Block Component Analysis

Lieven De Lathauwer

K.U.Leuven
Belgium
Lieven.DeLathauwer@kuleuven-kortrijk.be
Lieven.DeLathauwer@esat.kuleuven.be

Overview

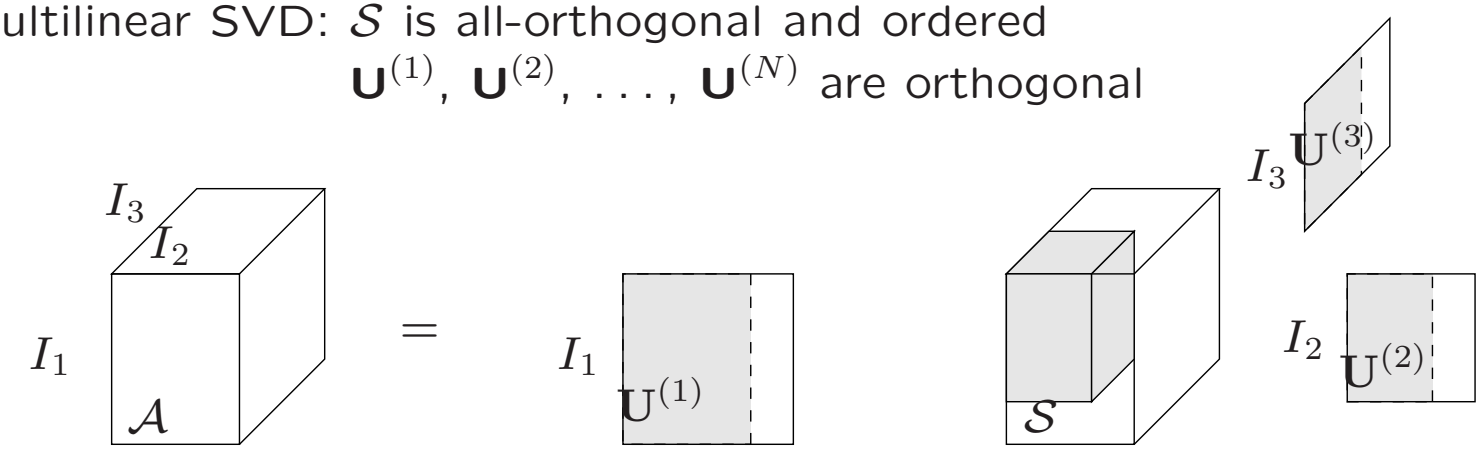
- Algebraic preliminaries
  - Multilinear rank and associated decomposition
  - Rank and associated decomposition
- Factor analysis and blind source separation
- Block Term Decompositions
- Block Component Analysis
- Variants
- Conclusion

Multilinear rank and associated decomposition

Multilinear rank: (dimension column space, dimension row space, . . .)
[*Hitchcock, 1927*]

Decomposition: $\mathcal{A} = \mathcal{S} \cdot_1 \mathbf{U}^{(1)} \cdot_2 \mathbf{U}^{(2)} \cdot_3 \ldots \cdot_N \mathbf{U}^{(N)}$

In multilinear SVD: $\mathcal{S}$ is all-orthogonal and ordered
$\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, . . . , $\mathbf{U}^{(N)}$ are orthogonal



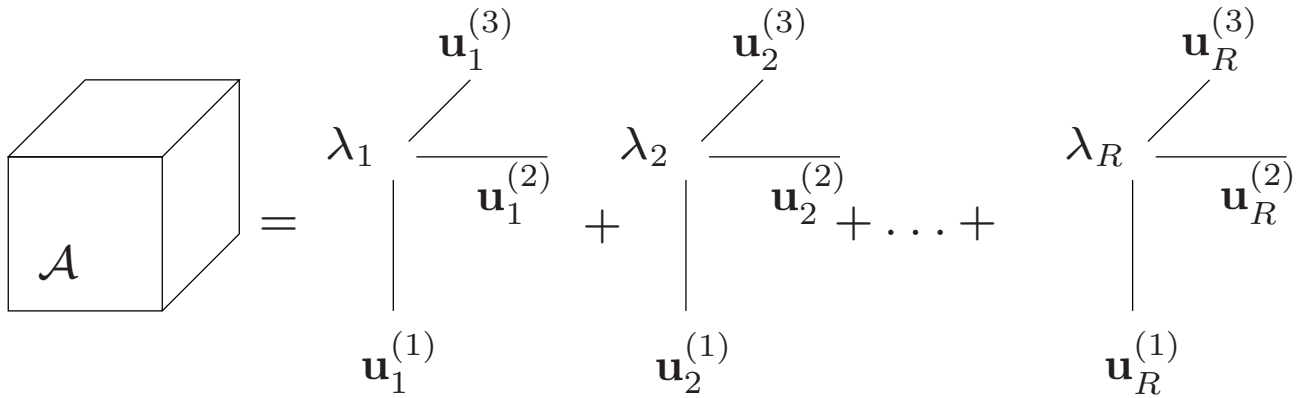
[*Tucker '64*], [*De Lathauwer '00*]

Cf. subspace variety

Rank and associated decomposition

Rank: minimal number of rank-1 terms [*Hitchcock, 1927*]

Canonical Polyadic Decomposition (CPD) of a tensor $\mathcal{A}$ is its decomposition in a minimal sum of rank-1 tensors



[*Harshman '70*], [*Carroll and Chang '70*]

Unique under mild conditions

Cf. R-th secant variety of Segre variety (unsymmetric), Veronese variety (symmetric)

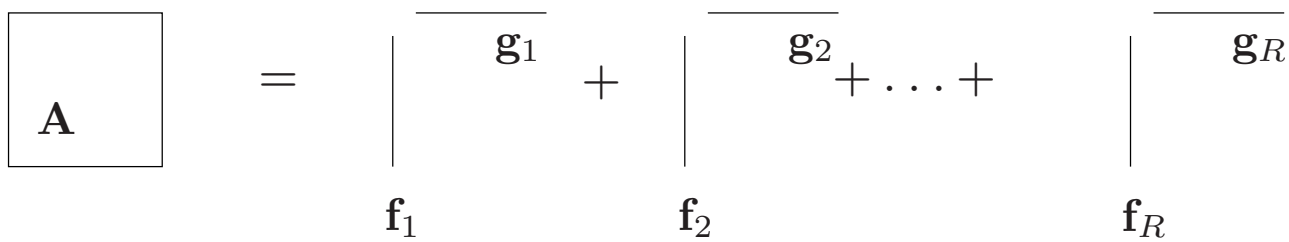
Overview

- Algebraic preliminaries
- Factor analysis and blind source separation
  - Principal Component Analysis (PCA)
  - Canonical Polyadic Analysis (CPA)
  - Independent Component Analysis (ICA)
- Block Term Decompositions
- Block Component Analysis
- Variants
- Conclusion

Factor analysis and blind source separation

- Decompose a data matrix in rank-1 terms
  E.g. independent component analysis, telecommunication, biomedical applications, chemometrics, exploratory data analysis, ...

$$\mathbf{A} = \mathbf{F} \cdot \mathbf{G}^T$$



- Decomposition in rank-1 terms is not unique

$$\begin{aligned} \mathbf{A} &= (\mathbf{F}\mathbf{M}) \cdot (\mathbf{M}^{-1}\mathbf{G}^T) \\ &= \tilde{\mathbf{F}} \cdot \tilde{\mathbf{G}}^T \end{aligned}$$

Principal Component Analysis

Exploitation of prior knowledge

PCA, SVD: uniqueness obtained by adding orthogonality constraints

$$\mathbf{A} = \mathbf{U}^{(1)} \cdot \mathbf{\Sigma} \cdot \mathbf{U}^{(2)^T}$$

$\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ orthogonal, $\mathbf{\Sigma}$ diagonal

Example: excitation-emission fluorescence in chemometrics

Matrix approach

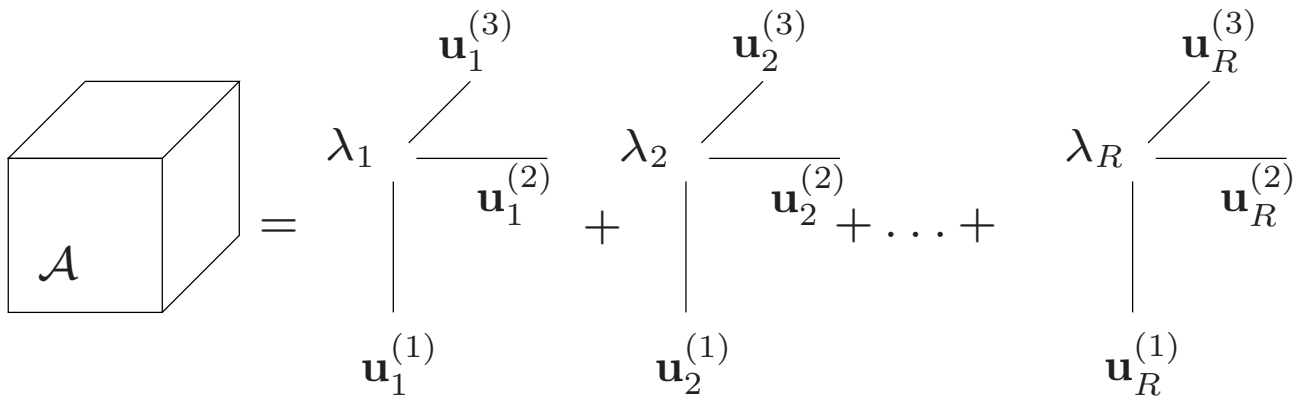
row vector $\sim$ emission spectrum
column vector $\sim$ excitation spectrum
coefficients $\sim$ concentrations

$$\boxed{\mathbf{A}} = \lambda_1 \, \mathbf{u}_1^{(1)} \, \overline{\mathbf{u}_1^{(2)}} + \lambda_2 \, \mathbf{u}_2^{(1)} \, \overline{\mathbf{u}_2^{(2)}} + \ldots + \lambda_R \, \mathbf{u}_R^{(1)} \, \overline{\mathbf{u}_R^{(2)}}$$

Tensor solution: CP Analysis

Tensorization: one matrix $\longrightarrow$ several matrices, stacked in tensor

row vector $\sim$ emission spectrum
column vector $\sim$ excitation spectrum
coefficients $\sim$ concentrations

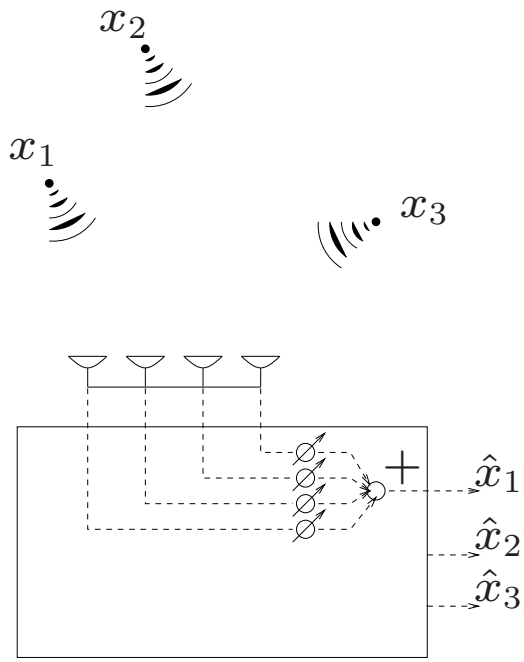


[*Smilde, Geladi, Bro '04*]

Independent Component Analysis (ICA)

Model:

$$Y = \mathbf{M}X$$



statistical independence implies: - the variables are uncorrelated
- additional conditions on the HOS

ICA: basic equations

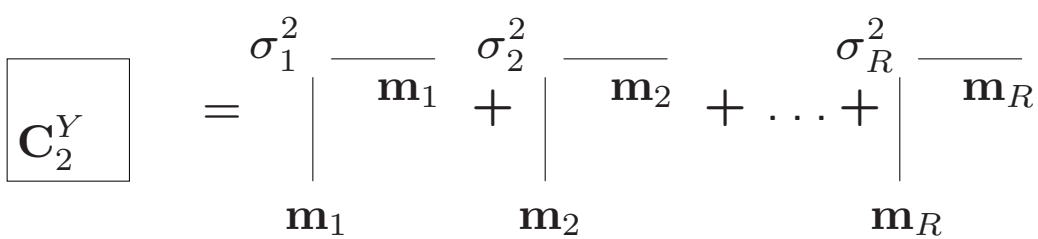
Model:

$$Y = \mathbf{M}X$$

Second order:

$$\begin{aligned}
\mathbf{C}_2^Y &= E\{YY^T\} \\
&= \mathbf{M} \cdot \mathbf{C}_2^X \cdot \mathbf{M}^{\mathbf{T}} \\
&= \mathbf{C}_2^X \cdot_1 \mathbf{M} \cdot_2 \mathbf{M}
\end{aligned}$$

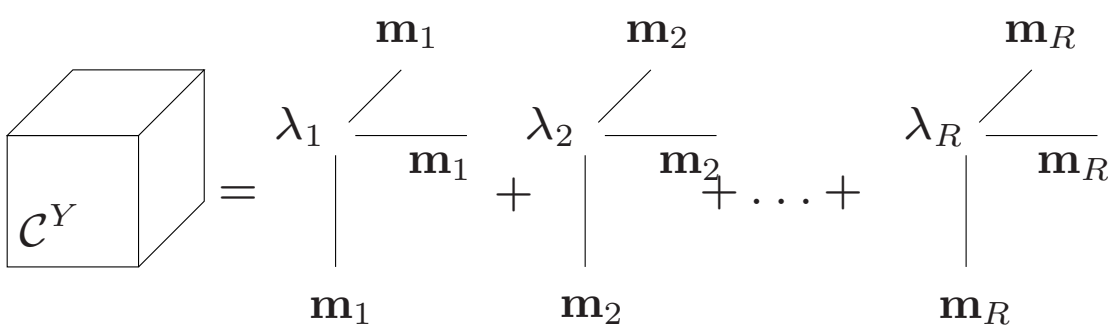
uncorrelated sources: $\mathbf{C}_2^X$ is diagonal
"diagonalization by congruence"



Higher order:

$$\mathcal{C}_4^Y = \mathcal{C}_4^X \cdot_1 \mathbf{M} \cdot_2 \mathbf{M} \cdot_3 \mathbf{M} \cdot_4 \mathbf{M}$$

independent sources: $\mathcal{C}_4^X$ is diagonal



Tensorization: decomposition data matrix $\longrightarrow$ CPD cumulant tensor

[*Comon '94*]

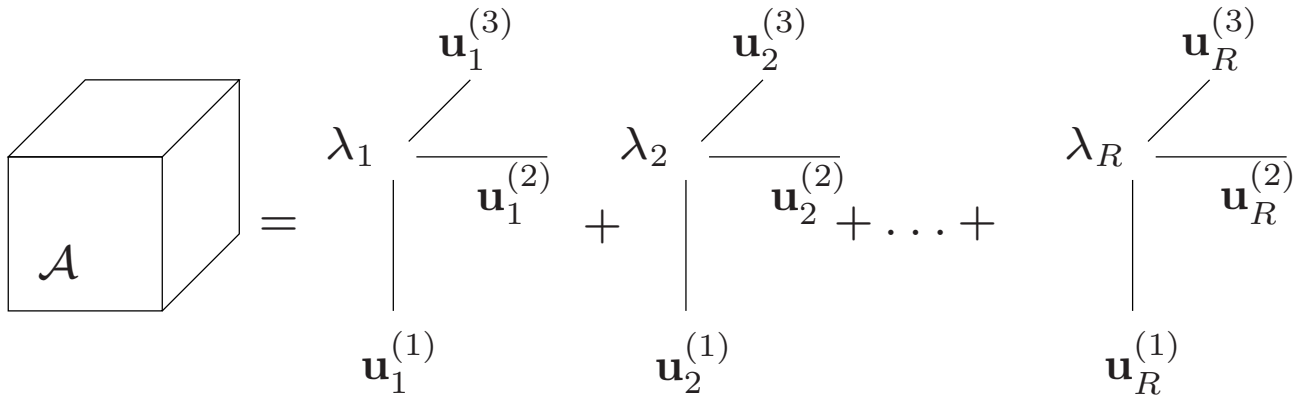
Overview

- Algebraic preliminaries
- Factor analysis and blind source separation
- Block Term Decompositions
- Block Component Analysis
- Variants
- Conclusion

Rank and associated decomposition

Rank: minimal number of rank-1 terms [*Hitchcock, 1927*]

Canonical Polyadic Decomposition (CPD) of a tensor $\mathcal{A}$ is its decomposition in a minimal sum of rank-1 tensors

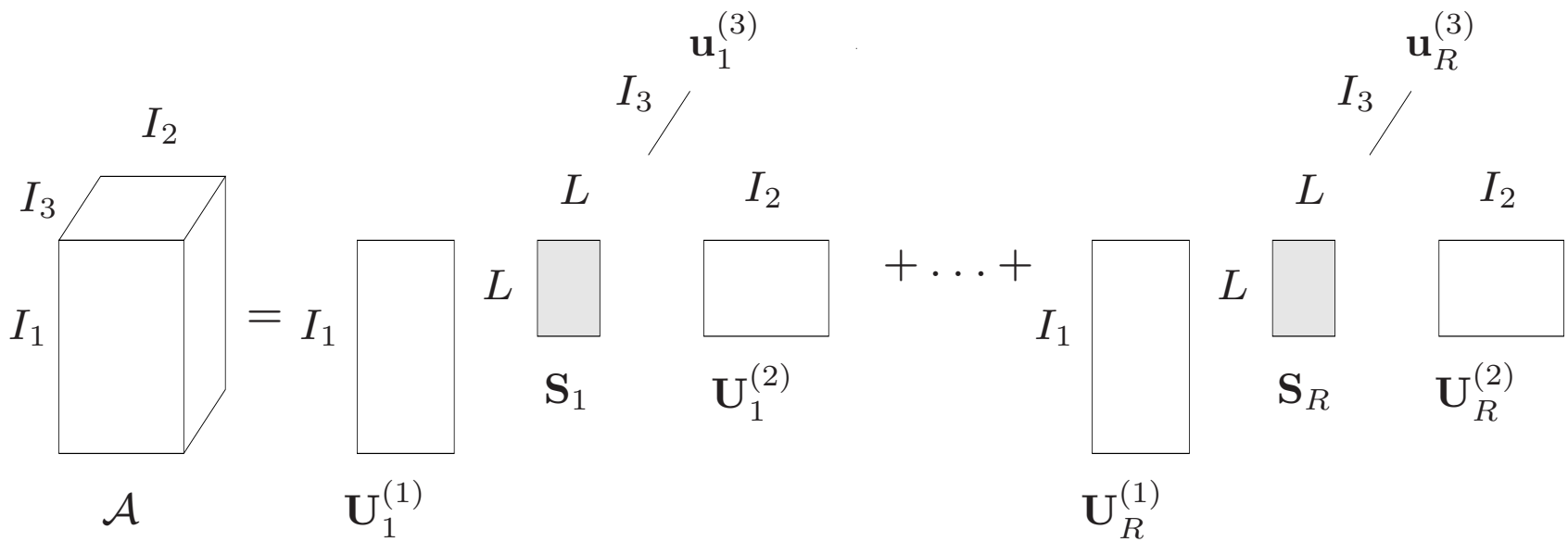


[*Harshman '70*], [*Carroll and Chang '70*]

Unique under mild conditions

Cf. R-th secant variety of Segre variety (unsymmetric), Veronese variety (symmetric)

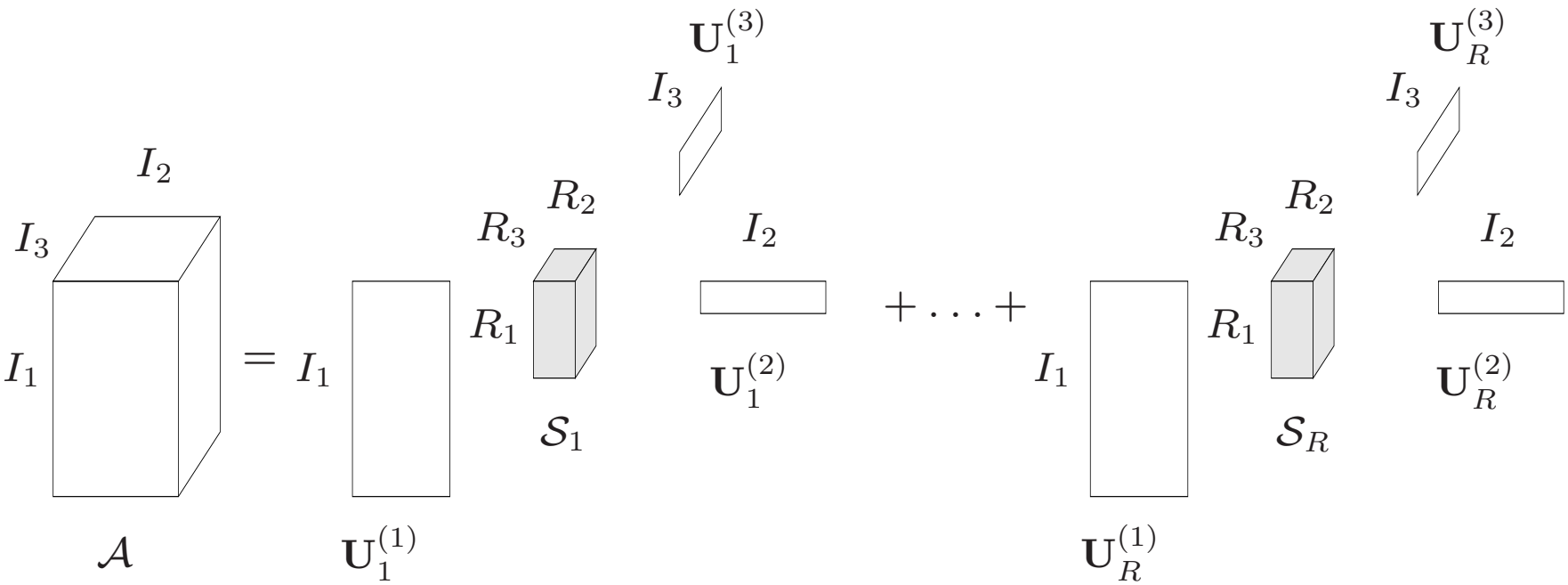
Decomposition in rank-$(L, L, 1)$ terms



Unique under mild conditions

[*De Lathauwer '06*]

Decomposition in rank-(R_1, R_2, R_3) terms



Unifies: multilinear rank and rank
Tucker and CPD

[*De Lathauwer '06*]

Overview

- Algebraic preliminaries
- Factor analysis and blind source separation
- Block Term Decompositions
- Block Component Analysis
- Variants
- Conclusion

Block Component Analysis

Demo

Wireless communication

e.g. Direct Sequence Code Division Multiple Access

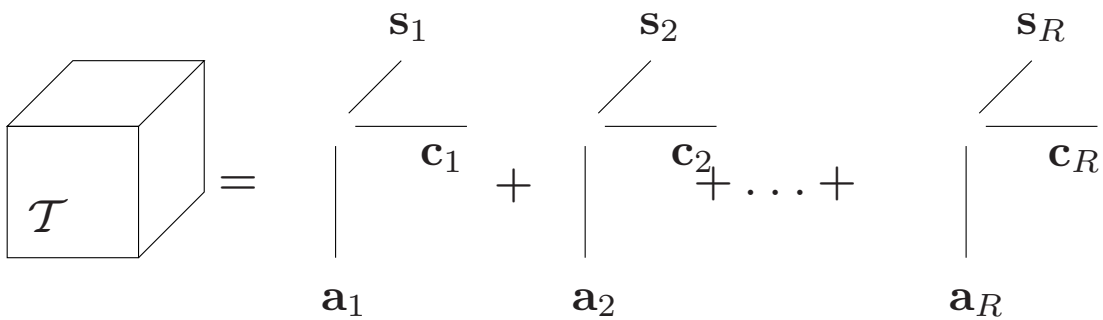
DS-CDMA signal $\sim$ rank-1 matrix

Signal transmitted by user r:

$$
\begin{aligned}
& s_{1r}\left[c_{1r} c_{2r} \cdots c_{Jr}\right] \\
& s_{2r}\left[c_{1r} c_{2r} \cdots c_{Jr}\right] \\
& s_{3r}\left[c_{1r} c_{2r} \cdots c_{Jr}\right] \\
& \qquad \vdots \\
& \sim \mathbf{s}_r \mathbf{c}_r^T = \mathbf{s}_r \circ \mathbf{c}_r^T
\end{aligned}
$$

DS-CDMA signal received by antenna array $\sim$ rank-1 tensor

Dimension 1 $\sim$ temporal diversity
Dimension 2 $\sim$ spectral diversity
Dimension 3 $\sim$ spatial diversity



Deterministic blind signal separation based on CPD

[*Sidiropoulos et al. '00*]

Generalization propagation model:

- CPD: line-of-sight propagation (no ISI)
- decomposition in rank-$(L_r, L_r, 1)$ terms: reflexions in far field of antenna array

[*De Lathauwer and de Baynast '08*]

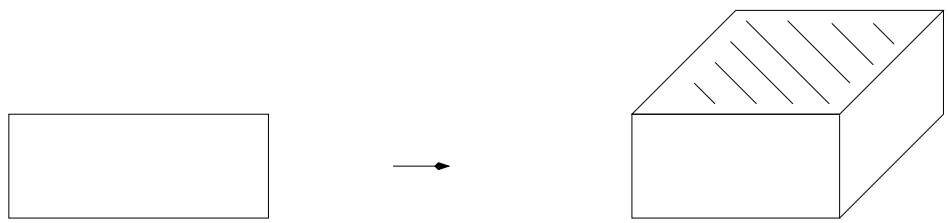
- decomposition in rank-$(L_r, M_r, \cdot)$ terms: reflexions not only in far field of antenna array

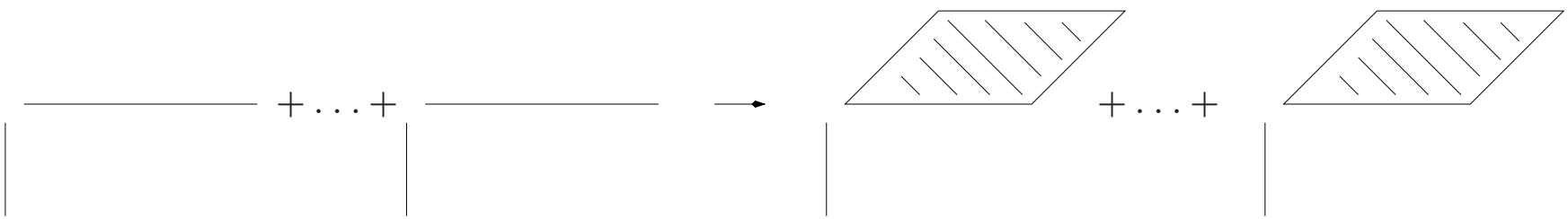
[*Nion and De Lathauwer '08*]

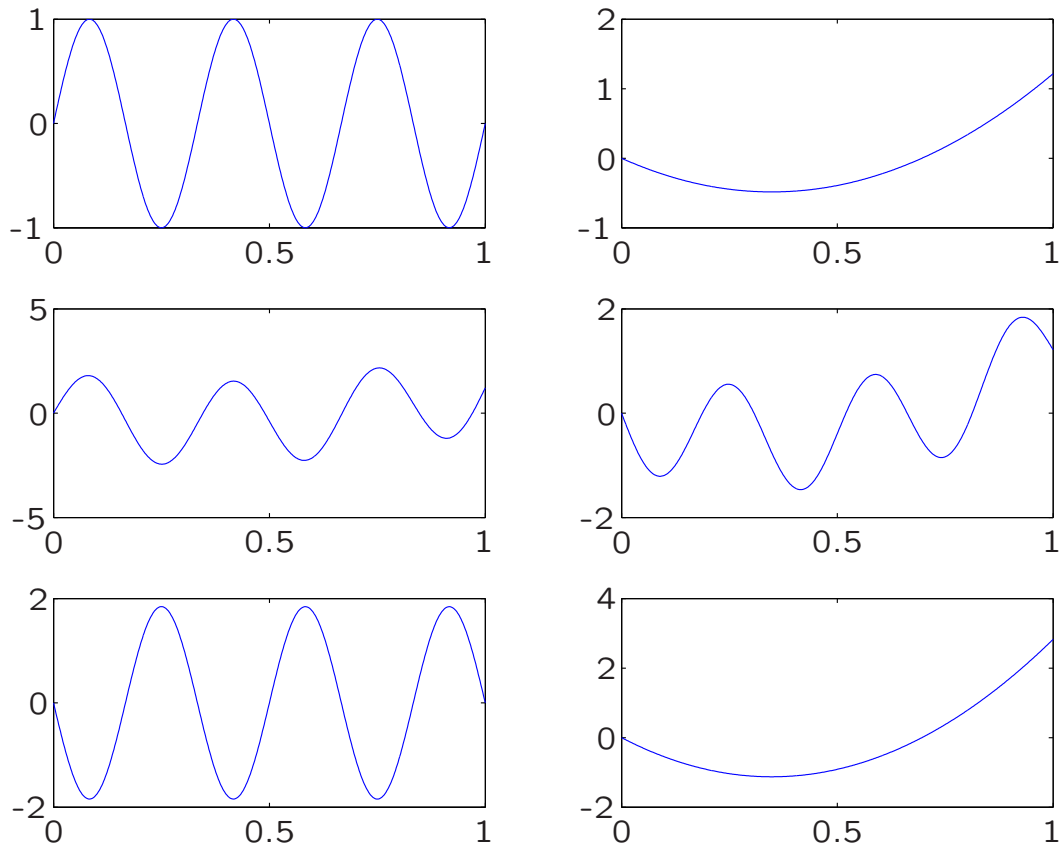
General comment: rank-1 terms are easiest to compute
rank-1 terms may be too simple

Exponentials, sinusoids, polynomials, exponential polynomials

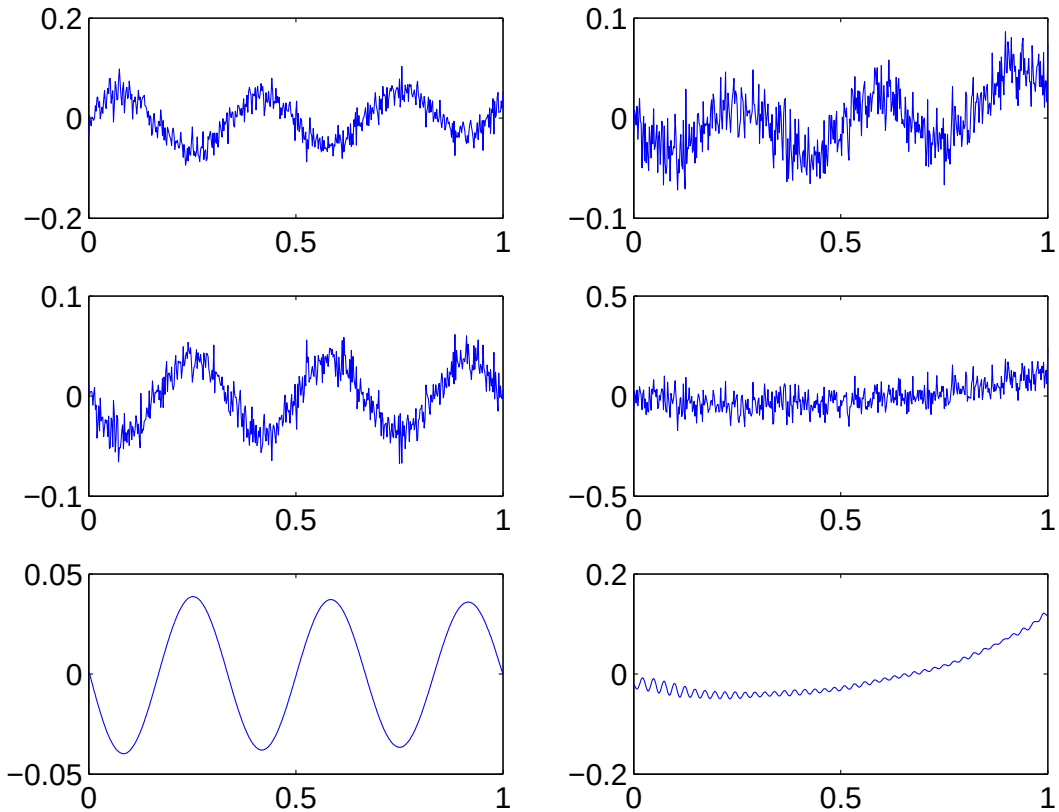
Principle: Map every row of $\mathbf{A} = \mathbf{F} \cdot \mathbf{G}^T$ to Hankel matrix
Hankel matrices are often very ill-conditioned
Hankel matrices generated by exponential polynomials are exactly low-rank





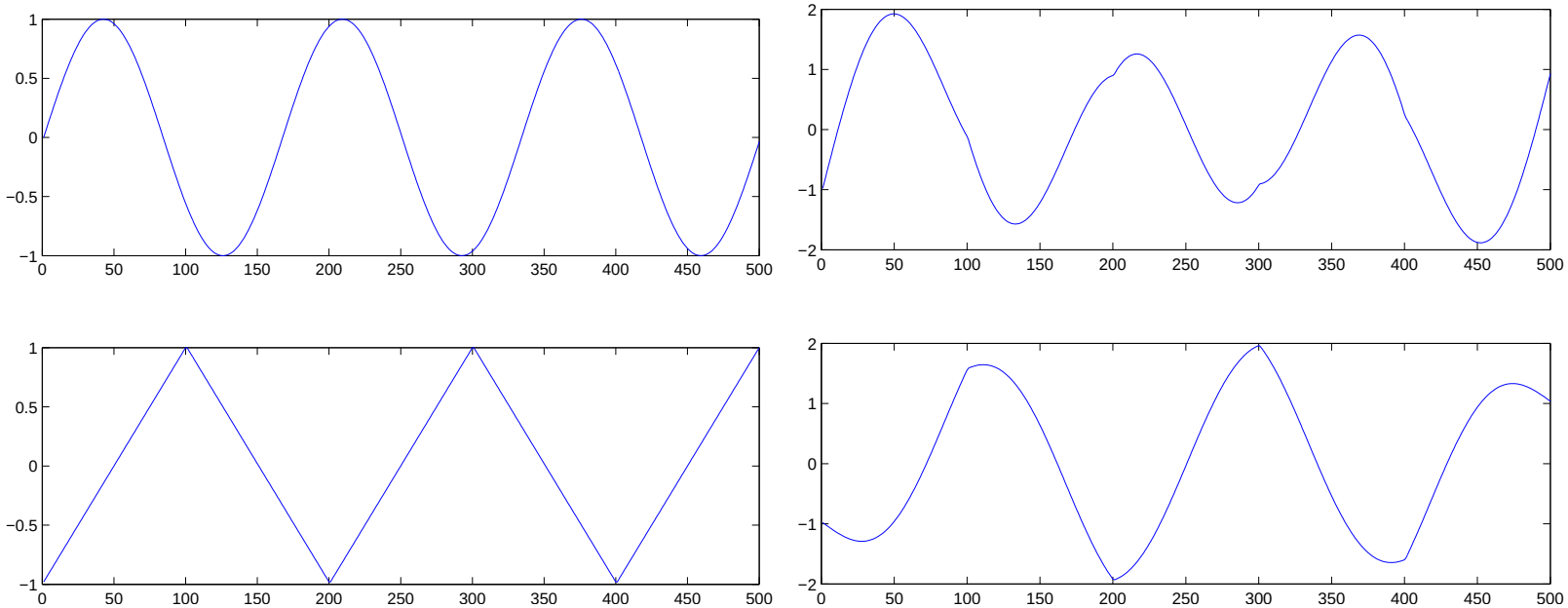


theoretical values: $(L_1, L_2) = (2, 3)$
perfect separation: $(L_1, L_2) = (2, 3), (3, 3), (2, 4), (3, 4), (4, 4)$
good separation: $(L_1, L_2) = (2, 2), (1, 2)$

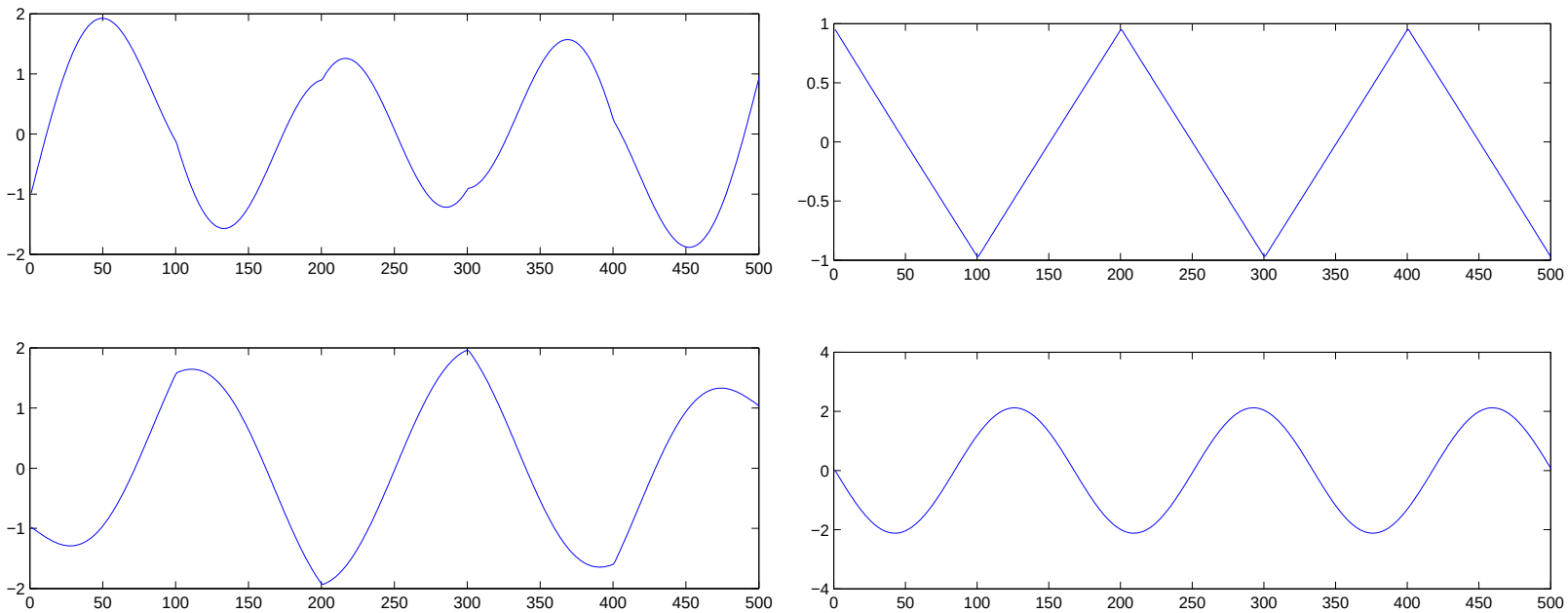


501 samples, SNR = 5 dB
good separation: $(L_1, L_2) = (1, 2), (2, 2), (2, 3)$

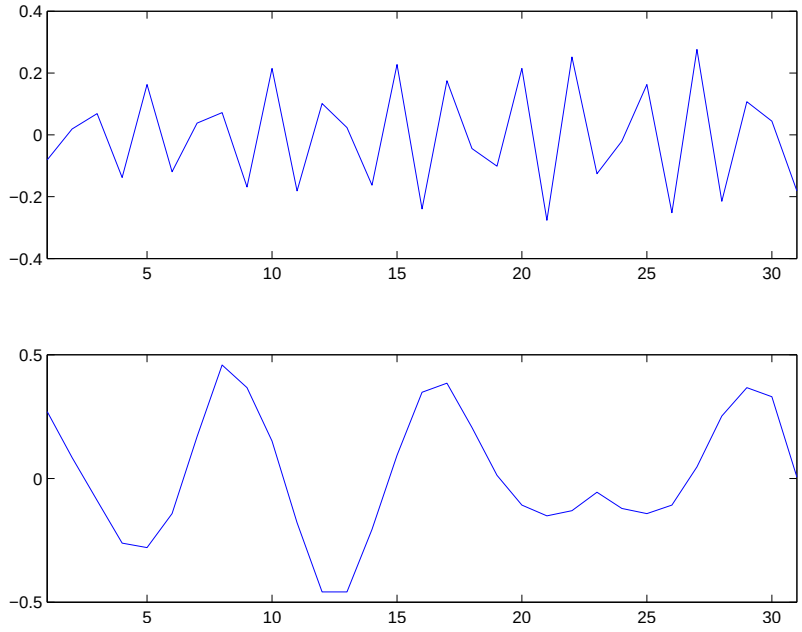
theoretical values: $L_1 = 2, L_2 = 251$



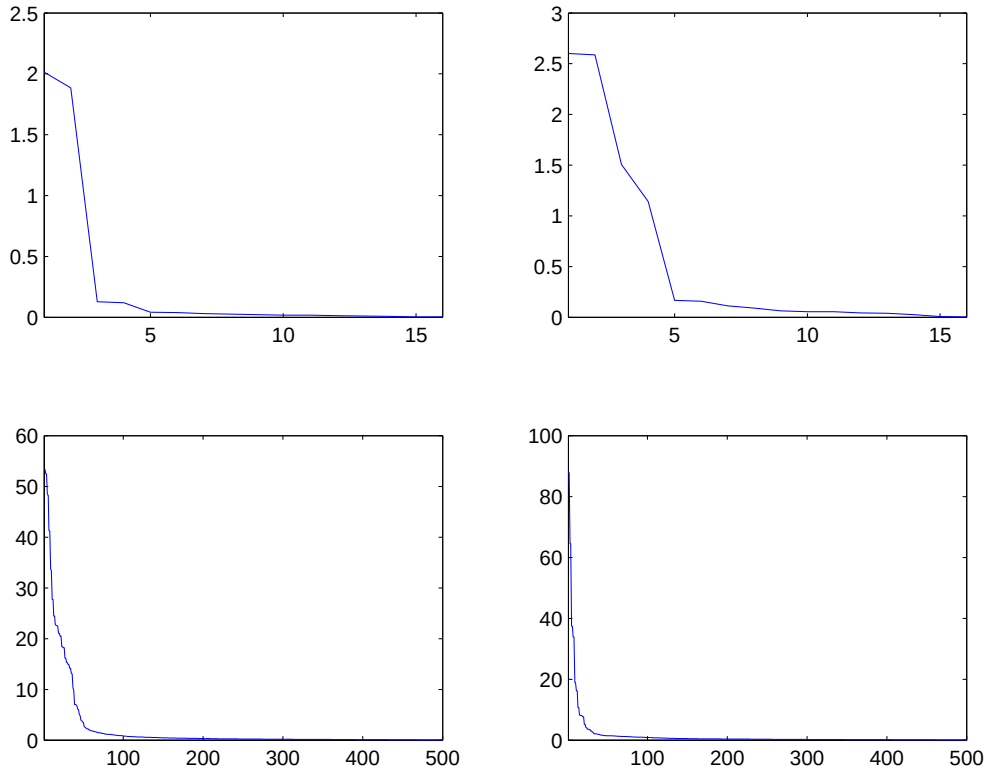
theoretical values: $L_1 = 2, L_2 = 251$
results: $L_1 = 2,\ L_2 = 2, 3, \ldots, 7$



Audio signals



`Chirp` (top) and `train` (bottom) signal, 31 samples



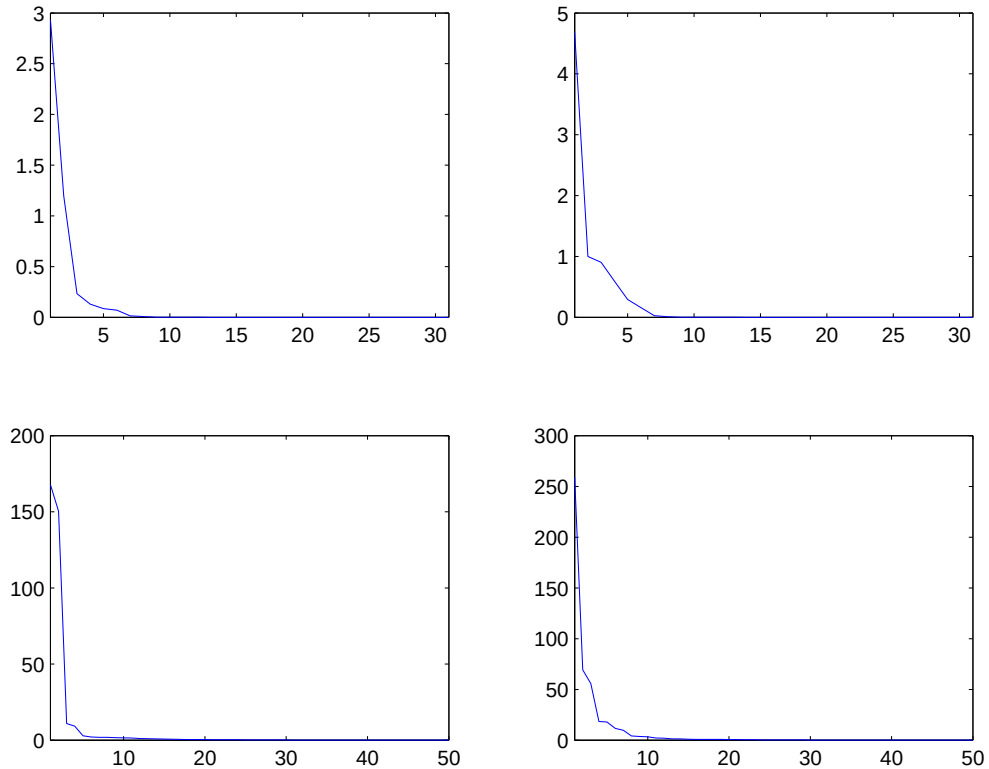
singular values of Hankel matrices generated by `chirp` (left) and `train` (right)
top: 31 samples; bottom: 1000 samples

L_1 / L_2	1	2	3	4	5	6	7
1	20	48	49	37	20	15	15
2	48	47	49	48	44	17	16
3	49	49	49	47	23	20	19
4	37	48	47	47	47	20	18
5	20	44	23	47	45	29	16
6	15	17	20	20	29	25	33
7	15	16	19	18	16	33	24

mean SIR [dB] (Hankel, noiseless) (ICA: COM2: 15 dB, JADE: 14 dB)

L_1 / L_2	1	2	3	4	5	6	7
1	49	47	49	51	51	19	13
2	47	47	50	49	51	38	22
3	49	50	49	48	49	47	45
4	51	49	48	47	48	46	44
5	51	51	49	48	48	46	44
6	19	38	47	46	46	46	47
7	13	22	45	44	44	47	44

median SIR [dB] (Hankel, noiseless) (ICA: COM2: 15 dB, JADE: 14 dB)



singular values of wavelet matrices generated by `chirp` (left) and `train` (right)
top: 31 samples; bottom: 1000 samples

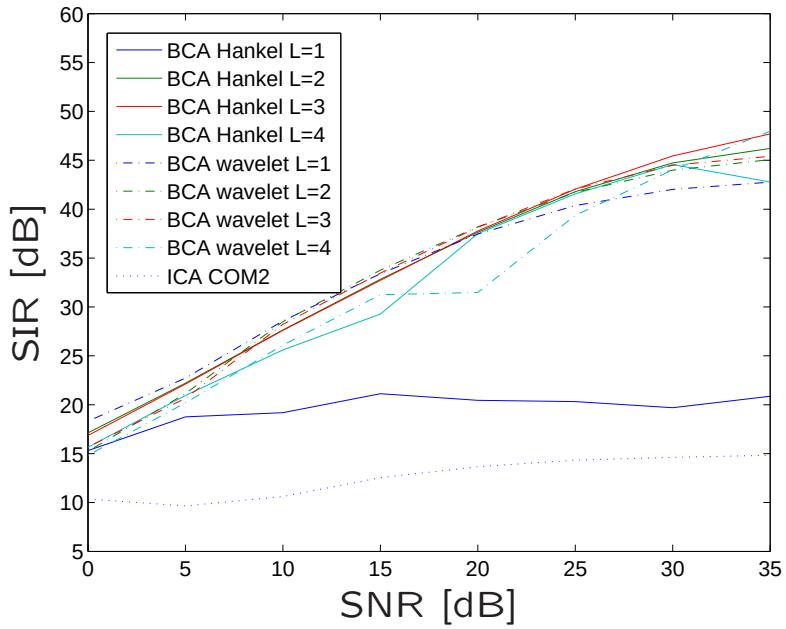
L_1 / L_2	1	2	3	4	5	6	7
1	43	41	19	19	16	14	13
2	41	46	47	48	33	17	14
3	19	47	46	27	25	18	13
4	19	48	27	52	16	47	13
5	16	33	25	16	13	28	12
6	14	17	18	47	28	46	17
7	13	14	13	13	12	17	17

mean SIR [dB] (wavelet, noiseless)

L_1 / L_2	1	2	3	4	5	6	7
1	43	41	41	41	16	13	12
2	41	46	49	52	33	18	17
3	41	49	46	41	32	33	12
4	41	52	41	52	21	48	12
5	16	33	32	21	13	35	12
6	13	18	33	48	35	47	25
7	12	17	12	12	12	25	20

median SIR [dB] (wavelet, noiseless)

Results for noisy data:



Conclusion

- In BCA signals are separated on the basis of low intrinsic dimensionality
- Foundation related to Pareto principle, compressed sensing, etc.
  Q: How many samples are needed for separation?
- Data can be tensorized in many ways (HOS, Hankel, . . .)
- In some cases, the rank-1 hypothesis is questionable
- PCA, ICA, CPA, NMF, . . . : easier to use but assumptions should be verified
- Constrained BCA: nonnegativity, sparsity, orthogonality, statistical independence, . . .

  Related work: CPA with independence constraint
  [*De Vos, Van Huffel, De Lathauwer '10*]
  CPA with orthogonality constraint
  [*Sørensen, De Lathauwer, Deneire '10*]