A Causal Locally Competitive Algorithm for the Sparse Decomposition of Audio Signals

Adam Charles, Abbie Kressner & Christopher Rozell
Georgia Institute of Technology

2011 Digital Signal Processing Workshop

Audio Coding

- Standard coding: Fourier/Wavelet
- Modern processing uses sparsity
- Sparse audio decompositions:

$$x(t) = \sum_i \sum_m s_i^m \phi_i(t - \tau_i^m)$$

Dictionary Element

Signal

Amplitude

Time Shift

- Make s_i^m sparse!

Motivation

- Potential applications for sparse inference
  - Audio coding
  - Audio enhancement
  - Hearing aids and cochlear implants
- How can we find s_i^m?

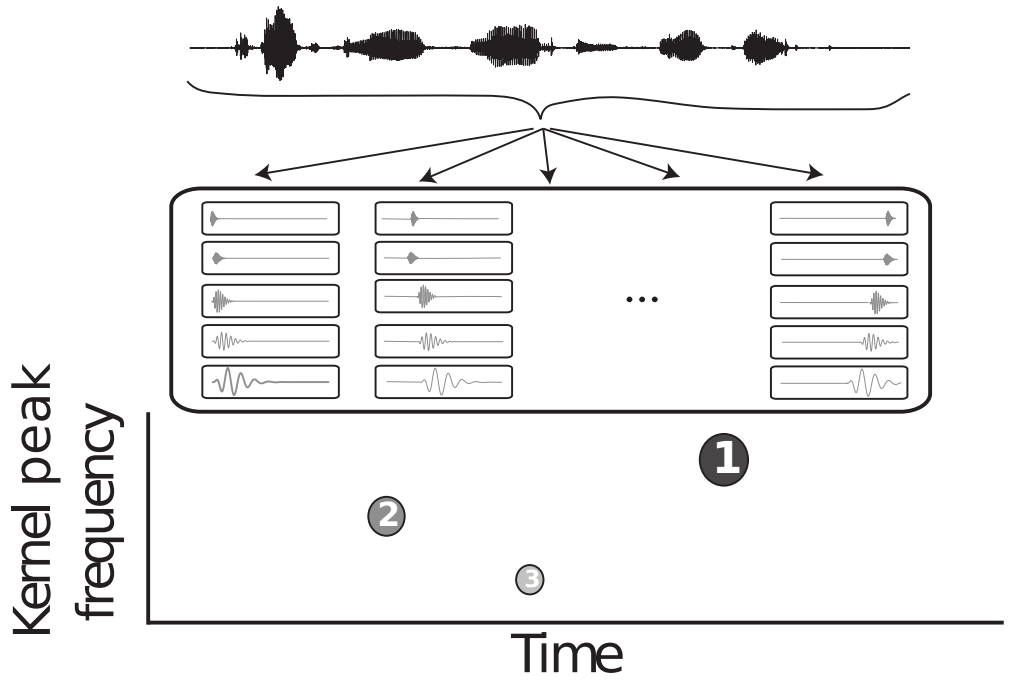
Matching Pursuit

- Vector-Matrix form

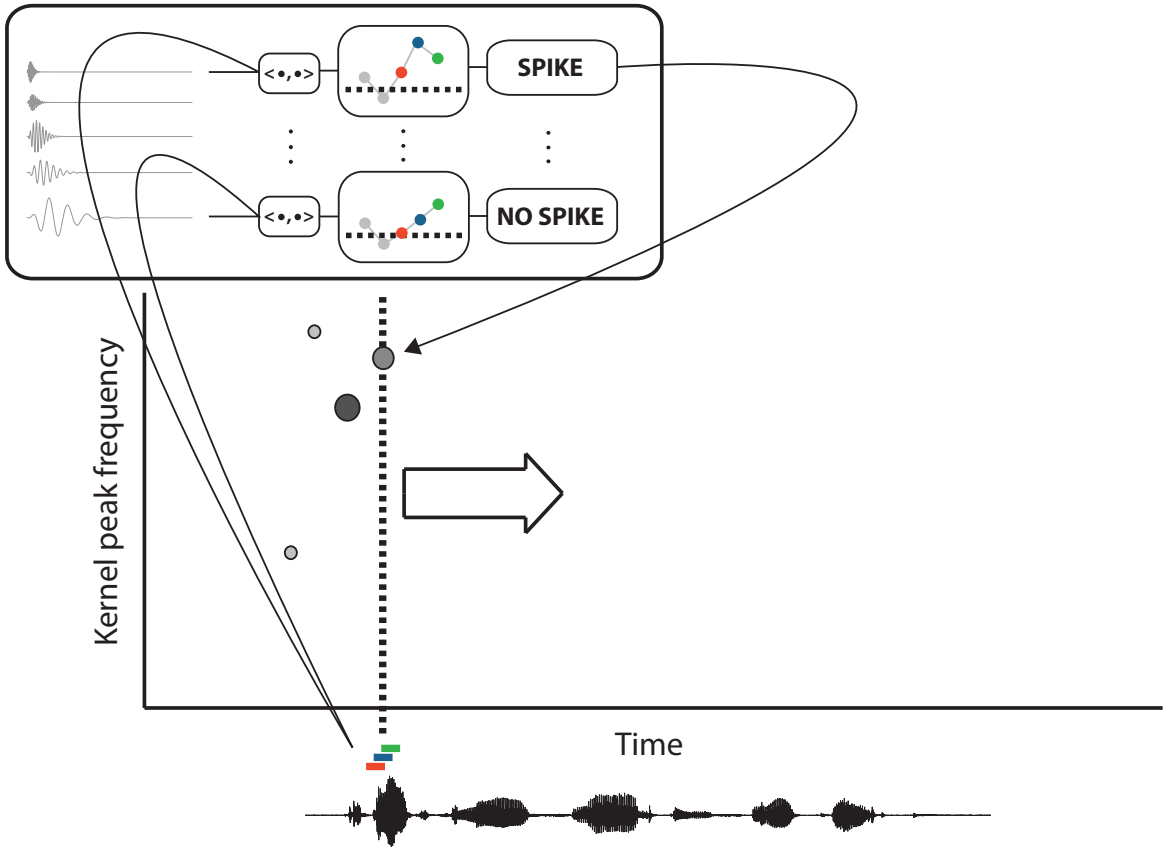
$$\boldsymbol{x} = \boldsymbol{\Phi a}$$

- Algorithm
  - Pick best a_i at time n
  - Calculate the residual
  - Repeat

(Mallat and Zhang 1993)



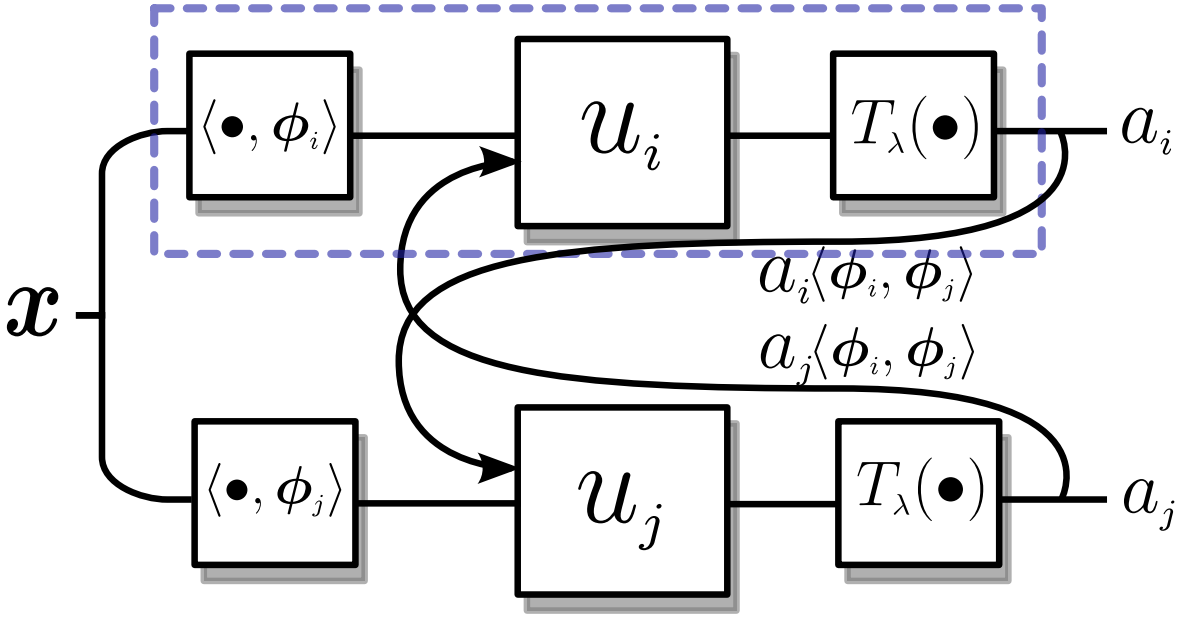
Filter and Threshold



(Smith and Lewicki, 2005)

LCA Structure

1. Use feedback to sparsify outputs while retaining signal integrity:



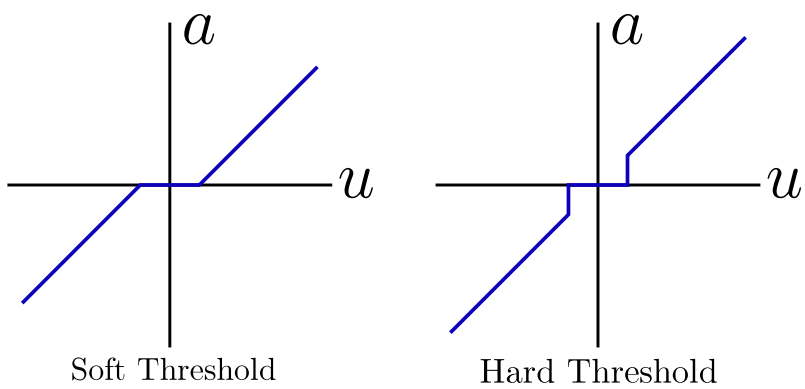
(Rozell et. al. 2008)

LCA Dynamics

$$1) \quad \dot{u}_i(t) = \frac{1}{\tau}\left(\langle \boldsymbol{x}, \boldsymbol{\phi}_i \rangle - u_i(t) - z_i(t)\right)$$

$$\text{where} \quad z_i(t) = \sum_{j \neq i} \langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \rangle a_j(t)$$

$$2) \quad a_i(t) = T_\lambda\left(u_i(t)\right)$$



frametitleCorrelations

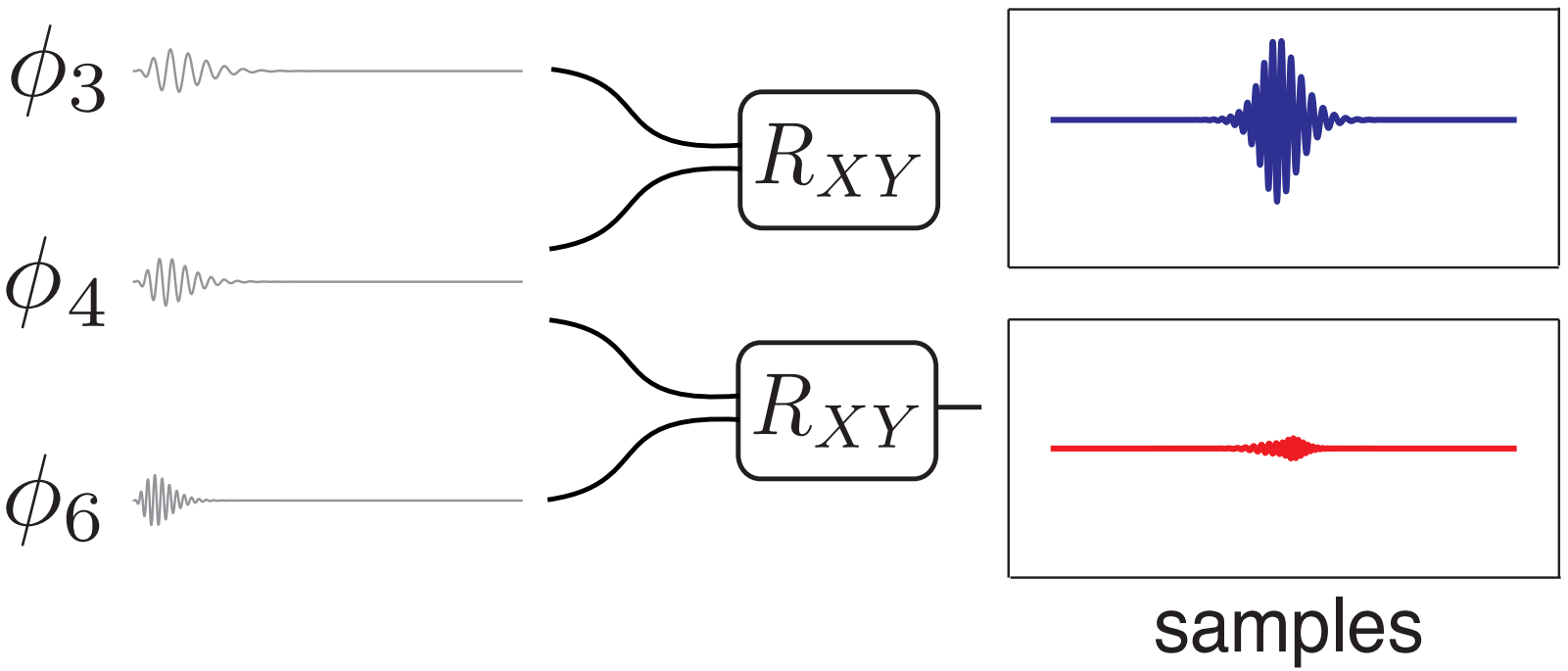


Figure: Basis Correlation Functions for ϕ_3, ϕ_4 and ϕ_6

Buffer

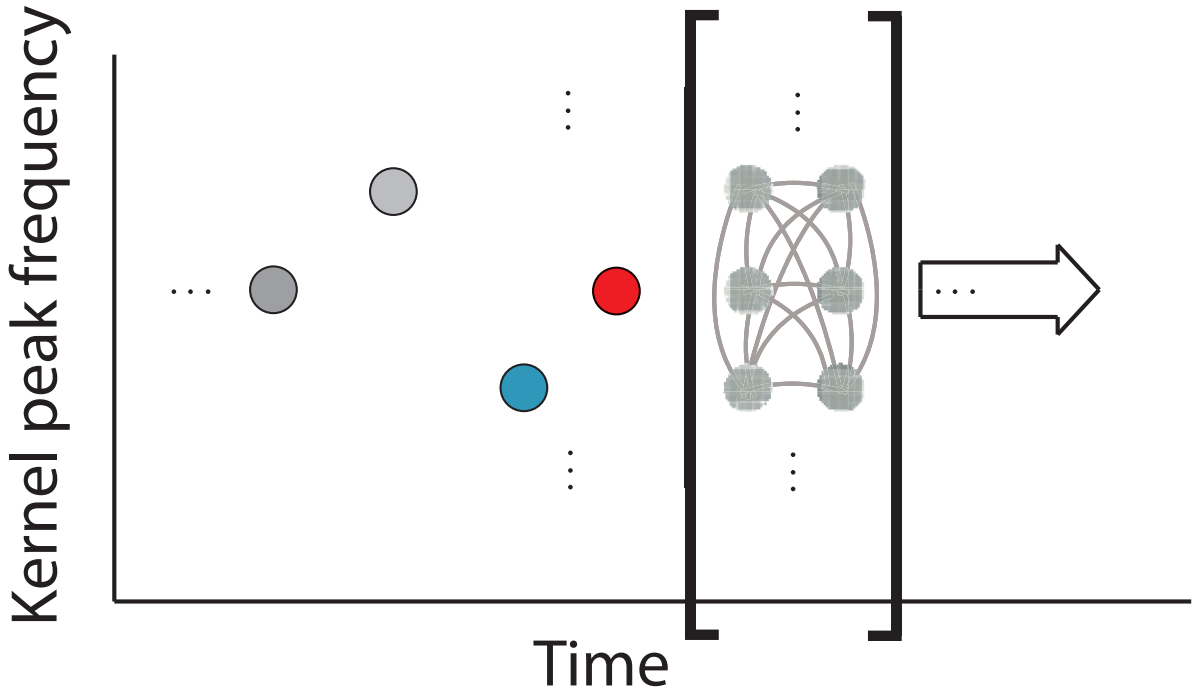


Figure: Track correlations through **space & time**

Causal LCA Architecture

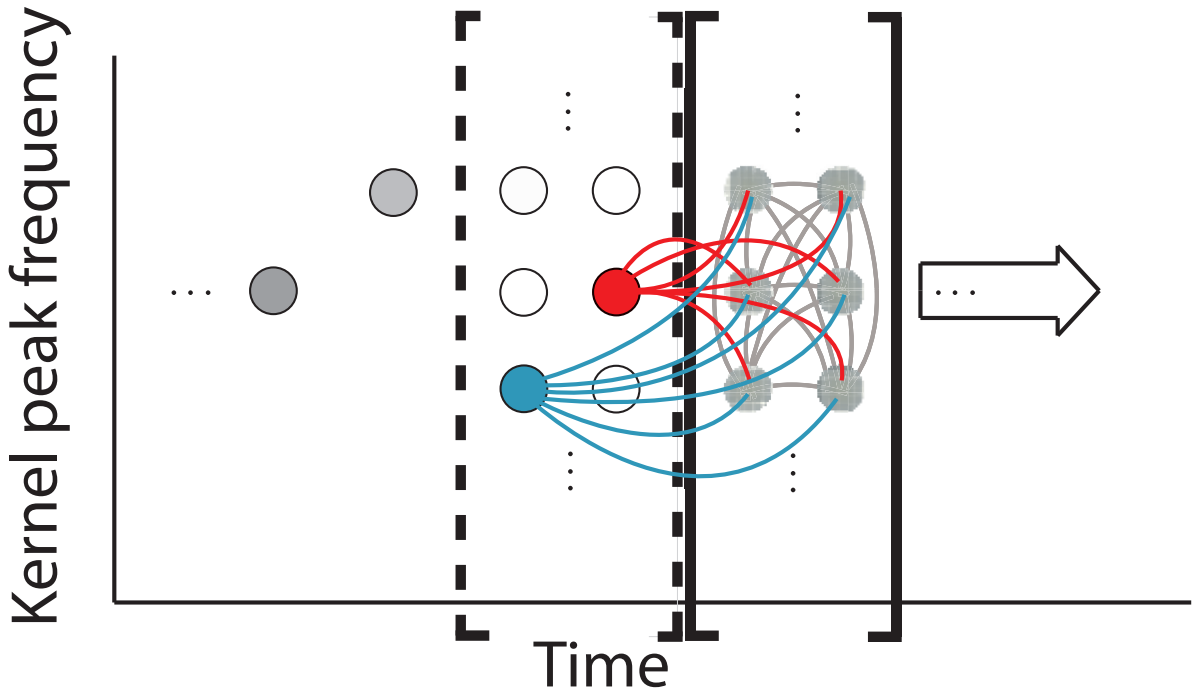


Figure: Recently written coefficients continue inhibiting

CLCA Dynamics

1. Read new sample and move sliding window
2. Allow LCA to converge at time *n*,

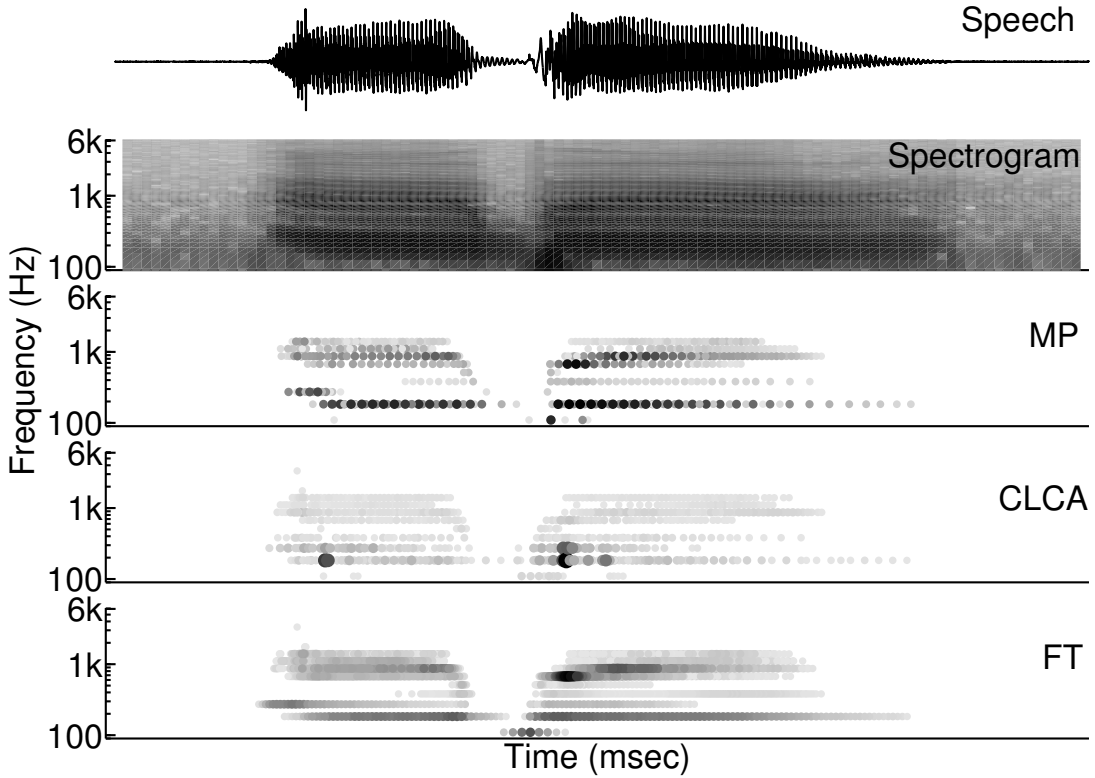
$$2a)\, \dot{u}_i(t) = \frac{1}{\tau}\left(\langle \boldsymbol{x}, \boldsymbol{\phi}_i \rangle - u_i(t) - z_i(t)\right)$$

$$\text{where} \;\; z_i(t) = \overbrace{\sum_k \langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_k \rangle \hat{a}_k}^{\text{already active}} + \overbrace{\sum_{j \neq i} \langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \rangle a_j(t)}^{\text{in buffer}}$$

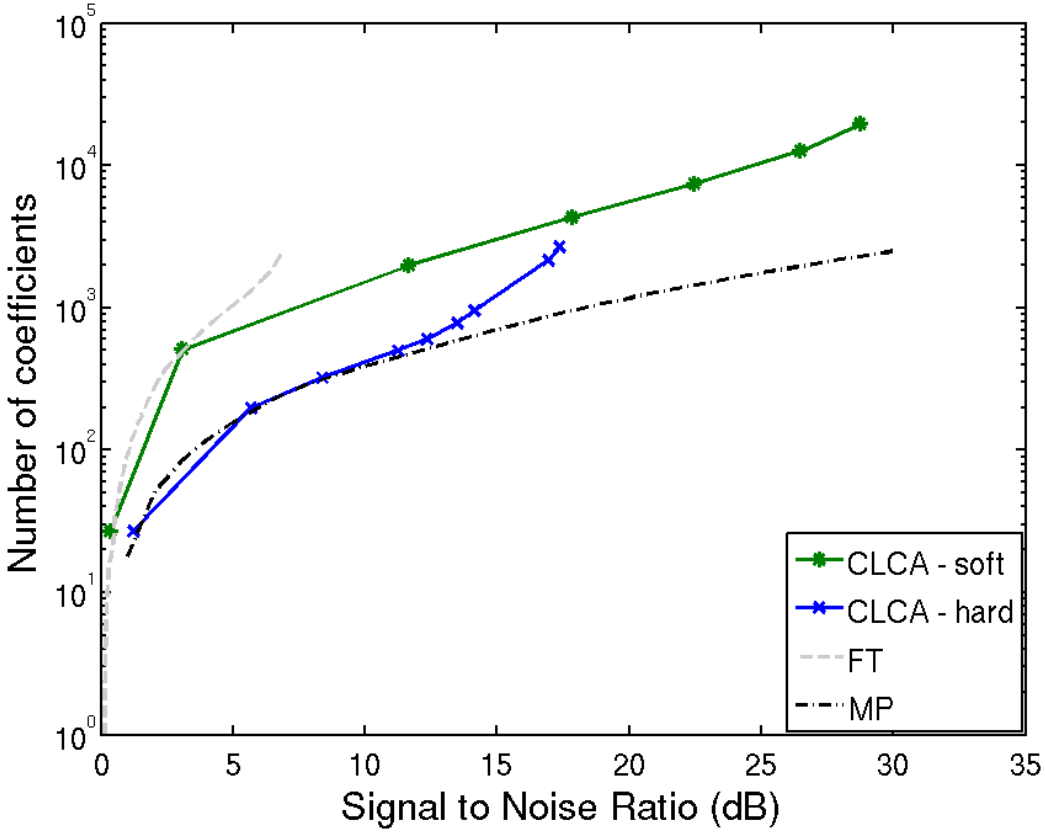
$$2b)\;\; a_i(t) = T_\lambda\left(u_i(t)\right)$$

3. Write last coefficients in the buffer and move all other values back a timestep

Spikegrams


Speech
Spectrogram
MP
CLCA
FT
6k
1k
100
Frequency (Hz)
Time (msec)

Rate Distortion Curve



Conclusions

- Sparsity with causality
- Analog system: low power and real-time (50KHz)
- 10ms window: within lip sync tolerance

Thank you

Q&A

email: acharles6@gatech.edu

Spikegrams

Matching Pursuit
15 db SNR
17 coefficients

Filter and threshold
7 db SNR
84 coefficients

Causal LCA
15 db SNR
162 coefficients