Identifiability of component analysis via tensor eigenvectors

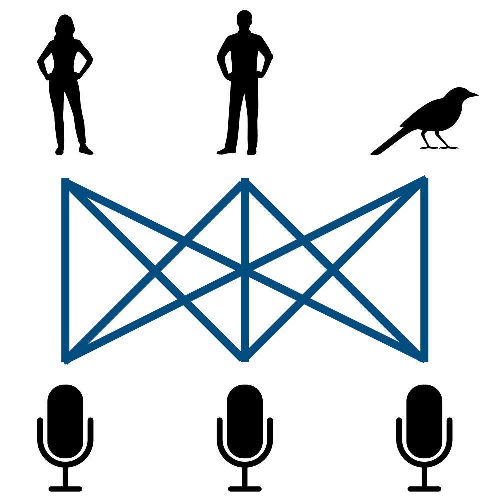
Álvaro Ribot, Anna Seigal and Piotr Zwiernik

New Directions in Algebraic Statistics
IMSI, Chicago

July 22, 2025

Component analysis

- Setup: $Y = AZ$
  - $A \in \mathbb{R}^{n \times n}$ is a fixed invertible matrix
  - Y is an observed random vector
  - Z is a latent random vector
- **Goal:** Recover A and Z, given samples from Y



Theorem[1]: If Z is independent and at most one Z_i is Gaussian, then A is identifiable (up to row permutation and sign flip)

Proof idea: reduce problem to decomposing a cumulant tensor

Can we relax the independence assumption and still get identifiability? [2]

[1] Comon, 1994. Independent component analysis, a new concept?
[2] Mesters, Zwiernik, 2024. Non-independent component analysis.

Principal component analysis (PCA)

- Setup: $Y = AZ$
  - $A \in \mathbb{R}^{n \times n}$ is a fixed **orthogonal** matrix: $A^\top A = I$
  - Y is an observed random vector
  - Z is a latent random vector of **uncorrelated** sources: $\mathrm{Cov}(Z)$ diagonal
- **Goal:** Recover A and Z, given samples from Y

Solution: Spectral decomposition of $\mathrm{Cov}(Y) = V \Lambda V^\top \implies A = V$

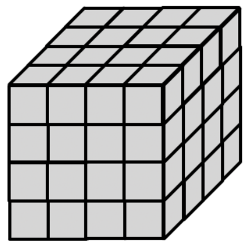
- **Whitening:** $Y_w = \Lambda^{-\frac{1}{2}} V^\top (Y - \mathbb{E}(Y)) \implies \mathbb{E}(Y_w) = 0, \mathrm{Cov}(Y_w) = I$

Corollary: In $Y = AZ$, we can assume that Y and Z are whitened

$$\implies I = \mathrm{Cov}(Y) = A\ \mathrm{Cov}(Z)\ A^\top = AA^\top$$

so we can restrict our search space to $\mathrm{O}(n) = \{Q \in \mathbb{R}^{n \times n} \mid QQ^\top = I\}$

Cumulant tensors



- A **tensor** is a multidimensional array $T \in (\mathbb{R}^n)^{\otimes d} \cong \mathbb{R}^{n \times \cdots \times n}$
  - $d = 2$: matrix $M \in \mathbb{R}^{n \times n}$
- A tensor is **symmetric** if $T_{i_1,\ldots,i_d} = T_{i_{\sigma(1)},\ldots,i_{\sigma(d)}}$ for all $\sigma \in S_d$
  - $d = 2: \;\; M_{ij} = M_{ji} \iff M = M^\top$
  - $d = 3: \;\; T_{ijk} = T_{ikj} = T_{jik} = T_{jki} = T_{kij} = T_{kji}$
- Let $S^d(\mathbb{R}^n) \subset (\mathbb{R}^n)^{\otimes d}$ denote the space of symmetric tensors
- $\mathrm{O}(n)$ acts on $S^d(\mathbb{R}^n)$ as $[Q \bullet T]_{i_1,\ldots,i_d} = \sum_{j_1,\ldots,j_d} Q_{i_1 j_1} \cdots Q_{i_d j_d} T_{j_1,\ldots,j_d}$
  - $d = 2: \;\; Q \bullet M = QMQ^\top$

- The cumulant-generating function of a random vector Z is $\kappa_Z(t) = \log \mathbb{E}(e^{t^\top Z})$
- The d-th **cumulant tensor** of Z is $K_d(Z) \in S^d(\mathbb{R}^n)$ given by

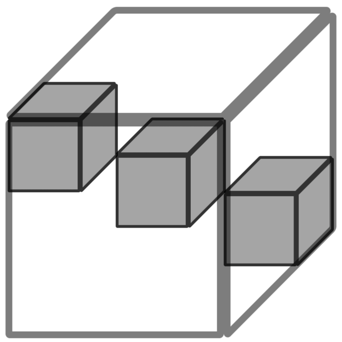
$$K_d(Z)_{i_1,\ldots,i_d} = \frac{\partial^d}{\partial t_{i_1} \cdots \partial t_{i_d}} \kappa_Z(t) \Big|_{t=0}$$

- Fact: $K_d(QZ) = Q \bullet K_d(Z)$

Independent component analysis (ICA)

Theorem[1]: $Y = AZ$. If Z is independent and at most one Z_i is Gaussian, then A is identifiable from Y (up to row permutation and sign flip)

Proof sketch:



- $Z = (Z_1, \ldots, Z_n)$ independent $\iff K_d(Z)$ is **diagonal** for all d
  - i.e. $K_d(Z)_{i_1,\ldots,i_d} \neq 0 \implies i_1 = \cdots = i_d$
  - More precisely, $K_d(Z) = K_d(Z_1)\, e_1^{\otimes d} + \cdots + K_d(Z_n)\, e_n^{\otimes d}$
- A symmetric tensor $T \in S^d(\mathbb{R}^n)$ is **orthogonally decomposable** (**odeco**) if

$$T = \lambda_1 a_1^{\otimes d} + \cdots + \lambda_n a_n^{\otimes d} \qquad \text{s.t.} \quad a_i \perp a_j \text{ for all } i \neq j$$

$$= A \bullet (\lambda_1 e_1^{\otimes d} + \cdots + \lambda_n e_n^{\otimes d}) \quad \text{s.t.} \quad A \in \mathrm{O}(n)$$

- $Y = AZ$ with $A \in \mathrm{O}(n) \implies K_d(Y) = A \bullet K_d(Z)$ is odeco for all d
- **Theorem**[3]: Odeco tensors have a unique odeco decomposition
- **Theorem**[4]: $\exists\, d_0$ s.t. $K_d(Z_i) = 0\ \forall\, d \geq d_0 \implies Z_i$ is Gaussian

[1] Comon, 1994. Independent component analysis, a new concept?
[3] Anandkumar et al, 2014. Tensor decompositions for learning latent variable models.
[4] Marcinkiewicz, 1939. Sur une propriété de la loi de Gauß.

Pairwise mean independent component analysis

Can we relax the independence assumption in ICA and still get identifiability?

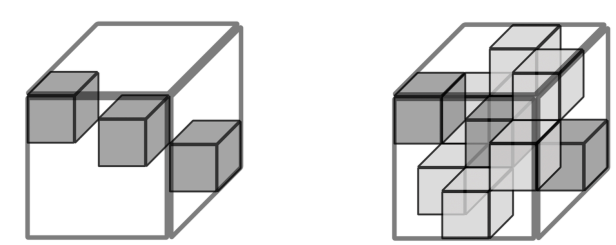
- Z is **pairwise mean independent** (**PMI**) if $\mathbb{E}(Z_i \mid Z_j) = \mathbb{E}(Z_i)$ for all $i \neq j$ [5]

Theorem[**R**SZ25+]: $Y = AZ$. If Z is *pairwise mean independent* and sufficiently general, then A is identifiable from Y (up to row permutation and sign flip)

- Conjectured in [2].
- **Proof idea**: reduce problem to decomposing a cumulant tensor with an *orthogonal basis of eigenvectors*
- Recall: $Z = (Z_1, \ldots, Z_n)$ independent $\iff K_d(Z) \in V_{\text{diag}}$ for all d
- $Z = (Z_1, \ldots, Z_n)$ PMI $\iff K_d(Z) \in V$ for all d, where

  $V = \{T \in S^d(\mathbb{R}^n) \mid T_{i,j,j,\ldots,j} = 0 \text{ for all } i \neq j\}$

- **Remark**: independent $\Longrightarrow$ PMI, i.e. $V_{\text{diag}} \subseteq V$



(A) Tensors in V_{diag} (B) Tensors in V

[2] Mesters, Zwiernik, 2024. Non-independent component analysis.
[5] Wooldridge, 2010. Econometric analysis of cross section and panel data.

Eigenvectors of tensors

- A unit vector $v \in \mathbb{R}^n$ is an **eigenvector** of $T \in S^d(\mathbb{R}^n)$ with eigenvalue $\lambda \in \mathbb{R}$ if

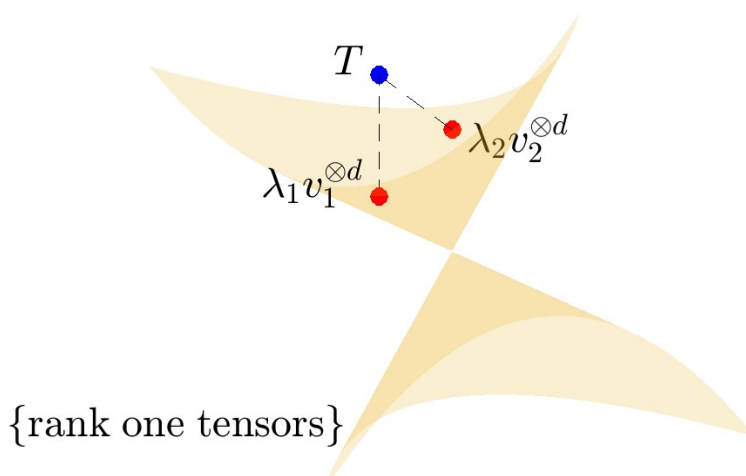
$$T(\,\cdot\,, v, \ldots, v) = \lambda v$$

where

$$T(\,\cdot\,, v, \ldots, v)_i = \sum_{j_2,\ldots,j_d} T_{i,j_2,\ldots,j_d} v_{j_2} \cdots v_{j_d}$$

  - $d = 2 : Mv = \lambda v$

- **Critical rank-one approximation**: critical point of $d_T(v) = \|T - v^{\otimes d}\|^2$



eigenvector

$\updownarrow$

critical rank-one approximation

Example: Odeco tensor $T = \lambda_1 q_1^{\otimes d} + \cdots + \lambda_n q_n^{\otimes d}$ with $q_i \perp q_j$ and $\|q_i\| = 1$.
Then, each q_i is an eigenvector of T with eigenvalue λ_i (there are more [6])

[6] Robeva, 2016. Orthogonal decomposition of symmetric tensors.

Orthogonal eigenvectors of tensors

Spectral theorem: every $M \in S^2(\mathbb{R}^n)$ is odeco, i.e., it can be decomposed as

$$M = Q\Lambda Q^\top = \lambda_1 q_1^{\otimes 2} + \cdots + \lambda_n q_n^{\otimes 2} \quad \text{s.t. } \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),\ Q \in \mathrm{O}(n)$$

If M is generic ($\lambda_i \neq \lambda_j$), this decomposition is unique (up to sign and permutation)

How does this generalize to higher-order tensors $T \in S^d(\mathbb{R}^n)$?

Prop: The set of tensors in $S^d(\mathbb{R}^n)$ with an orthogonal basis of eigenvectors is

$$X := \mathrm{O}(n) \bullet V = \{Q \bullet T \mid Q \in \mathrm{O}(n), T \in V\},$$

where $V = \{T \in S^d(\mathbb{R}^n) \mid T_{i,j,j,\ldots,j} = 0 \text{ for all } i \neq j\}$

Proof sketch:

- $e_1, \ldots, e_n$ are eigenvectors of $T \in S^d(\mathbb{R}^n) \iff T \in V$
- Let $Q \in \mathrm{O}(n)$, v is an eigenvector of $T \iff Qv$ is an eigenvector of $Q \bullet T$

Orthogonal eigenvectors of tensors

Spectral theorem: every $M \in S^2(\mathbb{R}^n)$ is odeco, i.e., it can be decomposed as

$$M = Q\Lambda Q^\top = \lambda_1 q_1^{\otimes 2} + \cdots + \lambda_n q_n^{\otimes 2} \quad \text{s.t. } \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),\ Q \in \mathrm{O}(n)$$

If M is generic ($\lambda_i \neq \lambda_j$), this decomposition is unique (up to sign and permutation)

How does this generalize to higher-order tensors $T \in S^d(\mathbb{R}^n)$?

Theorem[**R**SZ25]: If $T \in S^d(\mathbb{R}^n)$ is a generic tensor with an orthonormal basis of eigenvectors $v_1, \ldots, v_n$, then this basis is unique (up to sign).

Proof idea:

- Want to show: given $T \in V \subseteq S^d(\mathbb{R}^n)$ generic and $Q \in \mathrm{O}(n)$,

$$Q \bullet T \in V \iff Q \in \mathrm{SP}(n) = \{\mathrm{diag}(\pm 1, \ldots, \pm 1)P \mid P \text{ permutation}\}$$

- $n = 2$: $\mathrm{codim}((Q \bullet V) \cap V, V)$ is high enough for $Q \in \mathrm{O}(n) \backslash \mathrm{SP}(2)$
- $n \geq 3$: use normal form of orthogonal transformation and the case $n = 2$

[**R**SZ25] Ribot, Seigal, Zwiernik, 2025. Orthogonal eigenvectors and singular vectors of tensors

Summary: Identifiability of PMICA

Theorem[**R**SZ25+]: $Y = AZ$. If Z is *pairwise mean independent* and sufficiently general, then A is identifiable from Y (up to row permutation and sign flip)

- Z is **pairwise mean independent** (**PMI**) if $\mathbb{E}(Z_i \mid Z_j) = \mathbb{E}(Z_i)$ for all $i \neq j$

Proof sketch:

- $Z = (Z_1, \ldots, Z_n)$ PMI $\iff K_d(Z) \in V$ for all d
- The set of tensors with an orthogonal basis of eigenvectors is $X = \mathrm{O}(n) \bullet V$
- Orthogonal basis of eigenvectors are generically unique

ICA vs PMICA

ICA $\leftrightarrow K_d(Y) \in X_{\text{odeco}} = \mathrm{O}(n) \bullet V_{\text{diag}} \subset S^d(\mathrm{R}^n)$, and $\dim X_{\text{odeco}} = \binom{n+1}{2}$

PMICA $\leftrightarrow K_d(Y) \in X = \mathrm{O}(n) \bullet V \subset S^d(\mathrm{R}^n)$, and $\dim X = \binom{n+d-1}{d} - \binom{n}{2}$

Minimality of pairwise mean independence

- **Q:** Can we relax the independence assumption in ICA and still get identifiability?
- **A:** YES! It is enough to assume pairwise mean independence (PMI)
- **Q:** Can we relax the PMI assumption in PMICA and still get identifiability?

Theorem[**R**SZ25+]: $Y = AZ$. If Z_i is *mean independent* of Z_j for all $i \neq j$ except for $(i, j) = (1,2)$, then A is generically unidentifiable from $K_d(Y)$ for any $d \geq 3$

Remark: $\mathbb{E}(Z_2 \mid Z_1) = \mathbb{E}(Z_2)$ does **not** imply $\mathbb{E}(Z_1 \mid Z_2) = \mathbb{E}(Z_1)$

Alternative reformulation:

- Let $d \geq 3$ and a linear space $V \subsetneq W \subseteq S^d(\mathbb{R}^n)$ given by zero restrictions.
- Let $T \in W$ be generic.
- Then, there exists $Q \in \mathrm{O}(n) \backslash \mathrm{SP}(n)$ such that $Q \bullet T \in W$

Conjecture: the "given by zero restrictions" assumption can be dropped

What does it mean to be generic?

When does a tensor T have a unique orthogonal basis of eigenvectors?

Theorem[**R**SZ25+]: Let $T \in V \subset S^d(\mathbb{R}^n)$

- $d = 2$: T is generic $\iff$ distinct diagonal entries
- $d = 3$: T is generic $\iff$ at most one diagonal entry is zero
- $d = 4$: T is generic $\iff$ distinct diagonal entries
- $d = 5$: T is generic $\iff$ does not lie on a quadric

- Diagonal entries of cumulant tensors: $K_d(Z)_{i,\ldots,i} = K_d(Z_i)$
- In practice, PMI assumption may be relaxed if we only use $K_4(Y)$

Corollary: $Y = AZ$. If Z satisfies that $K_4(Z) \in V$ and $K_4(Z_i) \neq K_4(Z_j)$ for all $i \neq j$, then A is identifiable from $K_4(Y)$ (up to sign and permutation)

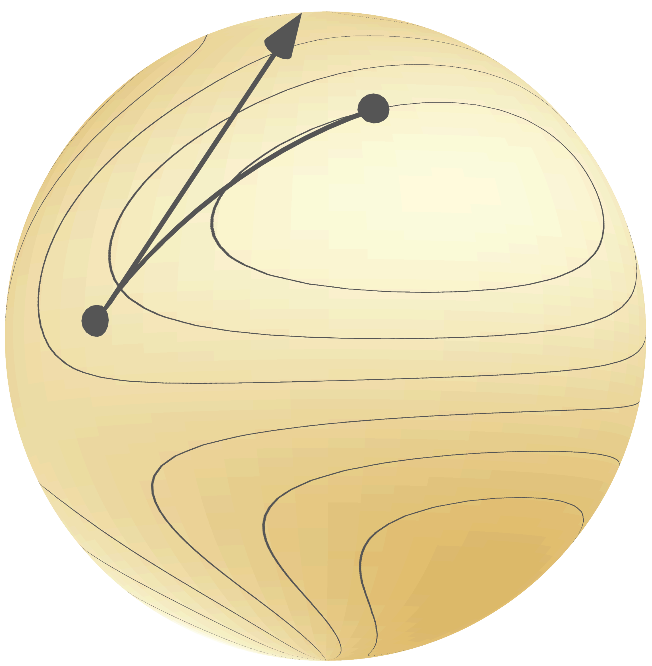
Computing PMI components

- Data analysis setup:
  - Model $Y = AZ$ with $A \in \mathrm{O}(n)$
  - We get samples $y_1, \ldots, y_m$
  - We estimate the d-th order cumulant tensor $\hat{K}_d \approx K_d(Y)$
  - **Goal:** Find $A \in \mathrm{O}(n)$ so that $A^\top \bullet \hat{K}_d$ lies (approximately) in V

Algorithm (RGD_PMICA):
- Input: tensor $\hat{K}_d \in S^d(\mathbb{R}^n)$
- Use Riemannian gradient descent to solve

$$\min_{A \in \mathrm{O}(n)} \ \mathrm{dist}(A^\top \bullet \hat{K}_d, V)^2$$

- Output: matrix $A \in \mathrm{O}(n)$



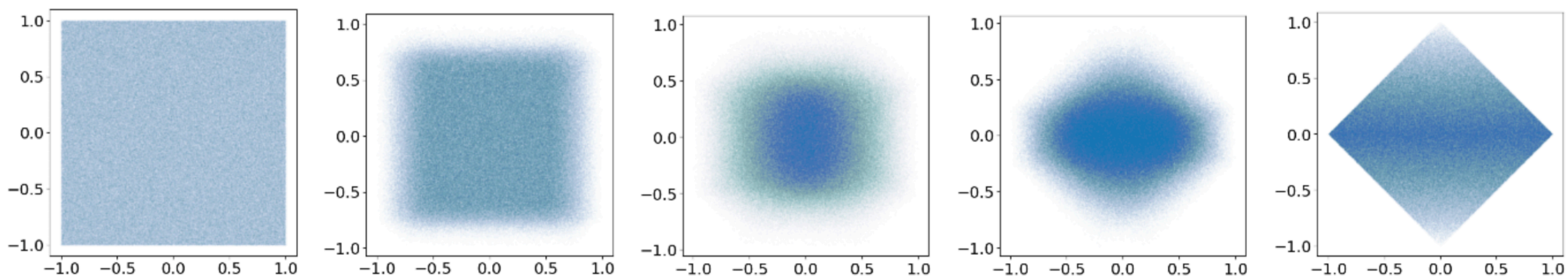
Pymanopt [7]

Remark: $\mathrm{dist}(T, V)^2 = \sum_{i \neq j} T^2_{i,j,\ldots,j}$

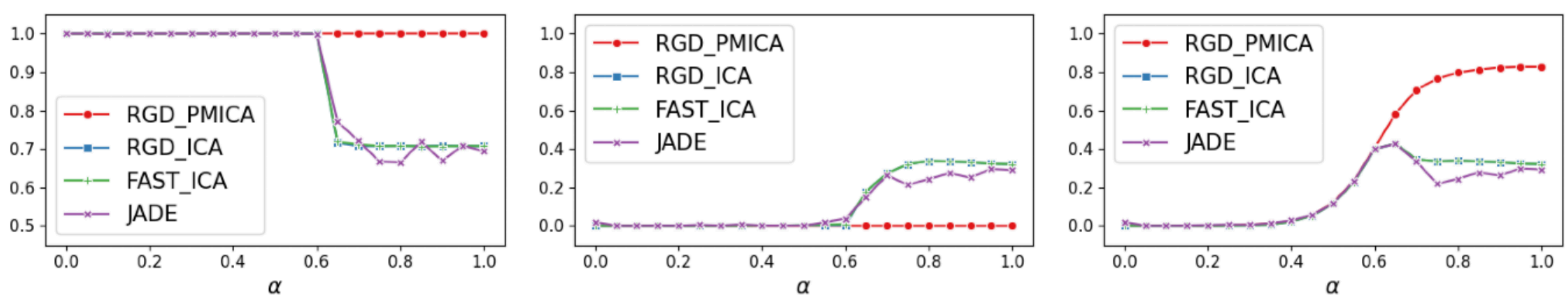
[7] Townsend et al, 2016. Pymanopt: A Python Toolbox for Optimization on Manifolds using Automatic Differentiation

Synthetic data experiments

- $n = 2$
- $Y^{(\alpha)} = AZ^{(\alpha)}$ for $\alpha \in (0,1)$, where $Z^{(\alpha)} = (1-\alpha)Z^{(0)} + \alpha Z^{(1)}$
- Use RGD_PMICA and RGD_ICA to estimate A from $\hat{K}_4(Y)$
- Compare with classical methods: FastICA[8] and JADE[9]



(A) $Z^{(0)}$ (B) $Z^{(0.25)}$ (C) $Z^{(0.5)}$ (D) $Z^{(0.75)}$ (E) $Z^{(1)}$



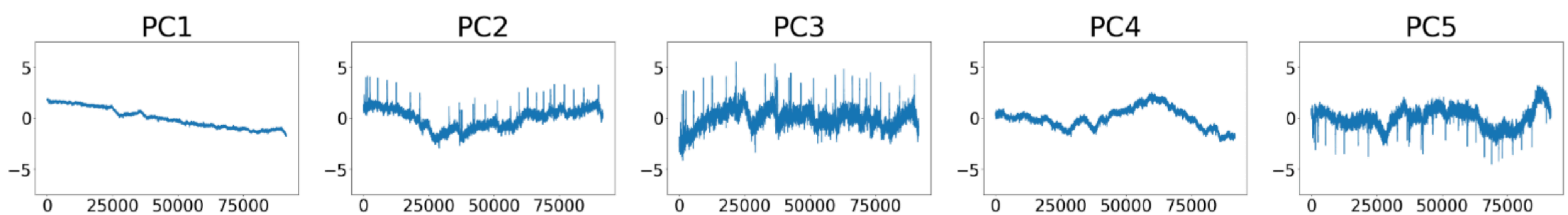
(A) $\mathrm{sim}(A, \hat{A})$ (B) Distance to PMI (C) Distance to independent

[8] Hyvärinen, Aapo, and Erkki Oja. Independent component analysis: algorithms and applications.
[9] Cardoso, Souloumiac, 1993. Blind beamforming for non-Gaussian signals.

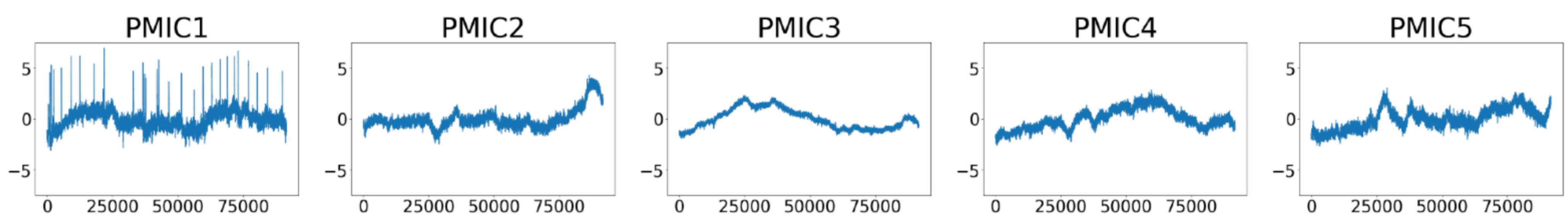
Real data experiments

- Fisher’s Iris flower dataset
- Fama-Fench Data Library: 25 Portfolios
- Chavarriaga’s Electroencephalogram (EEG) dataset

Tensor	Format	Distance to X	Distance to X_{odeco}
Iris	$4 \times 4 \times 4 \times 4$	0.23	0.67
Portfolios	$25 \times 25 \times 25 \times 25$	0.036	0.36
EEG	$64 \times 64 \times 64$	0.068	0.60



(A) Projection onto principal components



(B) Projection onto pairwise mean independent components

Takeaways and open questions

- Generic identifiability of pairwise mean independent component analysis
- Generic uniqueness of orthogonal basis of eigenvectors for symmetric tensors
- PMICA model is much larger than the ICA (odeco) model

- Future directions:
  - Understanding of the variety $X = \mathrm{O}(n) \bullet V$ (implicit description, degree)
  - Genericity conditions for all d
  - Algorithmic approaches for PMICA (maybe for some parametric family)
  - All of these for tensors in $\mathbb{R}^{n_1 \times \cdots \times n_d}$ with orthogonal singular vector tuples

Acknowledgments

- My coauthors: Anna Seigal and Piotr Zwiernik

Harvard John A. Paulson School of Engineering and Applied Sciences

"la Caixa" Foundation

RCC.HARVARD.EDU RCC

References

(1) Pierre Comon. Independent component analysis, a new concept? *Signal processing*,36(3):287–314, 1994.

(2) Geert Mesters and Piotr Zwiernik. Non-independent components analysis. *The Annals of Statistics*, 52(6):2506–2528, 2024.

(3) Anandkumar, Animashree, Rong Ge, Daniel J. Hsu, Sham M. Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *J. Mach. Learn. Res.* 15, no. 1 (2014): 2773-2832.

(4) Józef Marcinkiewicz. Sur une propriété de la loi de Gauß. *Mathematische Zeitschrift*, 44:612–618, 1939.

(5) Jeffrey Wooldridge. Econometric analysis of cross section and panel data. *MIT press*, 2010.

(6) Elina Robeva. Orthogonal decomposition of symmetric tensors. *SIAM Journal on Matrix Analysis and Applications* 37, no. 1 (2016): 86-102.

(7) James Townsend, Koep Niklas, and Sebastian Weichwald. Pymanopt: A python toolbox for optimization on manifolds using automatic differentiation. *Journal of Machine Learning Research* 17, no. 137 (2016): 1-5.

(8) Aapo Hyvärinen, and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks* 13, no. 4-5 (2000): 411-430.

(9) Jean-François Cardoso and Antoine Souloumiac. Blind beamforming for non-Gaussian signals. In *IEE proceedings F (radar and signal processing)*, volume 140, pages 362–370. IET, 1993.

[**R**SZ25] Alvaro Ribot, Anna Seigal, and Piotr Zwiernik. Orthogonal eigenvectors and singular vectors of tensors. *arXiv:2506.19009* (2025).