Learning complex-valued latent filters with absolute cosine similarity

Anh H.T. Nguyen, V.G. Reju, Andy W.H. Khong, and Ing Yann Soon
School of EEE, Nanyang Technological University, Singapore
nguyenha001@e.ntu.edu.sg, {reju, andykhong, eiysoon}@ntu.edu.sg.

ICASSP'2017,New Orleans, USA

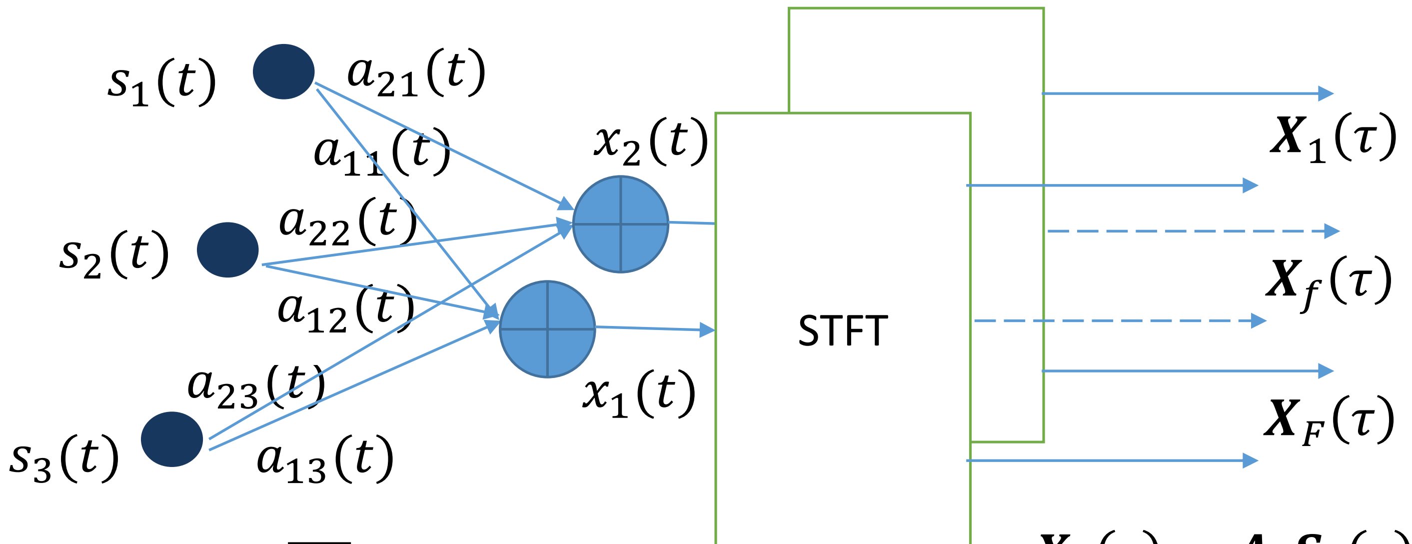
Overview

- Goal: *Under-determined convolutive* blind source separation
- Objective: Improve the accuracy of mixing matrix estimation
- Existing algorithms: Directional clustering and sparse coding
- Challenges: *Complex-valued mixing matrix* and non-convexity

NANYANG TECHNOLOGICAL UNIVERSITY

Blind separation of convolutive mixtures

- fMRI signals
- Multi-channel recording of real speeches



$$\boldsymbol{X}_f(\tau) = \boldsymbol{A}_f \boldsymbol{S}_f(\tau)$$

$$x_i(t) = \sum_j a_{ij}(t) * s_j(t)$$

Fig.: Under-determined convolutive mixtures

Complex-valued mixing model

- Linear and noiseless

$$\mathbf{x}[k] = \mathbf{A}\mathbf{s}[k] = \sum_{j=1}^{N} \mathbf{a}_j s_j[k]$$

where $\mathbf{x}[k] \in \mathbb{C}^{M\times 1}$ is the data we have

$\mathbf{s}[k] \in \mathbb{C}^{N\times 1}$ is the latent sources

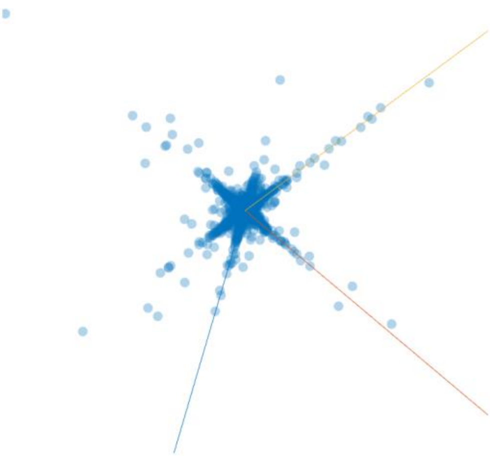
and $\mathbf{a}_j$ is *j*th latent filters/atom/factor/etc.

- M = N: determined mixing process
- M < N: *under-determined mixing process*

Assumptions

- The sources are sufficiently sparse so that observed data is directional.

$$\mathbf{x}[k] = \mathbf{A} \begin{bmatrix} s_1[k] \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \approx \mathbf{a}_1 s_1[k]$$



$$\because s_1[k] \gg \epsilon_j$$

- Infinite unit vectors having the same direction in complex vector space

$$D^2(\mathbf{x}[k], \mathbf{a}_1) = 1 - \cos^2 \theta_H(\mathbf{x}[k], \mathbf{a}_1) = 1 - \frac{\left|\mathbf{a}_1^H \mathbf{x}[k]\right|^2}{\|\mathbf{a}_1\|_2^2 \|\mathbf{x}[k]\|_2^2} \approx 0$$

Assumptions (cont.)

- The sources are zero-mean and unit-variance

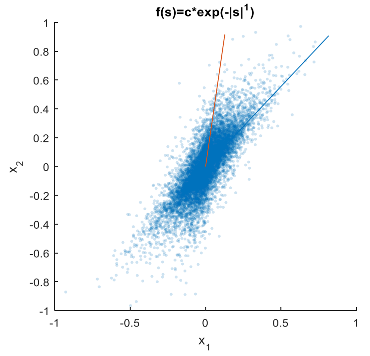


Fig.: Mixtures of Laplace sources

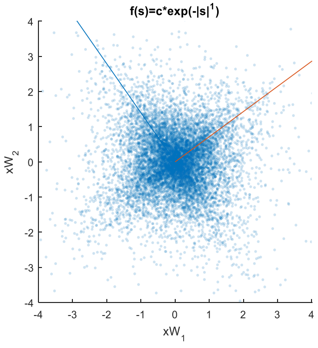


Fig.: Whitened mixtures

$$\mathbf{x}[k] = \mathbf{A}\mathbf{s}[k]$$

$$\mathbf{C}_s = E\{\mathbf{s}[k]\mathbf{s}^H[k]\} = \mathbf{I}$$

$$\Rightarrow$$

$$\mathbf{x}_W[k] = \mathbf{Q}\mathbf{x}[k] = \mathbf{Q}\mathbf{A}\mathbf{s}[k]$$

$$E\left\{\mathbf{x}_w[k]\mathbf{x}_w^H[k]\right\} = \mathbf{Q}\mathbf{A}\mathbf{A}^H\mathbf{Q}^H = \mathbf{I}$$

- The mixing matrix for pre-whitened data is semi-unitary/unitary

$$\mathbf{A}\mathbf{A}^H = \mathbf{I} \quad \Rightarrow \quad \sum_j \|\mathbf{a}_j\|_2^2 \cos^2\theta_H(\mathbf{x}[k], \mathbf{a}_j) = 1$$

- Observation: in under-determined case, minimizing the sparsity penalty of cosine similary is suboptimal for directional data.

Issues related to existing methods

- Sparse filtering uses an unsuitable sparsity enforcer for directional data.

$$\min_{\widehat{\mathbf{A}}} \; E\left\{ \left\| \widehat{\mathbf{A}}^H \mathbf{x} \right\|_1 / \left\| \widehat{\mathbf{A}}^H \mathbf{x} \right\|_2 \right\}$$

$$\longrightarrow \quad \min_{\widehat{\mathbf{A}}} \; E\left\{ \sum_{j=1}^{N} \|\widehat{\mathbf{a}}_j\|_2 \cos\theta_H(\mathbf{x}[k], \widehat{\mathbf{a}}_j) \right\}$$

- K-hyperlines works best for perfectly directional data

$$\min_{\widehat{\mathbf{A}}} \; E\left\{ \min_{j=1,\ldots,N} D^2(\mathbf{x}[k], \mathbf{a}_j) \right\}$$

- “Soft” extensions of K-hyperlines are computationally expensive
- Existing methods do not exploit the semi-unitary property of the mixing matrix.

Proposed algorithm

- Minimize the expected power mean of the phase-invariant cosine distance subject to semi-unitary constraint

$$\min_{\widehat{\mathbf{A}}} \ J(\widehat{\mathbf{A}};\ r), \ \text{s.t.}\ \widehat{\mathbf{A}}\widehat{\mathbf{A}}^H = \mathbf{I}_M.$$

where

$$J(\widehat{\mathbf{A}};\ r) = E\left\{\left[\frac{1}{N}\sum_{j=1}^{N}\left(D^2(\mathbf{x}[k], \widehat{\mathbf{a}}_j)\right)^r\right]^{1/r}\right\},\ \ r \in (-\infty, 1).$$

- Why the power mean?
  - Numerically stable: $\mu(y_1, y_2, \ldots, y_N;\ r) = y_{\min}\left[\frac{1}{N}\sum_{j=1}^{N} y_j^r / y_{\min}^r\right]^{1/r}$
  - Schur-concave (which acts as a sparsity enforcer)
  - Smooth surrogate of the minimum function (sparsity enforcer for approximately 1-sparse sources)

NANYANG TECHNOLOGICAL UNIVERSITY

Proposed algorithm (cont.)

- Semi-unitary constrained non-convex optimization problem
- Reparametrize semi-unitary constrained problems into *unconstrained ones in Euclidean space*

$$\min_{\widehat{\mathbf{A}}} \ f(\widehat{\mathbf{A}}), \ \text{s.t.} \ \widehat{\mathbf{A}}\widehat{\mathbf{A}}^H = \mathbf{I}_M$$

$$\Rightarrow \quad \min_{\mathbf{B}} \ f(\widehat{\mathbf{A}}) \ \text{s.t.} \ \widehat{\mathbf{A}} = (\mathbf{B}\mathbf{B}^H)^{-1/2}\mathbf{B}$$

- Unconstrained problem w.r.t. $\mathbf{B}$ which can be solved by off-the-shelf tools, e.g., L-BFGS, *NAG*, SGD, momentum, etc.
- $\hat{\mathbf{A}}$ is the nearest semi-unitary matrix of $\mathbf{B}$ (many-to-one mapping)
  - $\mathbf{B}$ must be full row rank
  - $\hat{\mathbf{A}}$ is always feasible
  - Same cost for all matrices that are mapped to the same $\hat{\mathbf{A}}$

Proposed algorithm (cont.)

- Backpropagation through nearest semi-unitary projector is practical.

Algorithm 1 Gradient of in-line row-wise decoupling scheme

1: $\mathbf{U},\ \boldsymbol{\Sigma},\ \mathbf{V} \leftarrow \mathrm{SVD}(\mathbf{B})$
2: $\boldsymbol{\sigma} \leftarrow \mathrm{diag}(\boldsymbol{\Sigma})$
3: $\widehat{\mathbf{A}} \leftarrow \mathbf{U}\mathbf{V}^H$
4: Find f and $\nabla_{\widehat{\mathbf{A}}^*} f$ for a batch or minibatch
5: $\mathbf{C} \leftarrow -(\boldsymbol{\Sigma}^{-1}\mathbf{U}^H(\nabla_{\widehat{\mathbf{A}}^*} f)\mathbf{V}) \oslash (\mathbf{1}\boldsymbol{\sigma}^T + \boldsymbol{\sigma}\mathbf{1}^T)$
6: $\nabla_{\mathbf{B}^*} f \leftarrow \mathbf{U}(\mathbf{C}^H + \mathbf{C})\boldsymbol{\Sigma}\mathbf{V}^H + \mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{U}^H\nabla_{\widehat{\mathbf{A}}^*} f$

- *One economy-size SVD per batch or minibatch.*
- May be useful for other signal processing applications or machine learning applications as well.

Proposed algorithm (cont.)

- Comparison to to optimization on Stiefel manifold

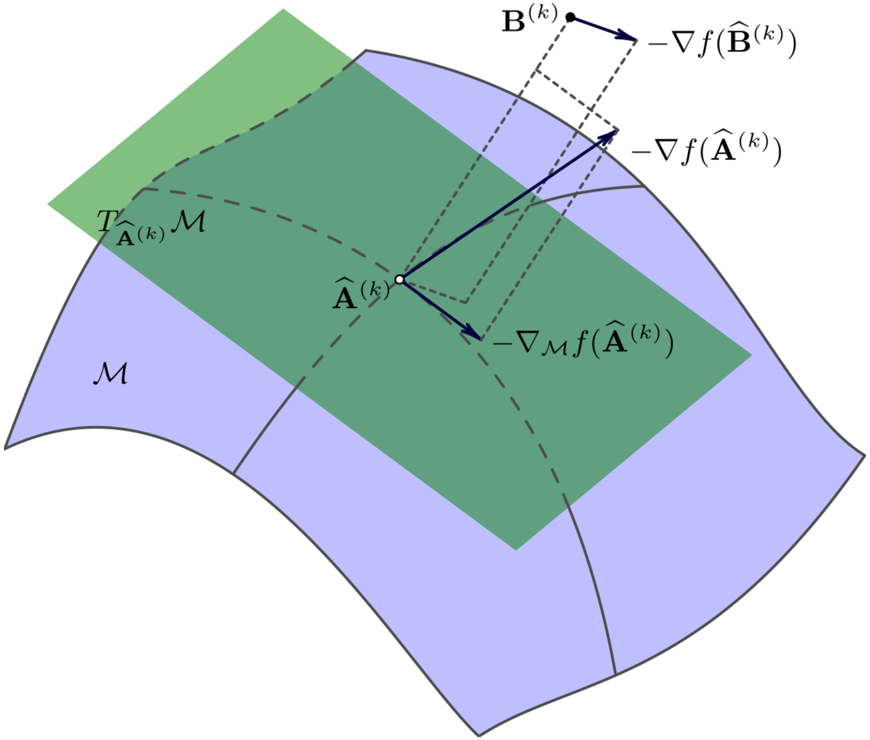


Fig.: Proposed reparameterization.

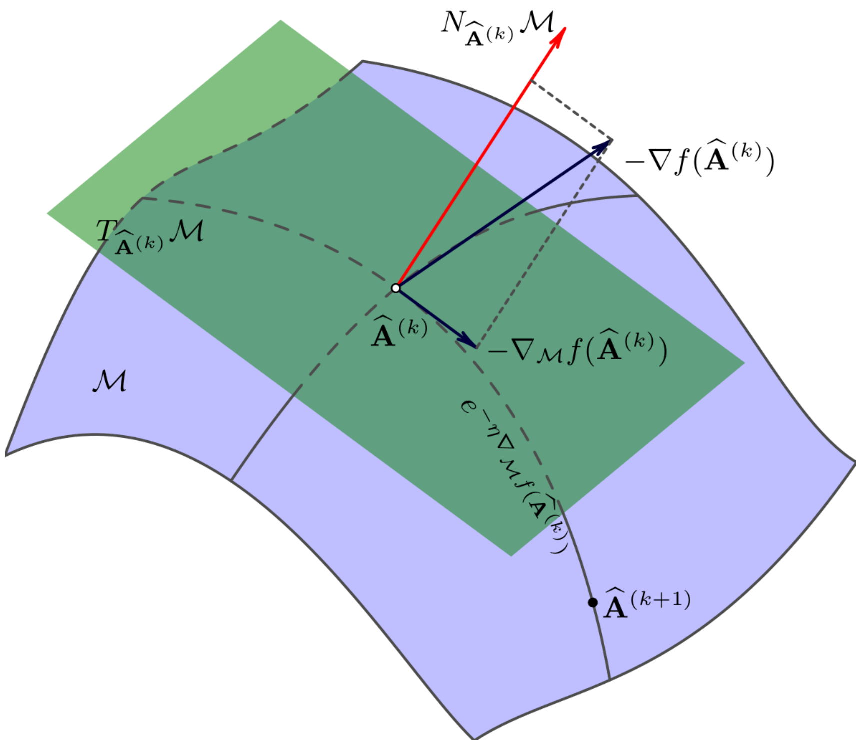


Fig.: Optimization on Stiefel manifold.

NANYANG TECHNOLOGICAL UNIVERSITY

Simulation results

- Performance criteria
  - Mixing-error-ratio (MER)

$$\mathrm{MER} = (20/N)\sum_j \log\left(\left\|\mathbf{a}_j^{\mathrm{coll}}\right\| / \left\|\mathbf{a}_j^{\mathrm{orth}}\right\| \right),$$

  - Signal-distortion-ratio (SDR) and signal-interference-ratio (SIR):

$$\hat{s}_{ij}(t) = s_{ij}(t) + e_{ij}^{\mathrm{spat}}(t) + e_{ij}^{\mathrm{interf}}(t) + e_{ij}^{\mathrm{artif}}(t)$$

$$SDR_j = 10\log\frac{\sum_{i,t} s_{ij}^2(t)}{\sum_{i,t}(e_{ij}^{\mathrm{spat}}(t)+e_{ij}^{\mathrm{interf}}(t)+e_{ij}^{\mathrm{artif}}(t))^2}$$

$$SIR_j = 10\log\frac{\sum_{i,t} s_{ij}^2(t)+(e_{ij}^{\mathrm{spat}}(t))^2}{\sum_{i,t}(e_{ij}^{\mathrm{interf}}(t))^2}$$

Simulation results (cont.)

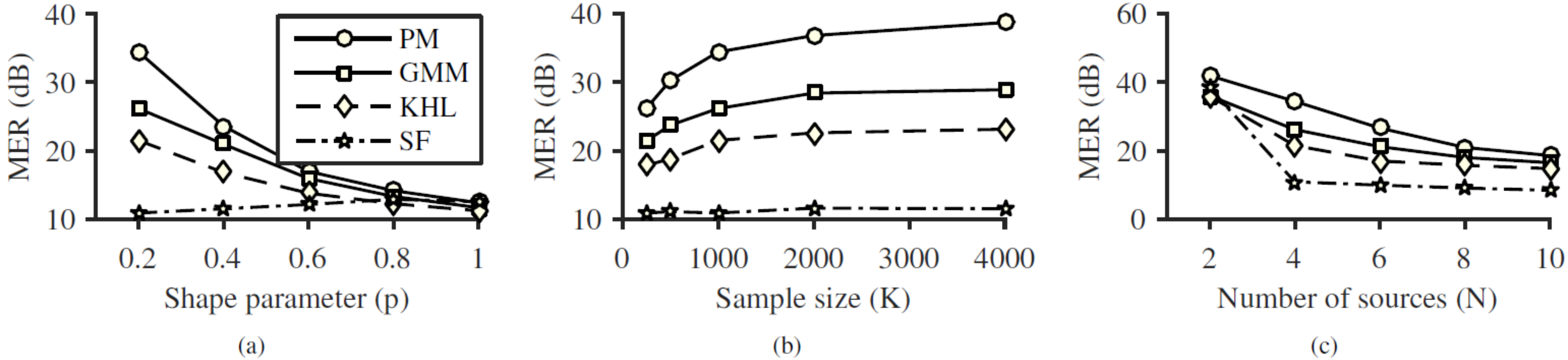


Fig. 1: Average MER in estimation of 2×4 mixing matrix w.r.t. : a) Sparseness. b) Sample size. c) Number of sources

- Synthesized data
  - Better estimation can be achieved.
  - Sparse filtering failed to recover the mixing matrix in under-determined case

Simulation results (cont.)

- Blind separations of under-determined live recording speeches

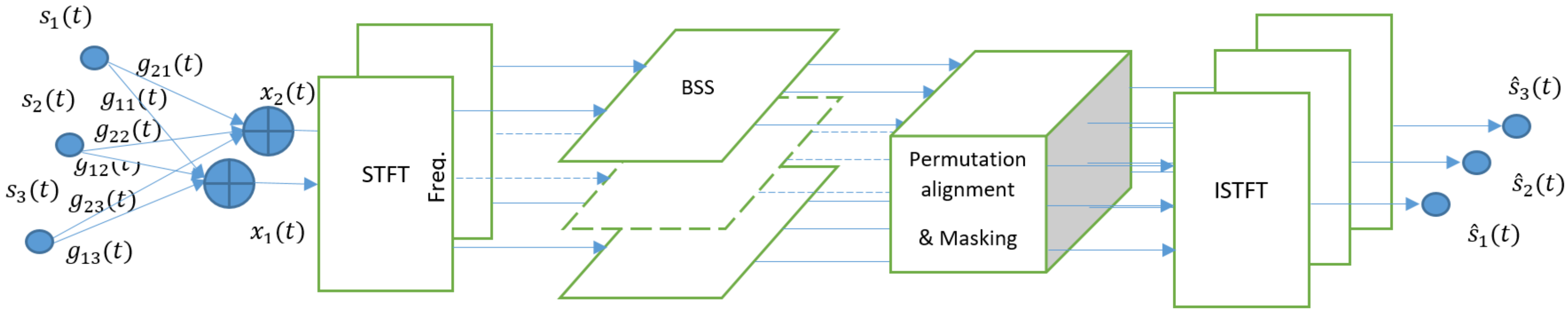


Table 1: Output SDR and SIR in dB for 2mic_4src_5cm subset of SiSEC dev1 dataset

RT60	130ms				250ms			
Source	4 males		4 females		4 males		4 females	
Perf. metric	SDR	SIR	SDR	SIR	SDR	SIR	SDR	SIR
PM	**4.55**	**8.27**	3.80	6.38	**3.67**	**6.06**	**3.57**	**5.36**
[22]	4.1	6.38	**4.47**	**6.48**	3.55	5.07	3.5	4.85
[9]	3.31	-	3.92	-	2.62	-	3.49	-
Input	-4.81	-4.60	-4.76	-4.68	-4.79	-4.64	-4.83	-4.71

Simulation results (cont.)

- Improvement in SIR is 14% (improvement in SDR is 2%) on SISEC dev1 dataset compared to the state of the art by Cho et. al.
- Much faster (up to 1 minute vs. up to 1 hours).

Male sources 2mic-4src, 130ms, 5cm	Mix.	Est src1	Est src2	Est src3	Est src4
Proposed method					
Cho et. al.					

NANYANG TECHNOLOGICAL UNIVERSITY

Conclusions

- Better estimation of complex-valued mixing matrix can be achieved by minimizing the expected power-mean of phase-invariant cosine distance subject to semi-unitary constraint.
- Semi-unitary constrained problems can be efficiently reparametrized into unconstrained problems in Euclidean space.

NANYANG TECHNOLOGICAL UNIVERSITY

THANK YOU FOR YOUR ATTENTION

Email: nguyenha001@e.ntu.edu.sg, {reju, andykhong, eiysoon}@ntu.edu.sg.