{1}{\sigma_c^2}\mathbf{I}_{28} \otimes \mathbf{I}_3\right) \end{cases}$$



Challenges in environmental data mining

- Constrained Compression

$$\begin{cases} \mathcal{G} = (\mathbf{A}_c \otimes \mathbf{B}_c \otimes \mathbf{C}_c)\mathcal{I}_R + \mathcal{E}_c \\ \mathbf{W}\mathbf{C}_c \in \mathcal{S}_C \end{cases}$$

- Multiway Data Fusion

$$\begin{cases} \mathcal{T}_i & = (\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i)\mathcal{I}_{R_i} + \mathcal{E}_i \\ \mathbf{A}_i & = f_i^{(\mathbf{A})}(\mathbf{A}^*), \quad f_i^{(\mathbf{A})} \in \mathcal{F}^{(\mathbf{A})} \\ \mathbf{B}_i & = f_i^{(\mathbf{B})}(\mathbf{B}^*), \quad f_i^{(\mathbf{B})} \in \mathcal{F}^{(\mathbf{B})} \\ \mathbf{C}_i & = f_i^{(\mathbf{C})}(\mathbf{C}^*), \quad f_i^{(\mathbf{C})} \in \mathcal{F}^{(\mathbf{C})} \end{cases}$$

Some definitions and properties : tensor space and rank

Definition (Tensor space)

A tensor space $\mathcal{E} \otimes \mathcal{F}$ is the linear space obtained by mapping all bilinear maps on $\mathcal{E} \times \mathcal{F}$ to linear maps. It is unique up to isomorphisms.

Definition (Tensor and rank)

A real valued tensor $\boldsymbol{\mathcal{T}}$ is a vector of a tensor space $(\mathbb{R}^K \otimes \mathbb{R}^L \otimes \mathbb{R}^M, \otimes)$. The rank of $\boldsymbol{\mathcal{T}}$ is the minimal number of elements $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$ needed to express $\boldsymbol{\mathcal{T}}$. A tensor can be considered low rank when R is much smaller than the dimensions K, L, M.

When the tensor product $\otimes$ is cast as the outer product $\circ$, tensors can be considered as multiway arrays without loss of generality.

Results : Emission spectra and chemical shifts