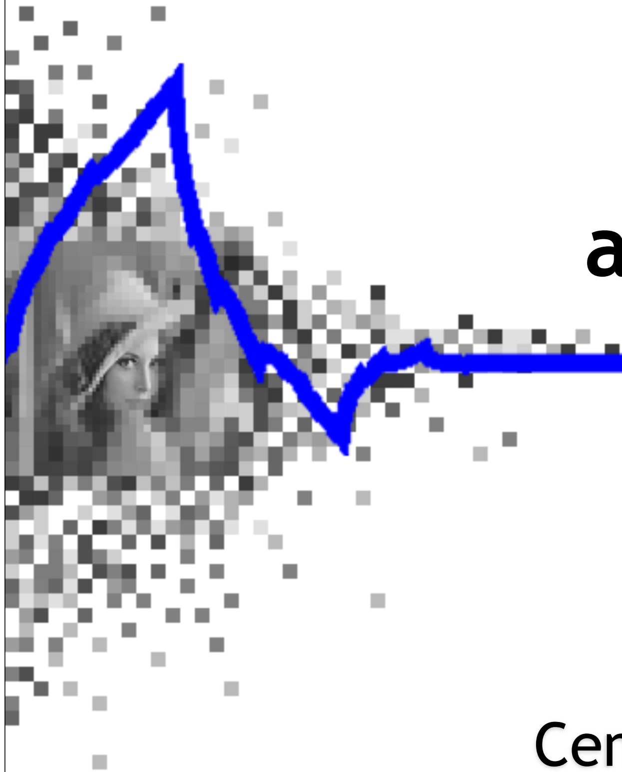
Aisenstadt Chair Course
CRM September 2009



Part V
Dictionary Learning and Source Separation

Stéphane Mallat

Centre de Mathématiques Appliquées
Ecole Polytechnique

Dictionary Learning

- PCA computes the orthonormal basis minimizing the average linear approximation error over of a signal set.

- Can we compute a redundant dictionary of size $P \quad N$ which minimizes the average non-linear approximation error over of a signal set ?

- NP-hard but greedy optimizations are possible.

- Are perception system learning redundant dictionaries to decompose input signals ?

Dictionary Update

- We want to optimize a dictionary $\mathcal{D} = \{\phi_p\}_{p\in\Gamma}$ to represent sparsely a training set of signals $\{f_k\}_{1\leq k\leq K}$:

$$\tilde{f}_k = \sum_{p\in\Gamma} a[k,p]\,\phi_p \quad \text{with} \quad \|f_k - \tilde{f}_k\| \leq \epsilon \ .$$

- Alternate optimization of the matrix of sparse decomposition coefficients $A = \{a[k,p]\}_{1\leq k\leq K, p\in\Gamma}$ and of the dictionary $\mathcal{D} = \{\phi_p\}_{p\in\Gamma}$ to minimize:

$$\sum_{k=0}^{K-1} \|f_k - \sum_{p\in\Gamma} a[k,p]\,\phi_p\|^2$$

- Minimum:

$$\{\phi_p[n]\}_{p,n} = (A^* A)^{-1}\, A^*\, \{f_k[n]\}_{k,n}$$

Greedy Optimization Algorithm

- **1.** *Initialization*: each $\phi_p[n]$ is a Gaussian white noise.

- **2.** *Sparse approximation:* matching or basis pursuit calculation of $A = \{a[k,p]\}_{1 \leq k \leq K, p \in \Gamma}$ satisfying

$$\Big\| f_k - \sum_{p \in \Gamma} a[k,p]\, \phi_p \Big\| \leq \epsilon \quad \text{for} \quad 1 \leq k \leq K \ .$$

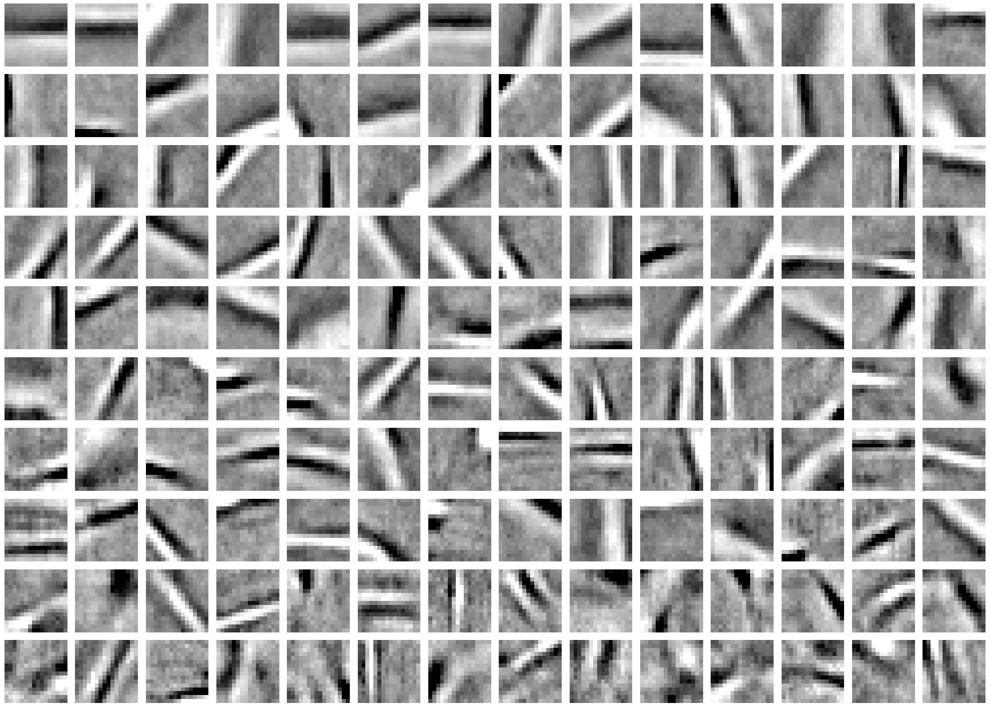
- **3.** *Dictionary update* to minimize the total error:

$$\{\phi_p[n]\}_{p,n} = (A^* A)^{-1} \, A^* \, \{f_k[n]\}_{k,n}$$

- **4.** *Dictionary normalization:* set $\|\phi_p\| = 1.$
- **5.** *Stop if* $\mathcal{D} = \{\phi_p\}_{p \in \Gamma}$ is marginally modified *or go to* **2.**

Dictionary from Natural Images

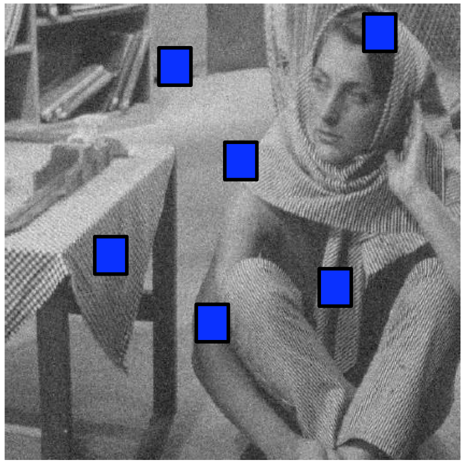
- Optimized dictionary obtained with fixed size vectors with a training set of natural images:



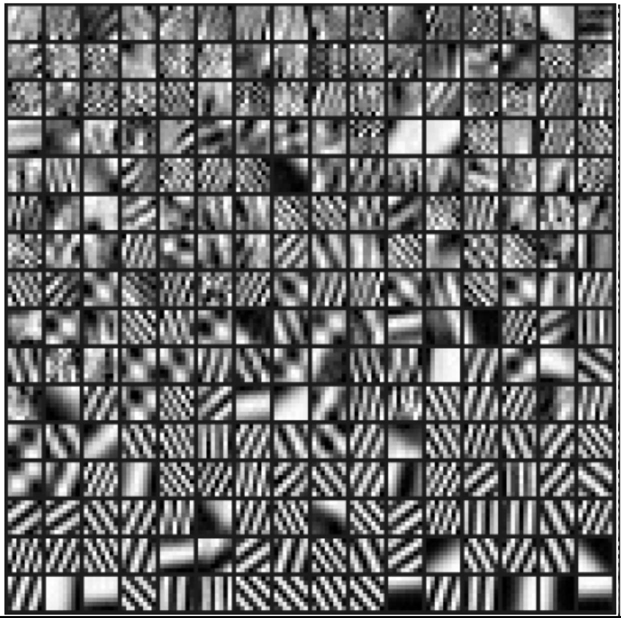
- Similar to the impulse response of simple cells neurons in the visual cortical area V1.

Image Denoising

- Training set: patches of the noisy image.



Elad & Aharon



The obtained dictionary after 10 iterations

$\sigma = 20$

- State of the art results: similar to Non-Local Mean.

Denoising Color

- Learning dictionary of color vectors: $\mathcal{D} = \{\phi_p\}_{p \in \Gamma}$

Mairad, Elad, Shapiro

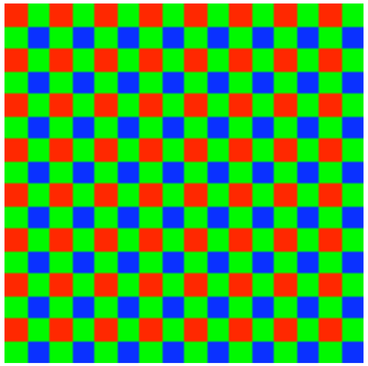
Original

Noisy: 12.8db

Estimated: 29.9db

Demoisaicing & Inpainting

- Demoisaincing: color pixels distributed on a subsampled Bayer grid in camera:



- Inpainting; missing pixels (in color images).
- Super-resolution recovery of color images using the image sparsity in a learned dictionary of color vectors.

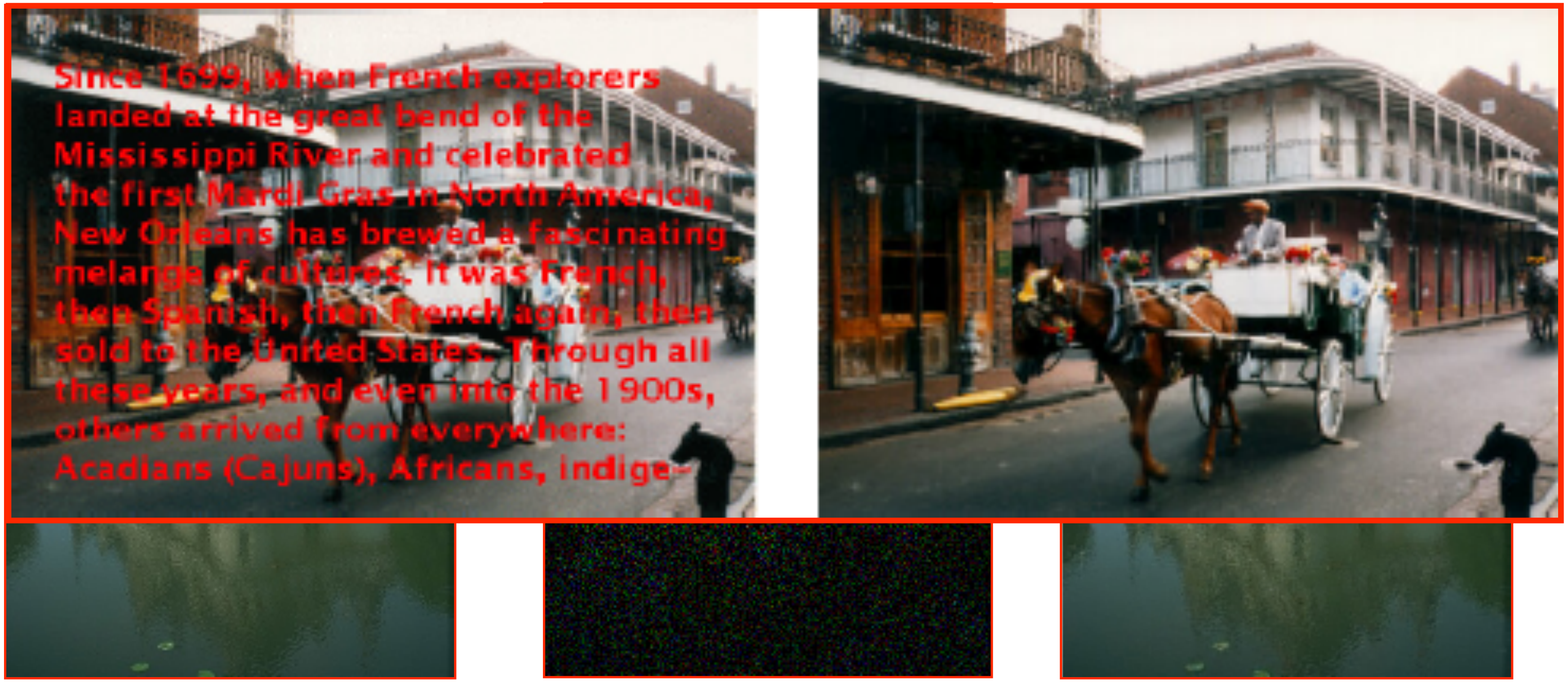
Color Demosaicing

Mairad, Elad, Shapiro

(a) Image 19 restored (b) Zoomed region (c) Image 17 restored (d) Zoomed region

Image Inpainting

Mairad, Elad, Shapiro



Original Missing 80% of the pixels Recovery

Conclusion to Dictionary Learning

- Dictionaries can be adapted to training signal sets with greedy algorithms.

- Efficient approach to build efficient signal models for compression, estimation and pattern recognition, as long as signals are highly compressible.

Blind Source Separation

- Separation of mixte signals from multiple channel measurements:
  - Audio separation of musical instruments in a stereo recording.
  - Electro-cardiogram discrimination of the heart beat of a featus from its mother.
- Blind source separation: recover *S* sources $\{f_s\}_{0 \leq s < S}$ from *K* channel measurements with unknown mixtures

$$Y_k[n] = \sum_{s=0}^{S-1} u_{k,s}\, f_s[n] + W_k[n] \quad \text{for} \quad 0 \leq k < K \ .$$

- If $K < S$, it is a super-resolution inverse problem which recovers $S\,N$ coefficients from $K\,N$.
- Sparse models versus Independent Component Analysis.

Multichannel Decomposition

- Multichannel signal vectors:

$$\vec{Y}[n] = (Y_k[n])_{0 \le k < K} \quad , \quad \vec{u}_s[n] = (u_{k,s}[n])_{0 \le k < K} \ .$$

- Multichannel mixing equation:

$$\vec{Y}[n] = \sum_{s=0}^{S-1} f_s[n] \, \vec{u}_s + \vec{W}[n] \ .$$

- Projection on a dictionary $\mathcal{D} = \{\phi_p\}_{p \in \Gamma}$

$$\langle \vec{Y}, \phi_p \rangle = \sum_{s=0}^{S-1} \langle f_s, \phi_p \rangle \, \vec{u}_s + \langle \vec{W}, \phi_p \rangle$$

with $\langle \vec{Y}, \phi_p \rangle = (\langle Y_k, \phi_p \rangle)_{0 \le k < K} \ .$

Sparse Support Separation

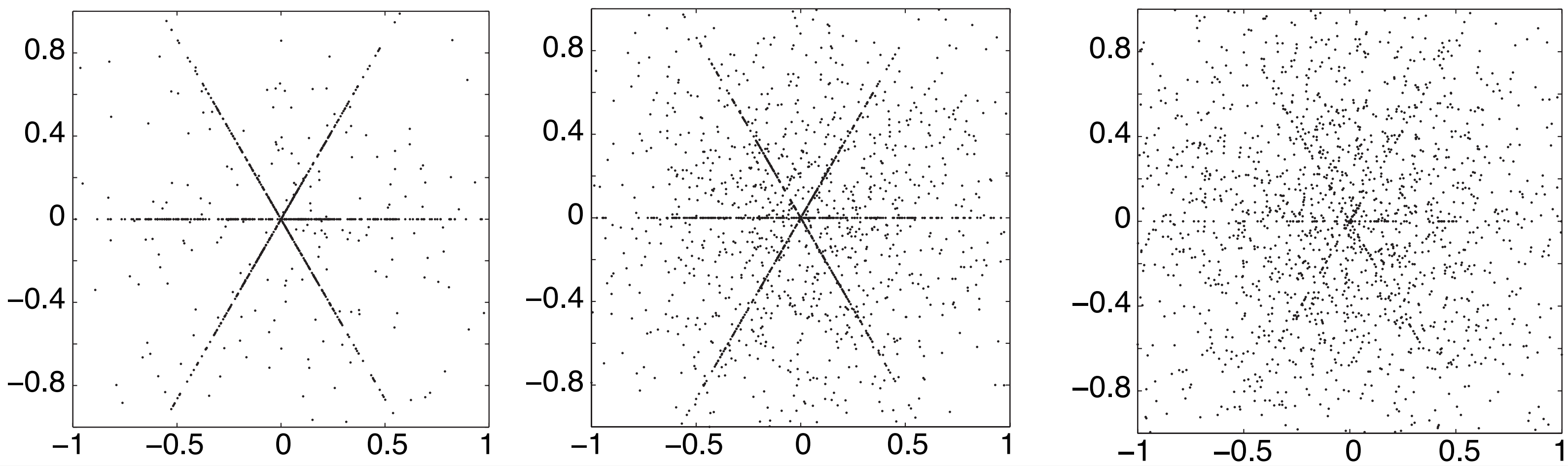
$$\langle \vec{Y}, \phi_p \rangle = \sum_{s=0}^{S-1} \langle f_s, \phi_p \rangle \, \vec{u}_s + \langle \vec{W}, \phi_p \rangle$$

- If the source representation $\{\langle f_s, \phi_p \rangle\}_{p \in \Gamma}$ is sparse, for any p it is likely that $|\langle f_s, \phi_p \rangle|$ is large for at most one s:

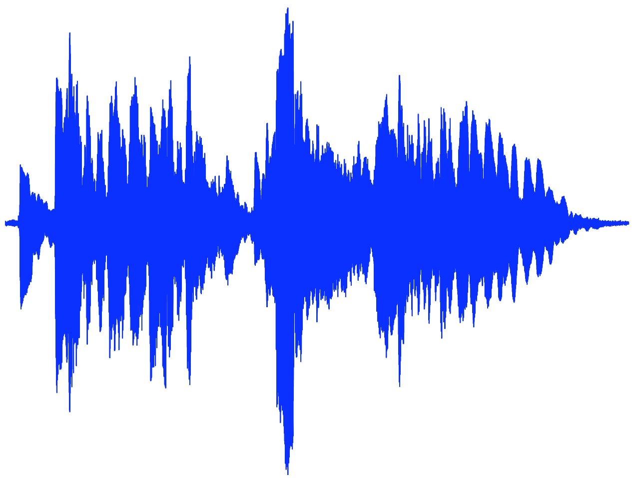
$$\langle \vec{Y}, \phi_p \rangle \approx \langle f_s, \phi_p \rangle \, \vec{u}_s + \langle W, \phi_p \rangle \;, \quad \text{so} \quad \frac{\langle \vec{Y}, \phi_p \rangle}{\| \langle \vec{Y}, \phi_p \rangle \|} = \vec{u}_s + \vec{\epsilon} \,.$$

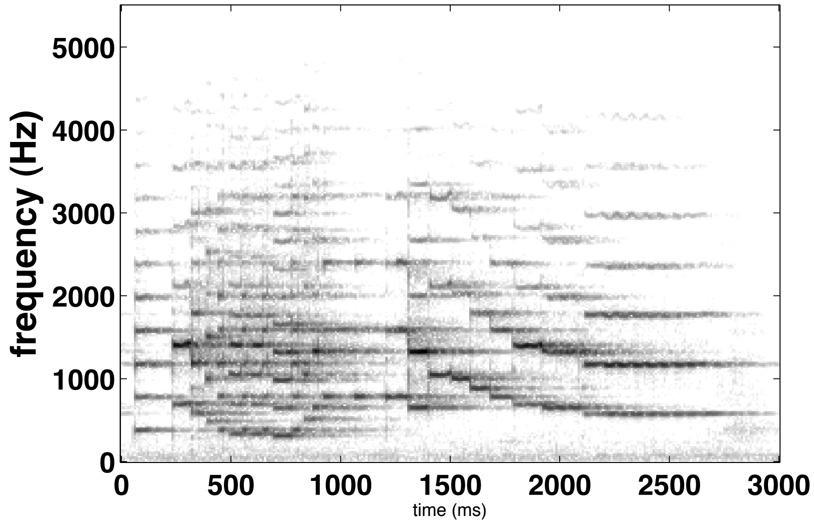
$M/N = 0.1$ $M/N = 0.4$ $M/N = 0.7$

$K = 2$
$S = 3$



Audio Log Spectrogram

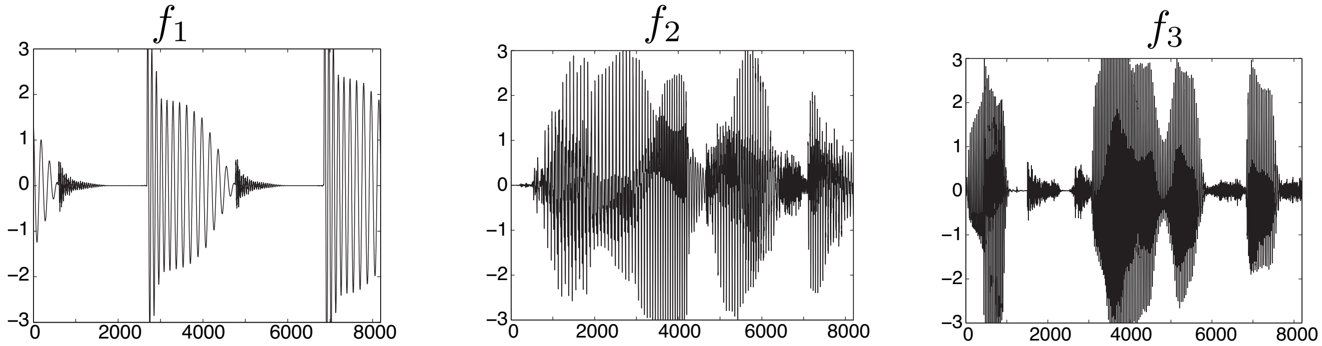




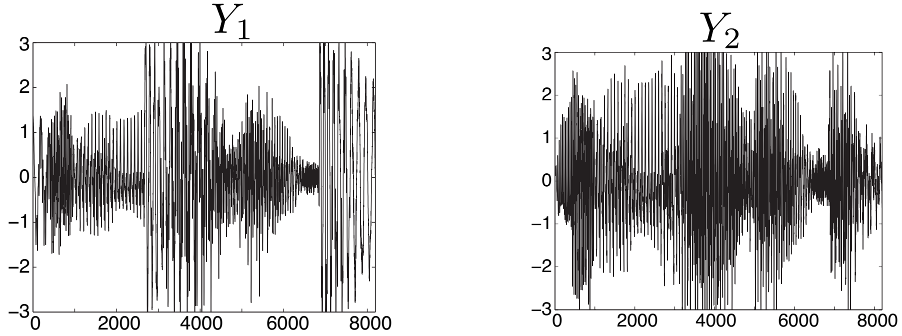
Identification of Mixing Directions

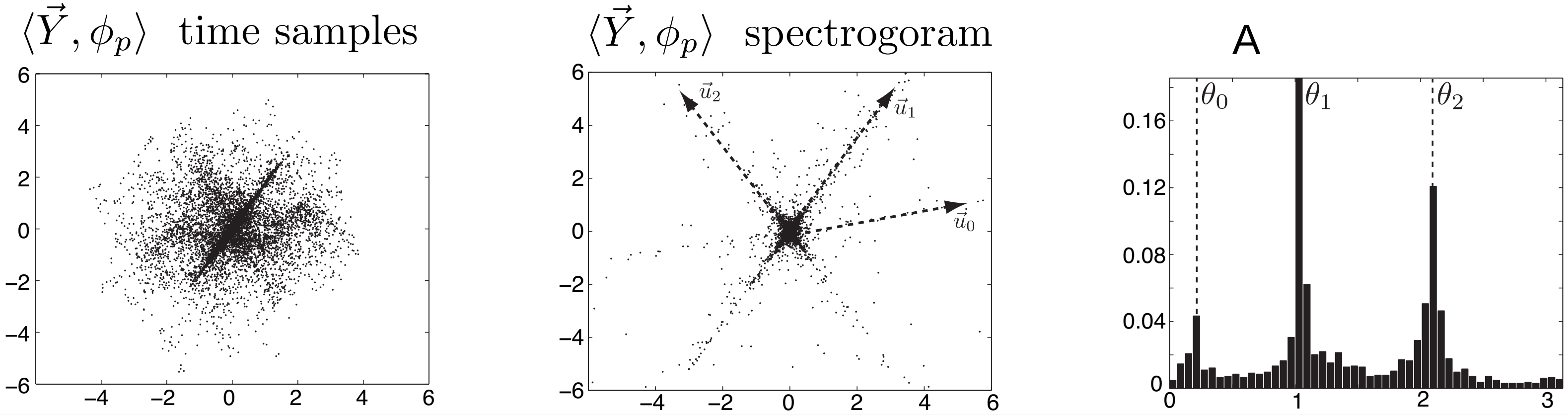
- Mixing directions $\vec{u}_s/\|\vec{u}_s\|$ are identified by voting:

Sources



Multichannel Measures





Source Separation

$$\langle \vec{Y}, \phi_p \rangle = \sum_{s=0}^{S-1} \langle f_s, \phi_p \rangle \, \vec{u}_s + \langle \vec{W}, \phi_p \rangle$$

- If $K < S$ there are less equations then unknown.

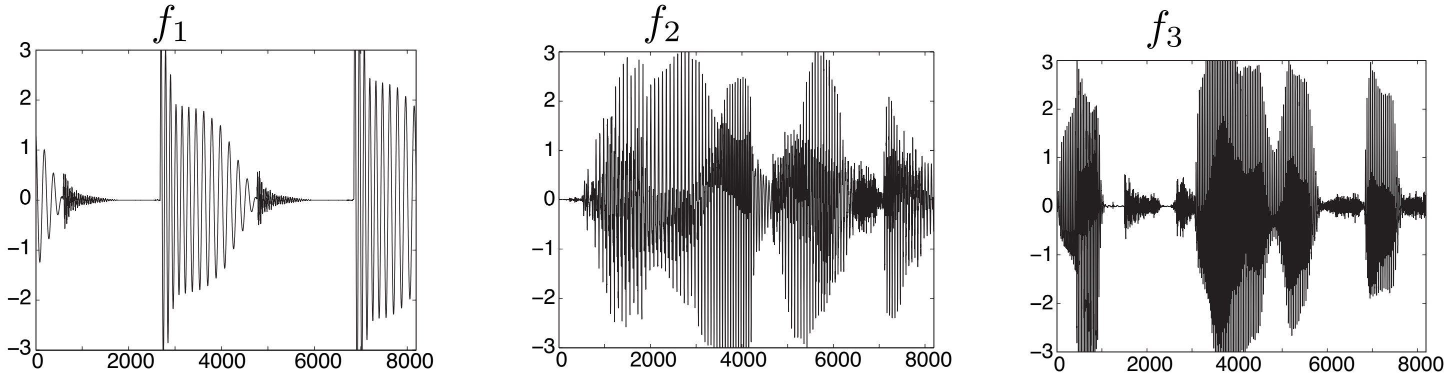
- For a given p there are few large $|\langle f_s, \phi_p \rangle|$
- Sparse decomposition of $\langle \vec{Y}, \phi_p \rangle$ in $\mathcal{D} = \{\vec{u}_s\}_{0 \le s < S}$
- Performed with an orthogonal matching pursuit:

$$\langle \vec{Y}, \phi_p \rangle = \sum_{s=0}^{S-1} \tilde{a}_s[p] \, \vec{u}_s \ .$$

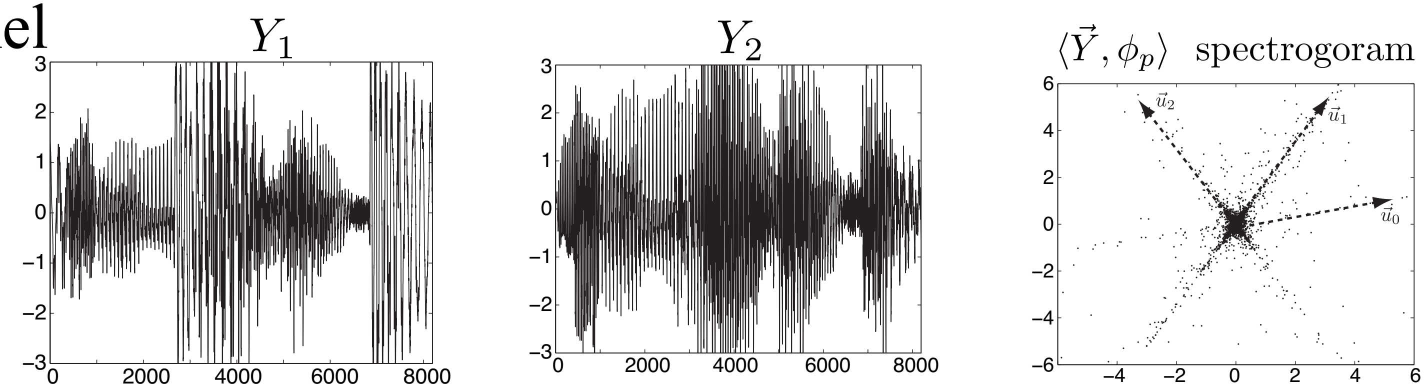
Source estimators: $\tilde{F}_s = \sum_{p \in \Gamma} \tilde{a}_s[p] \, \phi_p \ .$

Source Separation Example

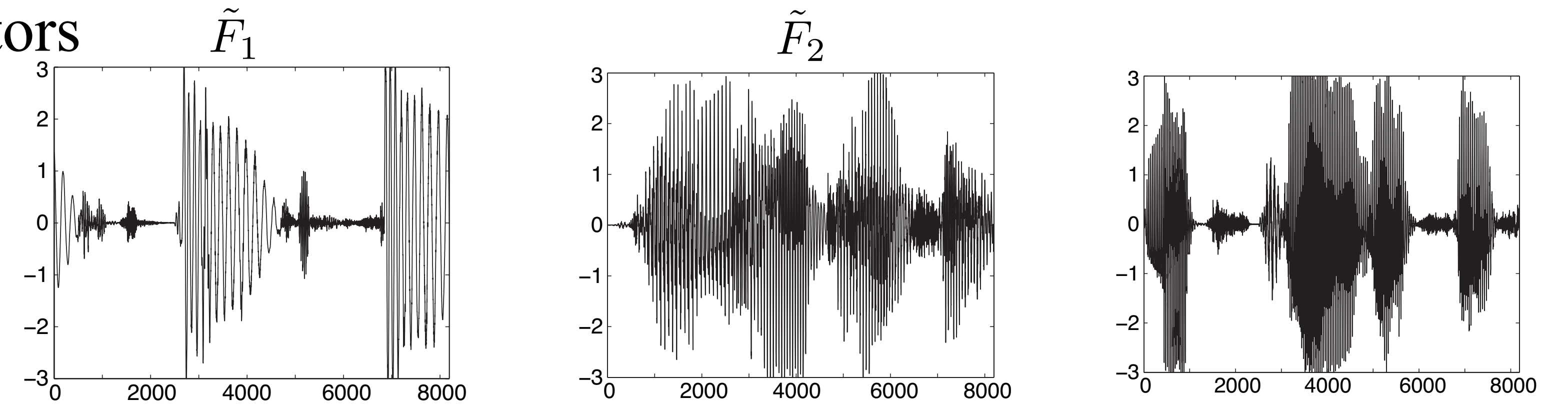
Sources



Multichannel Measures



Estimators



Conclusion to Source Separation

- Sparsity seems more effective then the concept of independence for separating signals.
- Improvements requires refined signal models that do not just rely on their sparsity.

Conclusion to Sparse Approximations

- Looking for sparse approximations is highly powerful to build effective signal models and solve low-level signal processing problems with fast algorithms:
  - Compression
  - Denoising
  - Inverse problems: with or without super-resolution
  - Compressive sensing

- Structured sparsity can further improve results: a current research direction.

- Sparse approximations also apply to classification and pattern recognition, for problems of limited complexity.