Sparse Representations
in
Biology and Signal Processing

Barak A. Pearlmutter

Hamilton Institute
NUI Maynooth
Co. Kildare
Ireland

Outline

Sparse representations ...

- as generic representational trick
- for acoustic source separation
- as model of neuronal sensory processing
- as predictive model of neuronal auditory processing

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^{3} Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km/hour}$

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\text{km}/\text{hour}$
- nerve pulse propagation: $60\,\text{km}/\text{hour}$

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^4 synapses/neuron

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km/hour}$
- nerve pulse propagation: $60\,\mathrm{km/hour}$
- 10^{10} neurons, 10^4 synapses/neuron
- 10^{10} transistors

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km/hour}$
- nerve pulse propagation: $60\,\mathrm{km/hour}$
- 10^{10} neurons, 10^4 synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: $200\,\mathrm{ms}$

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^{3} Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^{4} synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: $200\,\mathrm{ms}$
- Computer time to recognize sneaky animal: never

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^4 synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: 200 ms
- Computer time to recognize sneaky animal: never
- Brain time to tie shoelaces: 4 s

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^{3} Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^{4} synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: 200 ms
- Computer time to recognize sneaky animal: never
- Brain time to tie shoelaces: 4 s
- Computer time to tie shoelaces: never

Intel *vs* The Brain

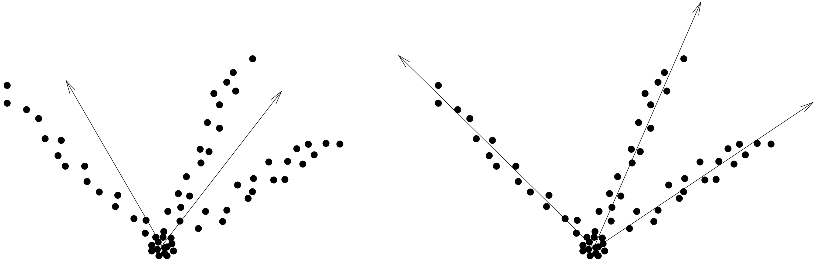
- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^4 synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: $200\,\mathrm{ms}$
- Computer time to recognize sneaky animal: never
- Brain time to tie shoelaces: $4\,\mathrm{s}$
- Computer time to tie shoelaces: never
- Brain sensory inference regime: noise limit of sensors across entire dynamic range and all ecological conditions

Intel *vs* The Brain

- Brain is made of snot (lipids and sugars and protein)
- VLSI transistor switching: 10^{12} Hz
- neuron switching: 10^3 Hz
- VLSI conductor pulse propagation: $100,000,000\,\mathrm{km}/\mathrm{hour}$
- nerve pulse propagation: $60\,\mathrm{km}/\mathrm{hour}$
- 10^{10} neurons, 10^4 synapses/neuron
- 10^{10} transistors
- Brain time to recognize sneaky animal: 200 ms
- Computer time to recognize sneaky animal: never
- Brain time to tie shoelaces: 4 s
- Computer time to tie shoelaces: never
- Brain sensory inference regime: noise limit of sensors across entire dynamic range and all ecological conditions
- Computer sensory inference regime: tuned systems sometimes work in carefully controlled demos

Sparse Representation

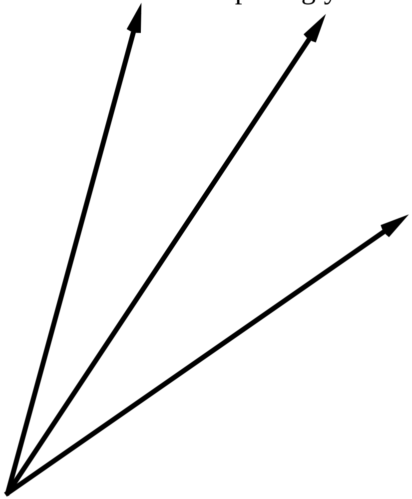
Using an *overcomplete basis* allows a sparseness assumption to capture surprisingly sophisticated statistical structure.



Complete basis (left), overcomplete basis (right).

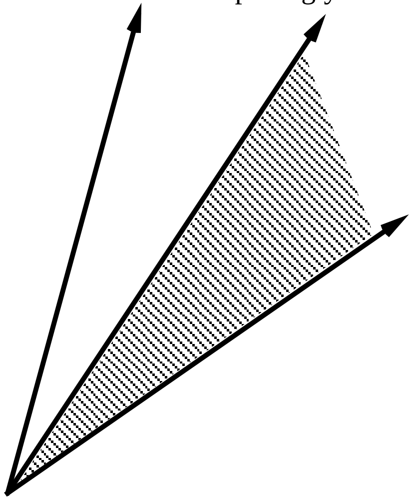
Sparseness in Higher Dimensions

"At most two" sparseness in 3D is surprisingly rich



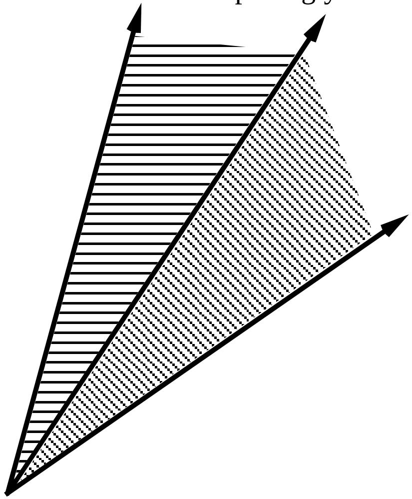
Sparseness in Higher Dimensions

"At most two" sparseness in 3D is surprisingly rich



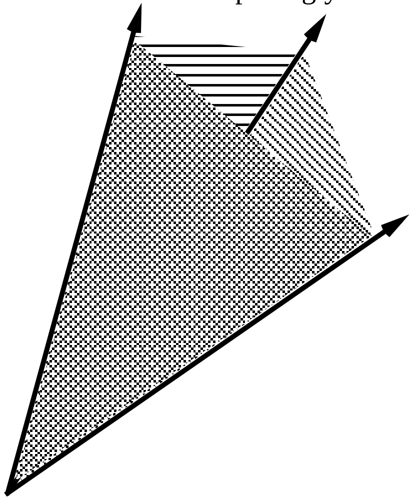
Sparseness in Higher Dimensions

"At most two" sparseness in 3D is surprisingly rich



Sparseness in Higher Dimensions

“At most two” sparseness in 3D is surprisingly rich



Sparseness

Conflicting representational requirements:

- strong prior information
- efficient optimisation

Sparseness

Conflicting representational requirements:

- strong prior information
- efficient optimisation

Fortunately:

- Sparse representations can efficiently represent statistical structure (Simoncelli and Olshausen, 2001).
- Sparse representations are found in the brain: V1 (Vinje and Gallant, 2000), A1 (DeWeese et al., 2003), *etc.*

Source Separation

$\mathbf{x}(t) = \mathbf{A}(\tau) * \mathbf{s}(t) + \text{noise}$

Think: $\mathbf{s}$ is sources $s_j(t)$
$\mathbf{x}$ is sensors $x_i(t)$
$a_{ij}(\tau)$ is filter $s_j \longmapsto x_i$

Problem: Given $\mathbf{x}$, recover $\mathbf{s}$.

Solution: Prior information; Bayes Rule.

Methods: generally indirect (find inverse map).

This Algorithm

One Microphone.

Direct Method.

This Algorithm

One Microphone.

Direct Method.

Just separation. Not localisation.

This Algorithm

One Microphone.

Direct Method.

Just separation. Not localisation.

Just separation.

This Algorithm

One Microphone.

Direct Method.

Just separation. Not localisation.

Just separation. Not localisation.

This Algorithm

One Microphone.

Direct Method.

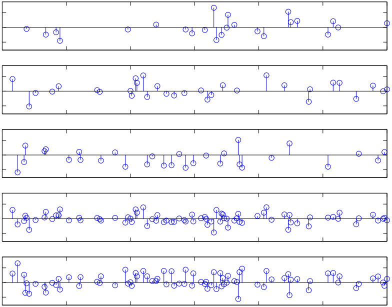
Just separation. Not localisation.

Just separation. Not localisation.

(But looks straightforward to extend...)

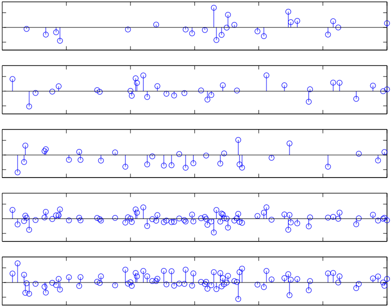
Instantaneous Mixture

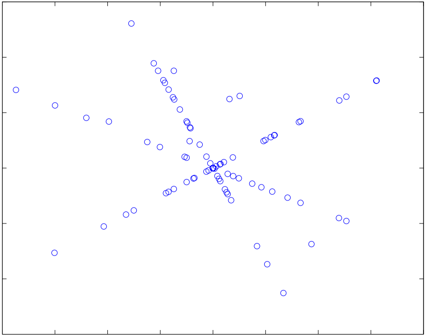
1. Move to basis where signals are sparse.
2. Columns of **A** are apparent (clustering, EM).
3. Sparse decomposition finds coefficients.



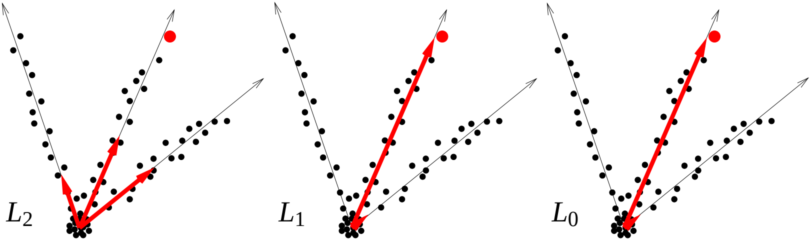
Instantaneous Mixture

1. Move to basis where signals are sparse.
2. Columns of **A** are apparent (clustering, EM).
3. Sparse decomposition finds coefficients.





Coefficients via L_1 Optimisation

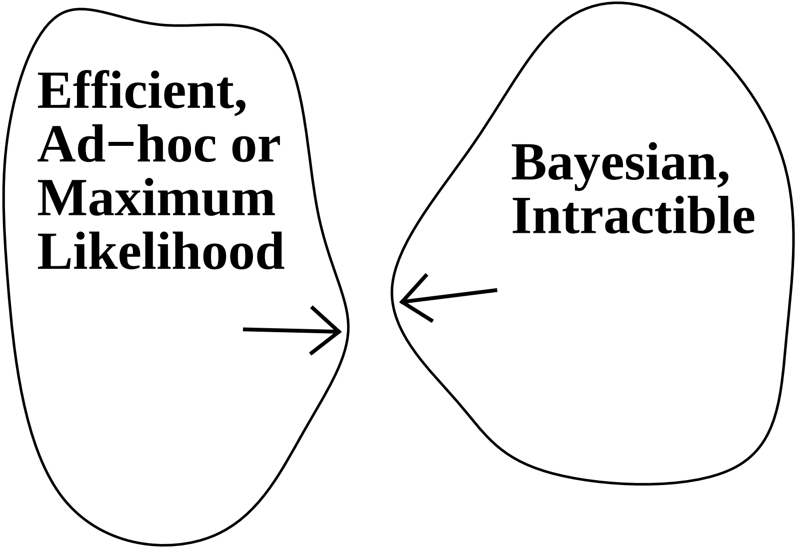


Formulation: $\min_{\mathbf{c}} \|\mathbf{c}\|_p$ subject to $\mathbf{D}\mathbf{c} = \mathbf{x}$ for some L_p

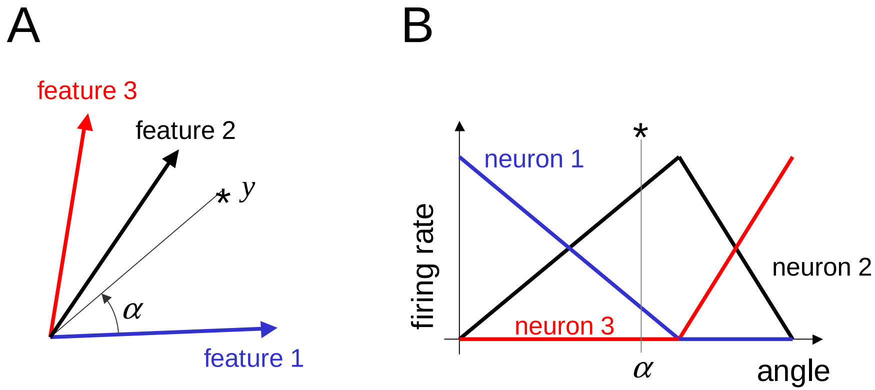
L_1-opt: convex approx. of L_0 (Chen et al., 1998).

L_2: pseudoinverse; L_0: NP-complete; L_1: Linear Programming.

Algorithm Development Philosophies



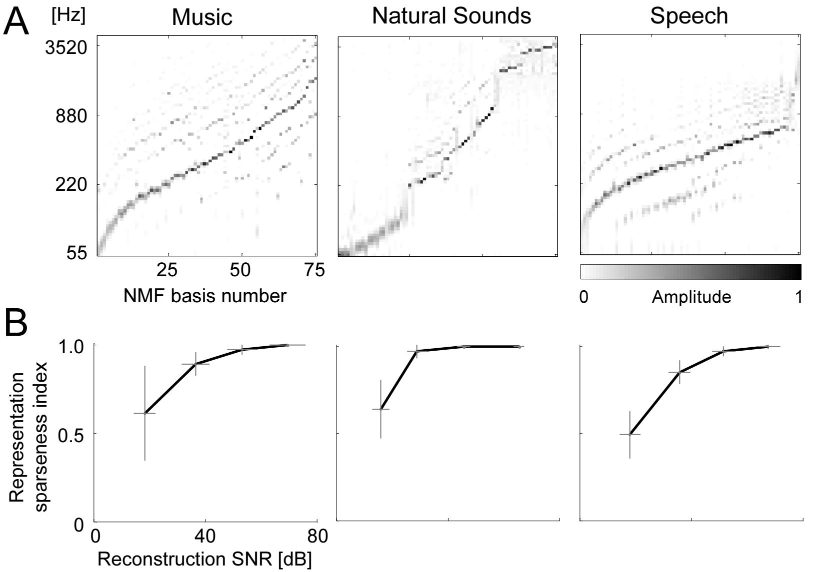
Piecewise Linear Encoding



(A) Encode input data '∗' in overcomplete basis (3 in 2D).

(B) Tuning curves: continuous, piecewise linear.

Finding a Sparse Overcomplete Spectral Basis



Sparse overcomplete spectral basis via NMF, using window size of one frame.

One-Ear Source Separation

Sensor vector becomes scalar:

$$x(t) = \sum_j s_j(t)$$

Hard! Two-microphone tricks don't work.

Known techniques use very strong priors.

- Factorial HMMs (Roweis, 2001)
- Basis/nonlinearity (Hochreiter and Mozer, 2001)
- Per-source basis (Jang and Lee, 2003)

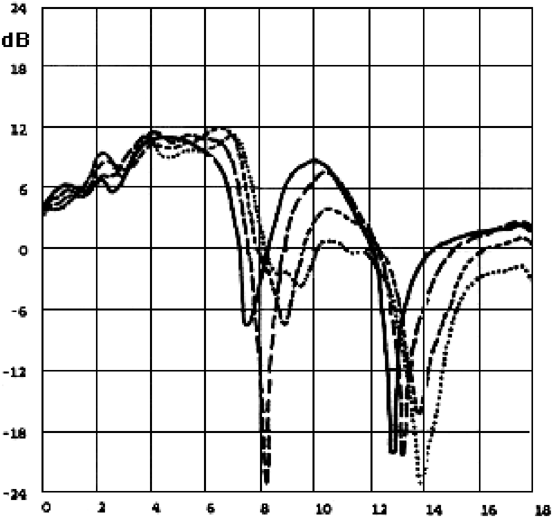
The HRTF

Animals use a variety of binaural and monaural cues for sound localisation, including intra-aural intensity and phase disparity, and monaural spectral filtering by the head and pinnae (head-related transfer function, or HRTF.)

$$x(t) = \sum_j h_j(\tau) * s_j(t)$$

HRTF studied for localisation of single source (Bregman, 1990; Yost et al., 1996) but not for separation of multiple sources.

Actual HRTFs



Different linear filters $h_j(\omega)$ for different angles of arrival.

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.

Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.
Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk} \, d_k(t)$$

Use filtered dictionary $\tilde{d}_{jk}(t) \equiv h_j(t) * d_k(t)$

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.
Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

Use filtered dictionary $\tilde{d}_{jk}(t) \equiv h_j(t) * d_k(t)$
to decompose observations in sensor space:

$$x(t) = \sum_j h_j(t) * \sum_k c_{jk}\, d_k(t)$$

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.
Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

Use filtered dictionary $\tilde{d}_{jk}(t) \equiv h_j(t) * d_k(t)$
to decompose observations in sensor space:

$$x(t) = \sum_j h_j(t) * \sum_k c_{jk}\, d_k(t) = \sum_{jk} c_{jk}\, \tilde{d}_{jk}(t)$$

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.
Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

Use filtered dictionary $\tilde{d}_{jk}(t) \equiv h_j(t) * d_k(t)$
to decompose observations in sensor space:

$$x(t) = \sum_j h_j(t) * \sum_k c_{jk}\, d_k(t) = \sum_{jk} c_{jk}\, \tilde{d}_{jk}(t)$$

$$\mathbf{c} = \arg\max_{\mathbf{c}} \|\mathbf{c}\|_1 \text{ such that } \mathbf{D}\mathbf{c} = \mathbf{x}$$

HRTF-Based Sparse Separation

Monaural source separation with **pinnae**.
Assume sources sparse in signal dictionary:

$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

Use filtered dictionary $\tilde{d}_{jk}(t) \equiv h_j(t) * d_k(t)$
to decompose observations in sensor space:

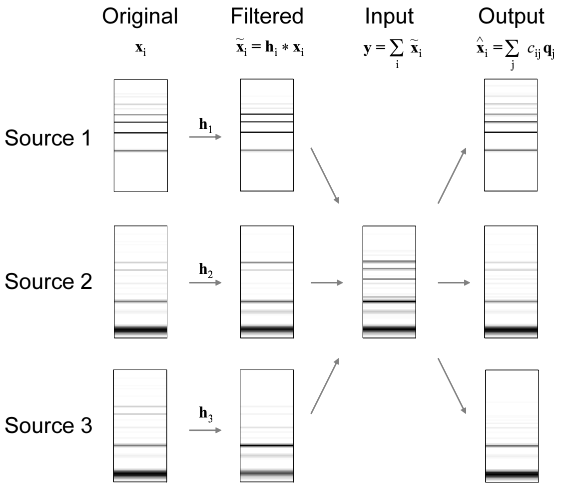
$$x(t) = \sum_j h_j(t) * \sum_k c_{jk}\, d_k(t) = \sum_{jk} c_{jk}\, \tilde{d}_{jk}(t)$$

$$\mathbf{c} = \arg\max_{\mathbf{c}} \|\mathbf{c}\|_1 \text{ such that } \mathbf{Dc} = \mathbf{x}$$

Reconstruct in source space (deconvolution):

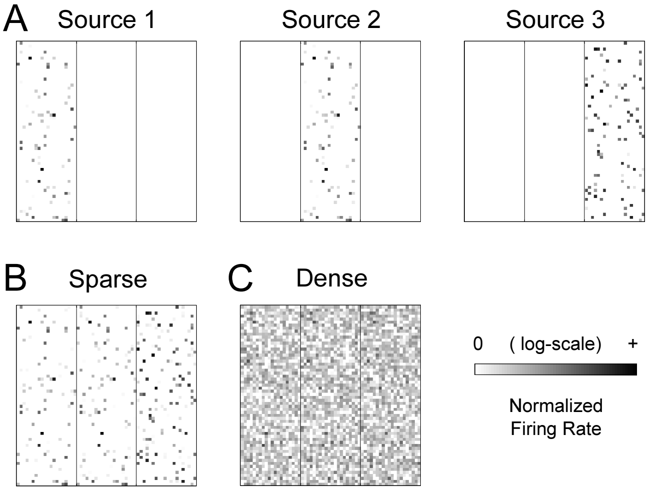
$$s_j(t) = \sum_k c_{jk}\, d_k(t)$$

HRTF-Based Sparse Separation (Cartoon)



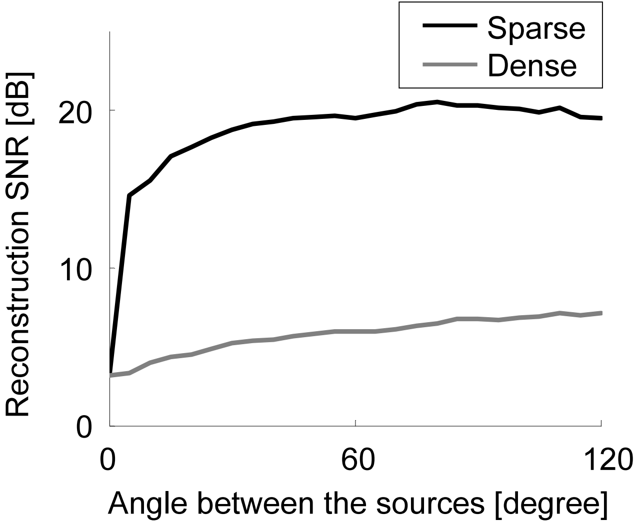
HRTF filters "colour" spectral features, allowing separation by sparse decomposition over basis HRTFs $*$ source dictionary elements.

Sparse *vs* Dense Coefficients



Coefficients recovered with min L_1-norm from three sources processed individually (top), their mixture (bottom left), and using the L_2-norm (bottom right).

Separation Performance



Noise

L_1-norm optimisation sensitive to noise.

Fix by including slop

$$\underset{\mathbf{c}}{\text{minimise}}\ \|\mathbf{c}\|_1 \text{ s.t. } \|\mathbf{Dc} - \mathbf{y}\|_p \leq \beta$$

Noise level is β, norm is $p = 1$, 2, or ∞.

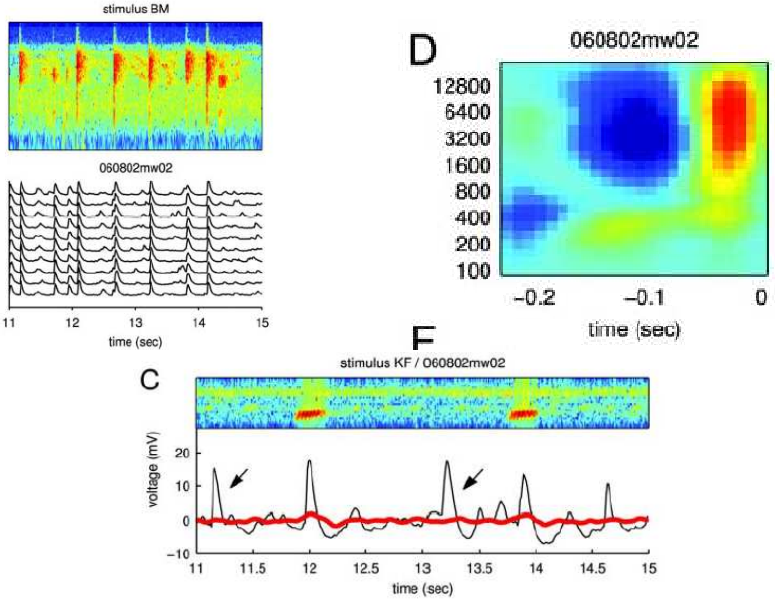
Gaussian ($p = 2$) using L_1+L_2, conic, or SDP.

Solve $p = 1$ or ∞ using linear programming.

Flexible Framework

- instantaneous and convolutive mixing
- binaural information (multiple ears)
- per-source dictionaries
- noise
- temporal structure on coefficients
- convolutive dictionaries
- adaptive signal dictionaries/filters
- room for advances in optimisation
- stronger prior? bigger dictionary!
- **contact with biology**

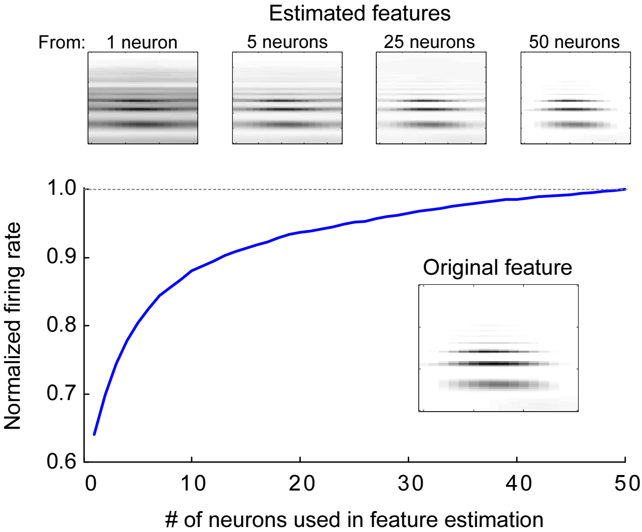
Physiology


stimulus BM
060802mw02
11
12
13
14
15
time (sec)
D
060802mw02
12800
6400
3200
1600
800
400
200
100
−0.2
−0.1
0
time (sec)
F
stimulus KF / 060802mw02
C
voltage (mV)
20
10
0
−10
11
11.5
12
12.5
13
13.5
14
14.5
15
time (sec)

Predictions

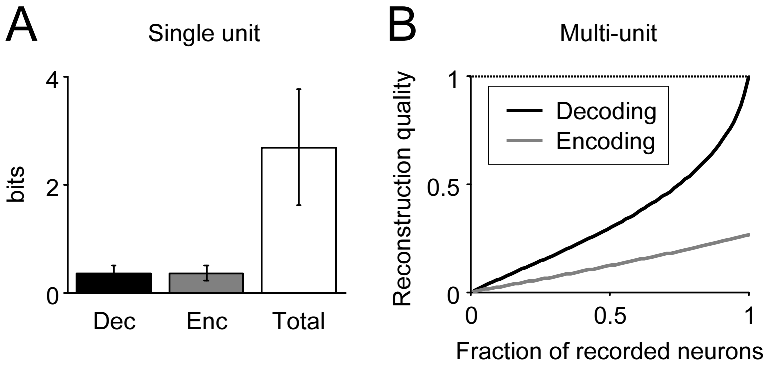
- Sparse activity
- **$>$#Neurons $\rightarrow$ Better Linear Stimulus Optimisation**
- **Linear decoding outperforms linear encoding (in bits)**
- **STRF depends on context**
- $\mathbf{r}(k\,\mathbf{x}) = k\,\mathbf{r}(\mathbf{x})$
- $\|\mathbf{r}(\mathbf{x}_1 + \mathbf{x}_2)\|_1 \leq \|\mathbf{r}(\mathbf{x}_1) + \mathbf{r}(\mathbf{x}_2)\|_1$
- Response shifts opposed to HRTF changes

Stimulus Optimisation



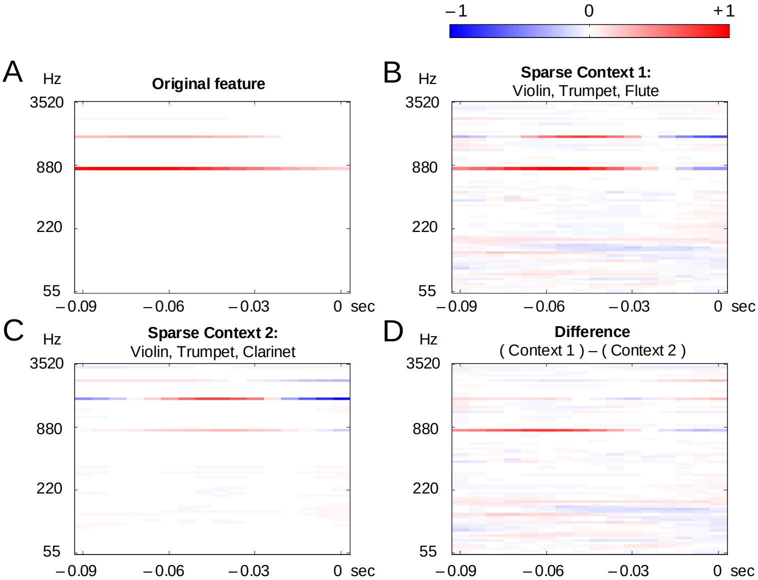
Ability of estimated stimulus to drive neuron improves with #neurons.
"Optimal feature" estimation requires multi-neuron recordings.

Linear Decoding vs Nonlinear Encoding



Linear decoding outperforms linear encoding for multi-neuron but not single neuron experiments.

Context-Dependence of STRF



Summary

Sparse representations

- ... were originally intended for efficient coding and learning, but can also support fast algorithms.
- ... are used in state-of-the-art signal processing.
- ... can make novel neuronal *pre*dictions.

Summary

Sparse representations

- ... were originally intended for efficient coding and learning, but can also support fast algorithms.
- ... are used in state-of-the-art signal processing.
- ... can make novel neuronal *pre*dictions.

Other domains:

- Control (*e.g.* switching)
- Statistics (*e.g.* sparse weights, shrinkage)
- Bayesian reasoning (*e.g.* explaining away)

Collaborators

- Anthony M. Zador
- Hiroki Asari
- Rasmus Olsson

References I

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, Massachusetts.

Chen, S. S., Donoho, D. L., and Saunders, M. A. (1998). Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61.

Deweese, M. R., Hromadka, T., and Zador, A. M. (2005). Reliability and representational bandwidth in the auditory cortex. *Neuron*, 48:479–488.

DeWeese, M. R., Wehr, M., and Zador, A. M. (2003). Binary spiking in auditory cortex. *J. Neurosci.*, 23(21):7940–7949.

Hochreiter, S. and Mozer, M. C. (2001). Monaural separation and classification of mixed signals: A support-vector regression perspective. In *3rd International Conference on Independent Component Analysis and Blind Signal Separation*, San Diego, CA.

Jang, G.-J. and Lee, T.-W. (2003). A maximum likelihood approach to single-channel source separation. *J. of Mach. Learn. Research*, 4:1365–1392.

O'Grady, P. D. and Pearlmutter, B. A. (2004). The LOST algorithms: Exploiting sparseness in the separation of linear subspaces. *EURASIP Journal on Signal Processing*. In press.

Pearlmutter, B. A., Asari, H., and Zador, A. M. (2005). Neuronal predictions of sparse linear representations. In *Forum Acusticum 2005*, Budapest, Hungary.

Pearlmutter, B. A. and Zador, A. M. (2004). Monaural source separation using spectral cues. In *Fifth International Conference on Independent Component Analysis*, LNCS 3195, pages 478–485, Granada, Spain. Springer-Verlag.

References II

Rickard, S. T. and Dietrich, F. (2000). DOA estimation of many W-disjoint orthogonal sources from two mixtures using DUET. In *Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing (SSAP2000)*, pages 311–314, Pocono Manor, PA.

Roweis, S. T. (2001). One microphone source separation. In *Adv. in Neu. Info. Proc. Sys. 13*, pages 793–799. MIT Press.

Simoncelli, E. P. and Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annu. Rev. Neurosci.*, 24(1):1193–1216.

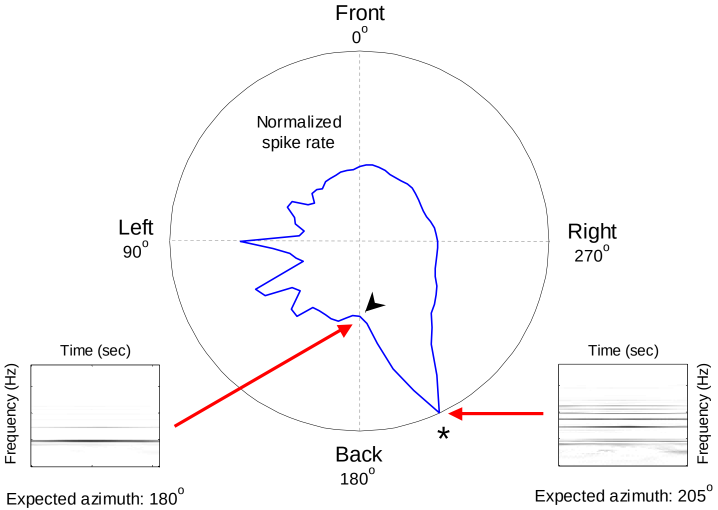
Tyler, M., Danilov, Y., and Bach-y Rita, P. (2003). Closing an open-loop control system: Vestibular substitution through the tongue. *Journal of Integrative Neuroscience*, 2(2):159–164.

Vinje, W. E. and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276.

Yost, W. A., Dye, Jr., R. H., and Sheft, S. (1996). A simulated "cocktail party" with up to three sound sources. *Percept Psychophys*, 58(7):1026–1036.

Zibulevsky, M. and Pearlmutter, B. A. (2001). Blind source separation by sparse decomposition in a signal dictionary. *Neu. Comp.*, 13(4):863–882.

Top-Down Neuronal Modulation



Normalised response of coefficient tuned to cello feature at 205° azimuth (‘∗’).

Sensory Substitution/Augmentation

Present sensory data using somatosensory modality: electrode array on tongue (Tyler et al., 2003).

Sensory Substitution/Augmentation

Present sensory data using somatosensory modality: electrode array on tongue (Tyler et al., 2003).

Potential applications: oscillopsia, blindness, hearing loss, new senses (magnetic, radar), data visualisation.

Need Good Signal Dictionaries

- separation and de-mixing performance depend on dictionary
- need dictionary optimisation algorithms for:
  - sparseness
  - separation
  - perceptual quality with lossy encoder
  - speech, music, video
  - particular acoustic sources
  - rapid dictionary adaptation

Current Dictionary Work

Goal: automatically find optimal dictionaries.

- NMF on spectra (Hiro Asari)
- Subspace clustering (O'Grady and Pearlmutter, 2004)
- Alg. Diff. of LP (Vamsi Potluru):
  $\nabla_{\mathbf{D}} \langle \| \mathsf{L1opt}(\mathbf{s}, \mathbf{D}) \|_1 \rangle$
  $\nabla_{\mathbf{D}} \| \mathsf{sparse\_separate}(\mathbf{A} * \mathbf{s}, \mathbf{D}) - \mathbf{s} \|_2^2$

Distant future: joint estimation of sources, HRTF, acoustic environment; adaptive source models.

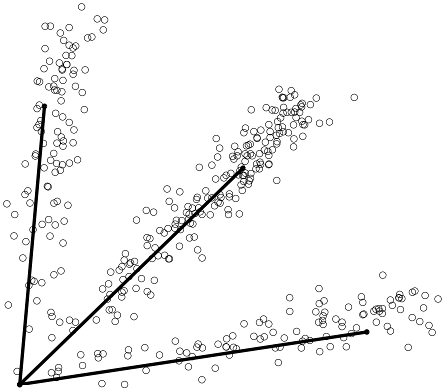
AD of Linear Programming

$$\begin{aligned}
\text{L1opt}(\mathbf{y}, \mathbf{D}) &= \begin{bmatrix}\mathbf{I} & -\mathbf{I}\end{bmatrix} \text{lp}\left(\mathbf{1}, -\mathbf{I}, \mathbf{0}, \mathbf{D}\begin{bmatrix}\mathbf{I} & -\mathbf{I}\end{bmatrix}, \mathbf{y}\right) \\
\text{lp}(\mathbf{w}, \mathbf{A}, \mathbf{a}, \mathbf{B}, \mathbf{b}) &= \arg\min_{\mathbf{z}} \ \mathbf{w}^\top \mathbf{z} \text{ s.t. } \mathbf{A}\mathbf{z} \le \mathbf{a} \ \& \ \mathbf{B}\mathbf{z} = \mathbf{b} \\
&= \text{lq}\left(\begin{bmatrix}\mathbf{P}_\alpha & 0 \\ 0 & \mathbf{P}_\beta\end{bmatrix}\begin{bmatrix}\mathbf{A} \\ \mathbf{B}\end{bmatrix}, \begin{bmatrix}\mathbf{P}_\alpha & 0 \\ 0 & \mathbf{P}_\beta\end{bmatrix}\begin{bmatrix}\mathbf{a} \\ \mathbf{b}\end{bmatrix}\right) \\
\text{lq}(\mathbf{M}, \mathbf{m}) &= \mathbf{z} \quad \text{s.t. } \ \mathbf{M}\mathbf{z} = \mathbf{m} \\
\overleftarrow{\text{lq}}(\mathbf{M}, \mathbf{m}, \mathbf{z}, \grave{\mathbf{z}}) &= \begin{bmatrix}\grave{\mathbf{M}} & \grave{\mathbf{m}}\end{bmatrix} \\
\grave{\mathbf{m}} &= \text{lq}(\mathbf{M}^\top, \grave{\mathbf{z}}) \\
\grave{\mathbf{M}} &= -\grave{\mathbf{m}}\mathbf{z}^\top
\end{aligned}$$

Dictionary Optimisation

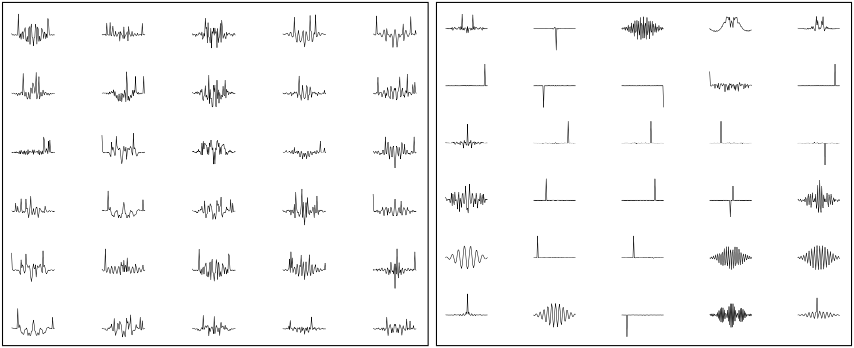
1. Draw $\mathbf{y}$ from signal distribution.
2. Calculate $E_{\mathbf{y}} = \|\text{L1opt}(\mathbf{y}, \mathbf{D})\|_1$
3. Calculate $\nabla_{\mathbf{D}} E_{\mathbf{y}}$ using $\overleftarrow{\text{L1opt}}(\cdot)$.
4. Step $\mathbf{D} := \mathbf{D} - \eta \nabla_{\mathbf{D}} E_{\mathbf{y}}$.
5. Normalise columns $\mathbf{d}_i$ of $\mathbf{D}$ with $\mathbf{d}_i := \mathbf{d}_i / \|\mathbf{d}_i\|_2$.
6. Repeat to convergence of $\mathbf{D}$.

Learnt Dictionary (1)



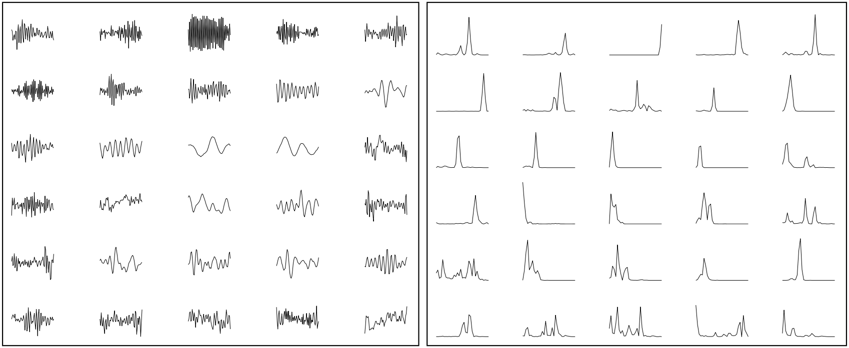
Overcomplete basis for 2D test data.
Noise biases basis towards regions of less coverage; fix with noise term.

Learnt Dictionary (2)



Artificial signal ensemble (left); learnt dictionary elements (right).
Data: sparse combinations of Gabor functions and delta pulses.
$\mathbf{D}$: 64×128.

Learnt Dictionary (3)



Learnt dictionary for speech: William Butler Yeats reciting *Coole Park and Ballylee* and *The Lake Isle of Innisfree*, 8 ms windows, $\mathbf{D} : 64 \times 128$. Dictionary elements (left) and power spectra 0–4 kHz (right).

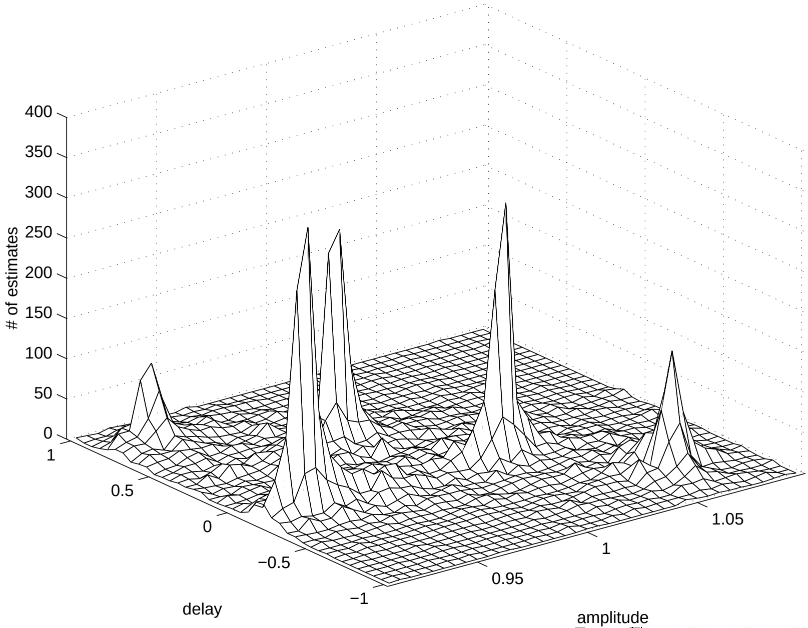
DUET algorithm

Anechoic mixture (Rickard and Dietrich, 2000). Assume narrowband: delays $|\delta_j| < $ wavelength.

$$x_1(t) = \sum_j s_j(t) \qquad x_2(t) = \sum_j a_j\, s_j(t - \delta_j)$$

- Mixing/delay parameters found by clustering in time-frequency domain (short-time DFT).
- Can recover #sources $>$ #mixtures.
- Practical almost-real-time algorithm.

DUET Coefficient Clusters



To Add

block diagram
more cells in cortex than in cochlea