Adaptive Filters and SIMO System Identification

Patrick A. Naylor
Imperial College London

Co-workers: Nick Gaubitch, Uttachai Manmontri, Andy Khong, Rehan Ahmad, Jimi Wen, Xiang (Shawn) Lin

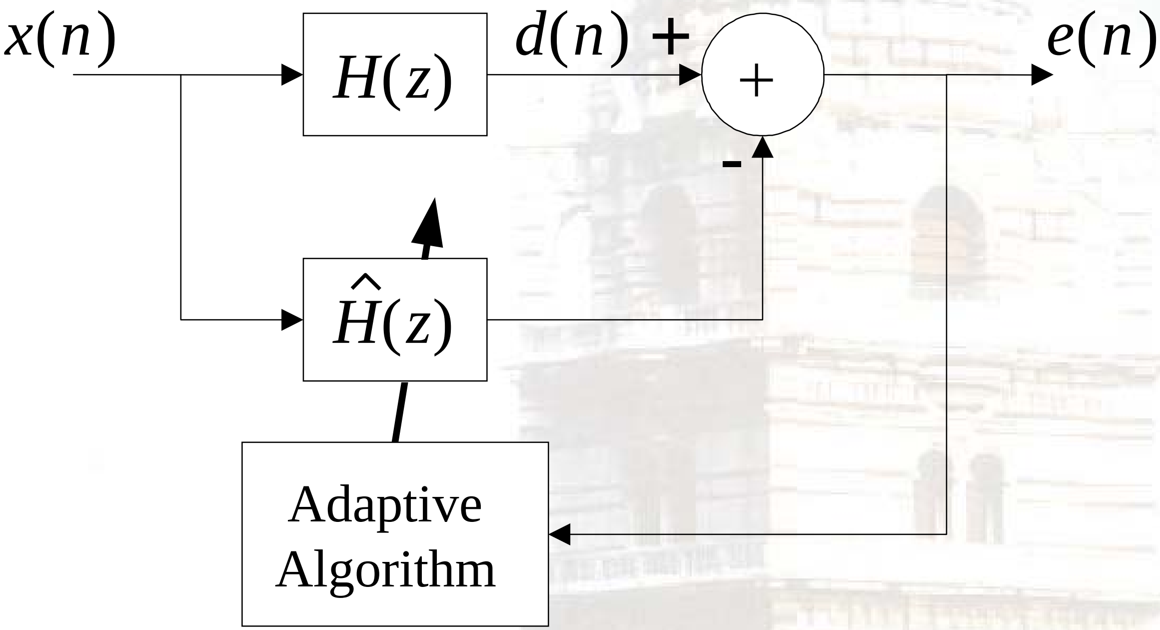
Download Video Version

Contents

- Adaptive signal processing (a DSP perspective)
  - Overview and simple application
  - Classes of techniques
- Supervised algorithms
  - Single channel system identification
- Unsupervised algorithms
  - Single channel
  - Multichannel
- Channel Inversion

Problem Formulation – system identification

- Given input $x(n)$ and output $d(n)$ of an unknown system $H(z)$, estimate a system $\hat{H}(z)$ such that $e(n)$ is minimized.
- The system is often assumed to be FIR
  - Can't go unstable
- The taps of the adaptive filter are adjusted by the adaptive algorithm



Notation

Input sample at sample n: $x(n)$

Tap input vector at sample n: $\mathbf{x}(n) = [x(n), x(n-1), x(n-2), \ldots, x(n-L+1)]^T$

Unknown system: $H(z)$

Estimated system: $\hat{H}(z)$

Coefficient vector at time n: $\mathbf{h}(n) = [h_0(n), h_1(n), \ldots, h_{L-1}(n)]^T$

$\hat{H}(z)$

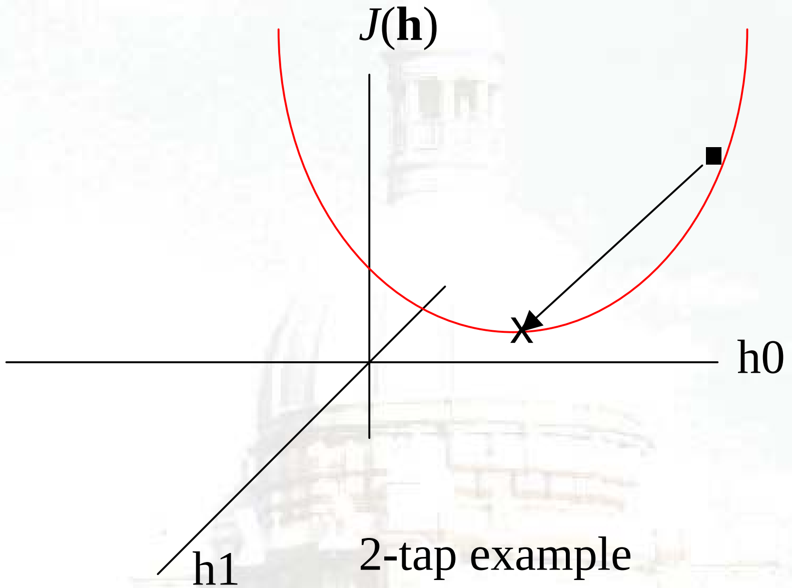
Steepest Descent

- Write the error as:

$$e(n) = d(n) - \sum_{k=0}^{L-1} \hat{h}_k(n) x(n-k)$$

- Form the cost function:

$$J = E\left[e^2(n)\right]$$



- Modify the coefficients so as to reduce the cost:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) - \frac{1}{2}\mu \frac{\partial J(\hat{\mathbf{h}})}{\partial \hat{\mathbf{h}}}$$

Gradients in all the dimensions of **h**

Step-size

Optimal Filtering

- Steepest descent is an iterative (adaptive) algorithm to find the minimum point in the cost function
- Steepest descent ideally attains the optimal solution given by the Wiener-Hopf equations

$$\hat{\mathbf{h}}_{opt} = \mathbf{R}^{-1}\mathbf{p}$$

where $\mathbf{R} = E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]$ is the autocorrelation matrix

and $\mathbf{p} = E\left[\mathbf{x}(n)d(n)\right]$ is the cross-correlation vector

The Least-Mean-Square Adaptive Algorithm

- Least-Mean-Square is an approximation to the steepest gradient approach.
  - True gradient

$$\frac{\partial J(\mathbf{h}_k)}{\partial h_k} = -2E\left[x(n-k)e(n)\right]$$

  - Instantaneous approximation uses the product

$$x(n-k)e(n)$$

  - The expectation has been removed – gradient noise is expected

- Tap Update Equation:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu \mathbf{x}(n)\, e(n)$$

The Normalized NLMS ‘Workhorse’

- The Normalized LMS (NLMS) is popular in practice
  - Performance is less dependent on data properties

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{\mu}{\left\|\mathbf{x}(n)\right\|^2 + \delta}\mathbf{x}(n)\, e(n)$$

  - This is the workhorse of adaptive DSP
    - Low complexity
    - Can be run using fixed point arithmetic
    - Noise robust

Classes of Algorithms

- Supervised
  - A reference signal used to generate an error that can be minimised
- Blind
  - No reference signal available
- Single Channel
  - One observation
  - Identifiy a single channel system
- Multichannel
  - More than one observation
  - Identify several systems

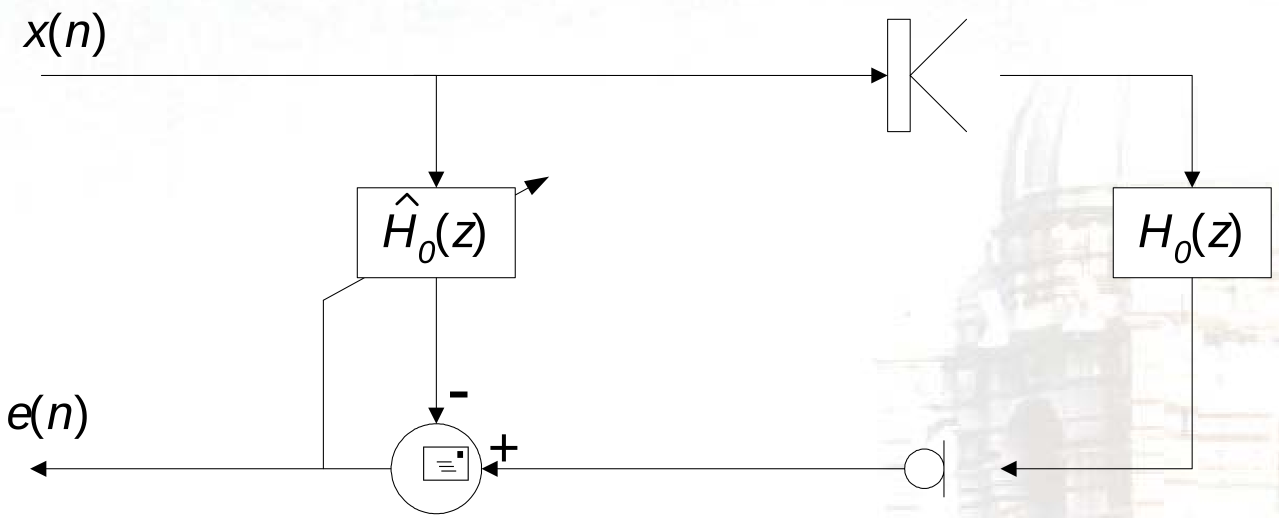
How much information is available?

Easier ◄──────────────────► Harder

How many assumptions must we make?

Single Channel System Identification

- Example: echo cancellation



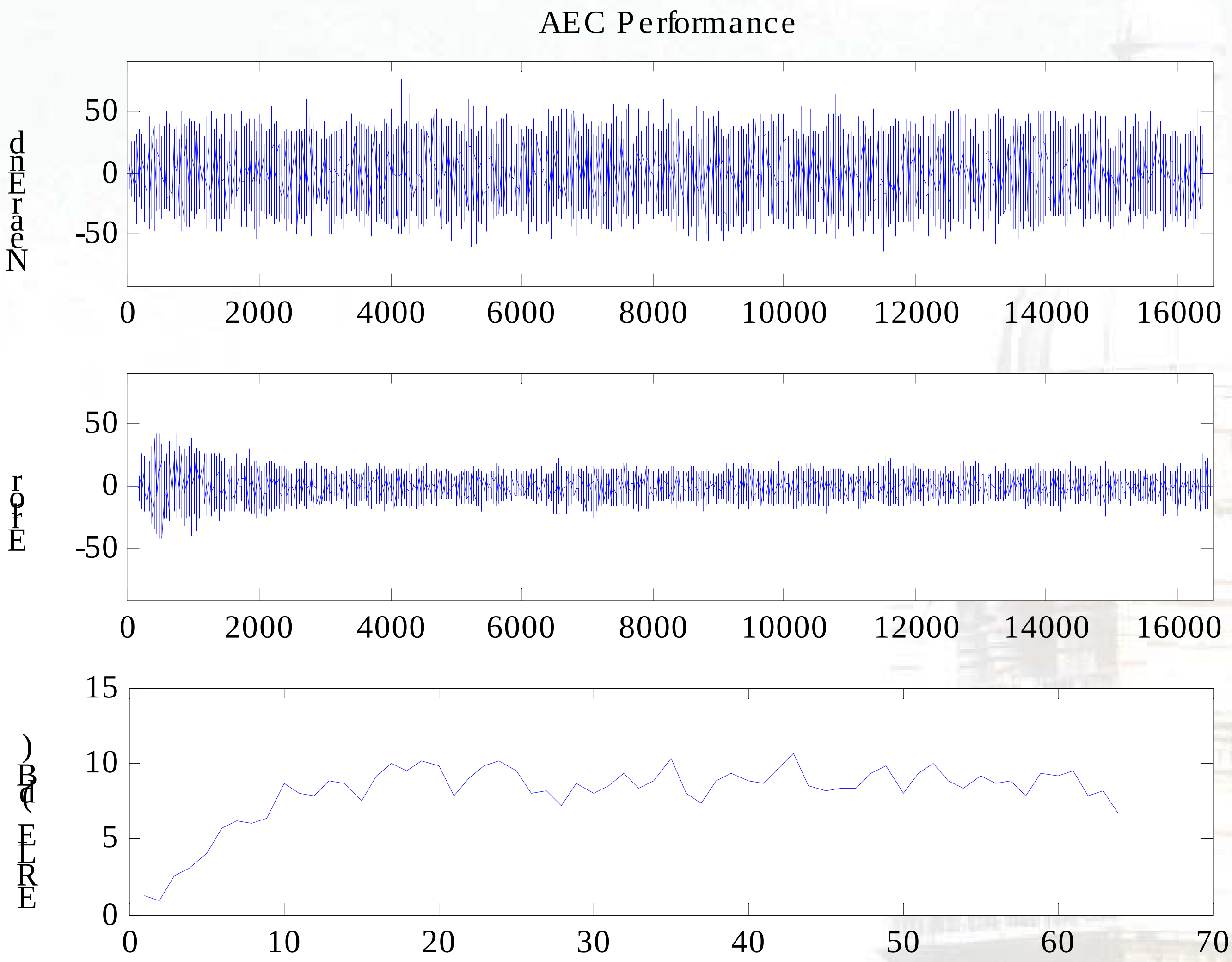
Performance Evaluation

- Segmental mean square error (MSE)
  - Measures the output error
  - Divide error signal into blocks of chosen length
  - Compute MSE in each block
  - Often expressed in dB as a power ratio between $d(n)$ and $e(n)$
- System Misalignment
  - Measures the error in the system identification

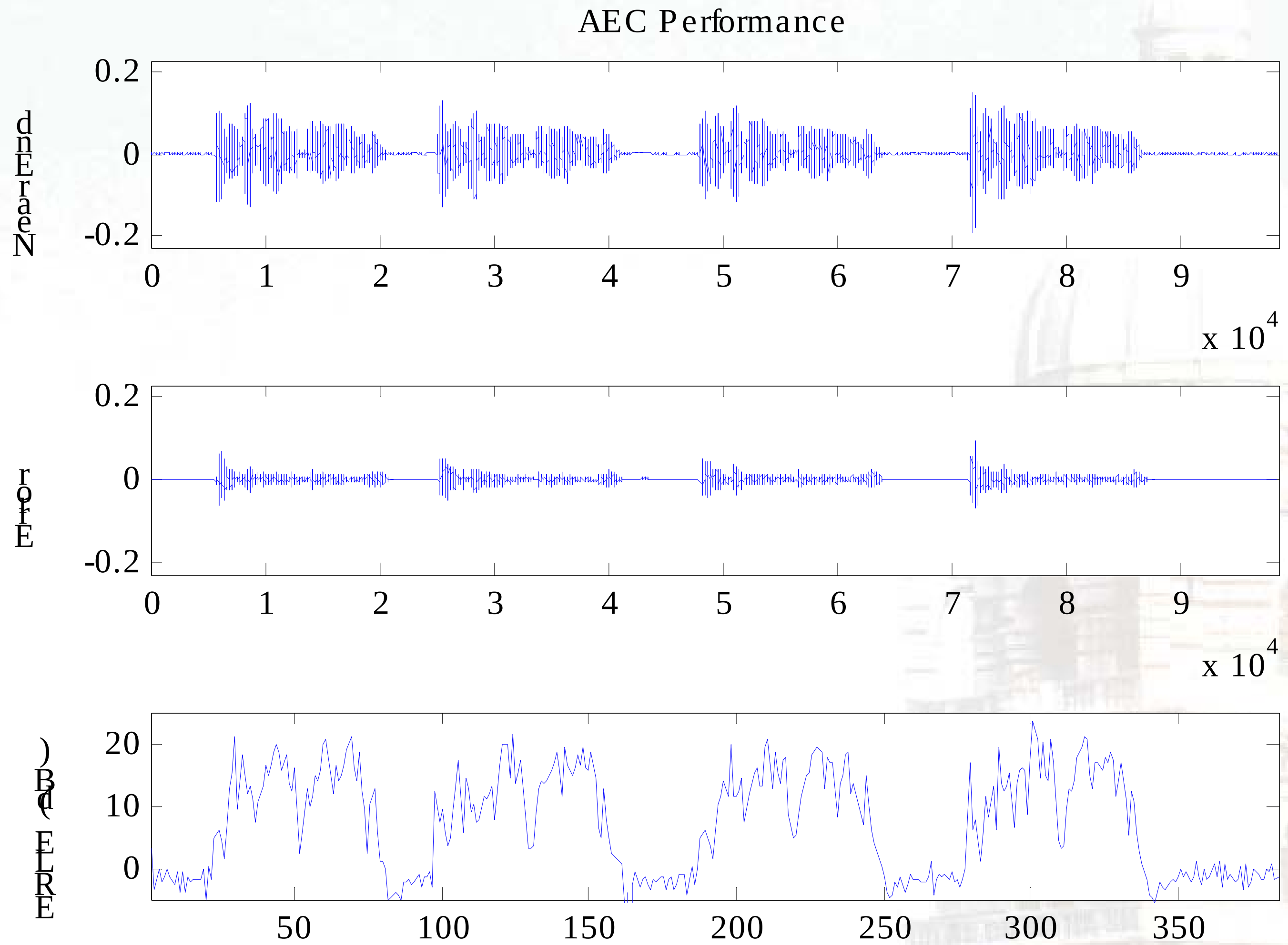
$$10\log_{10} E\left(\frac{\left\|\mathbf{h}-\hat{\mathbf{h}}(n)\right\|_2^2}{\left\|\mathbf{h}\right\|_2^2}\right) \text{ dB}$$

  - Where $\left\|\square\right\|_2^2$ is the sum of the squared elements of the vector

AEC Performance – White noise input



AEC Performance – speech input



Unsupervised Single Channel Technique

- Blind Signal Extraction
  - To extract a particular signal from a mixture of signals and noise
  - 'Blind' because we don't know the signals and we don't know how they have been mixed
- Consider N signals in a vector $\mathbf{s}(n) = [s_1(n), s_2(n), \ldots, s_N(n)]^T$ and an MxN mixing matrix **A**

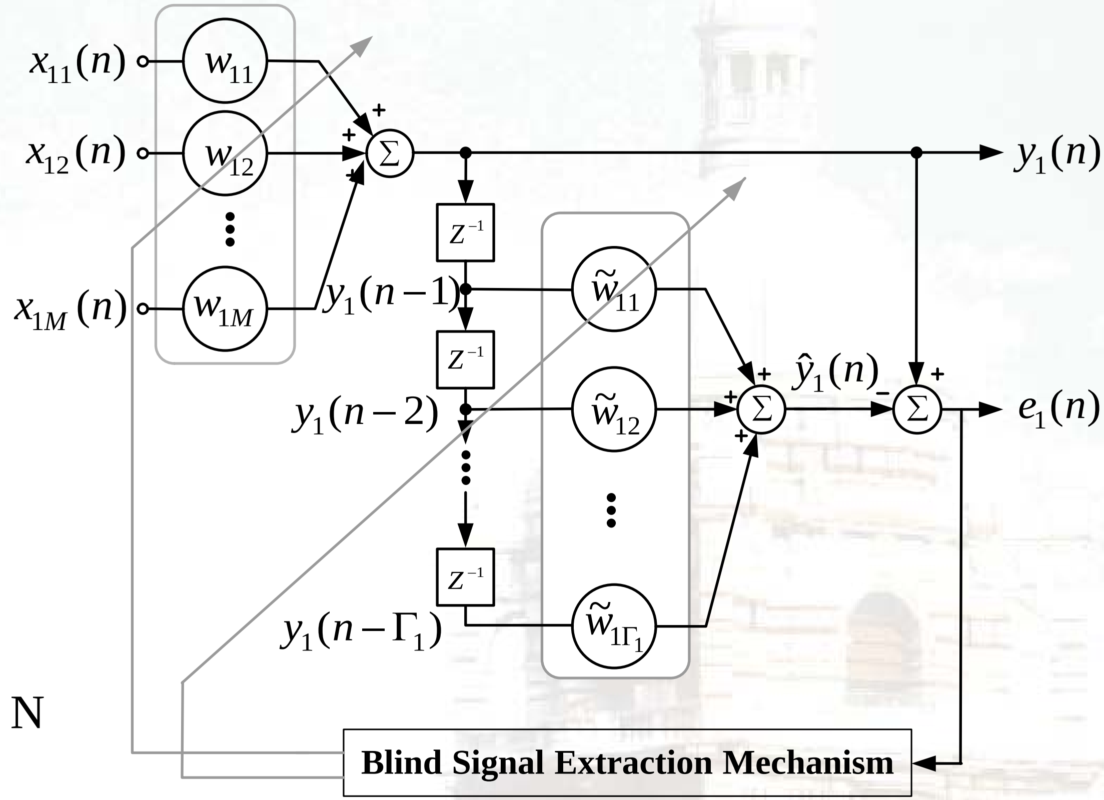
$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)$$

where $\mathbf{x}(n)$ are the observed mixtures

- Now consider demixed signals $\mathbf{y}(n)$ and a demixing matrix **W** applied to **x**

$$\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n)$$

Approach

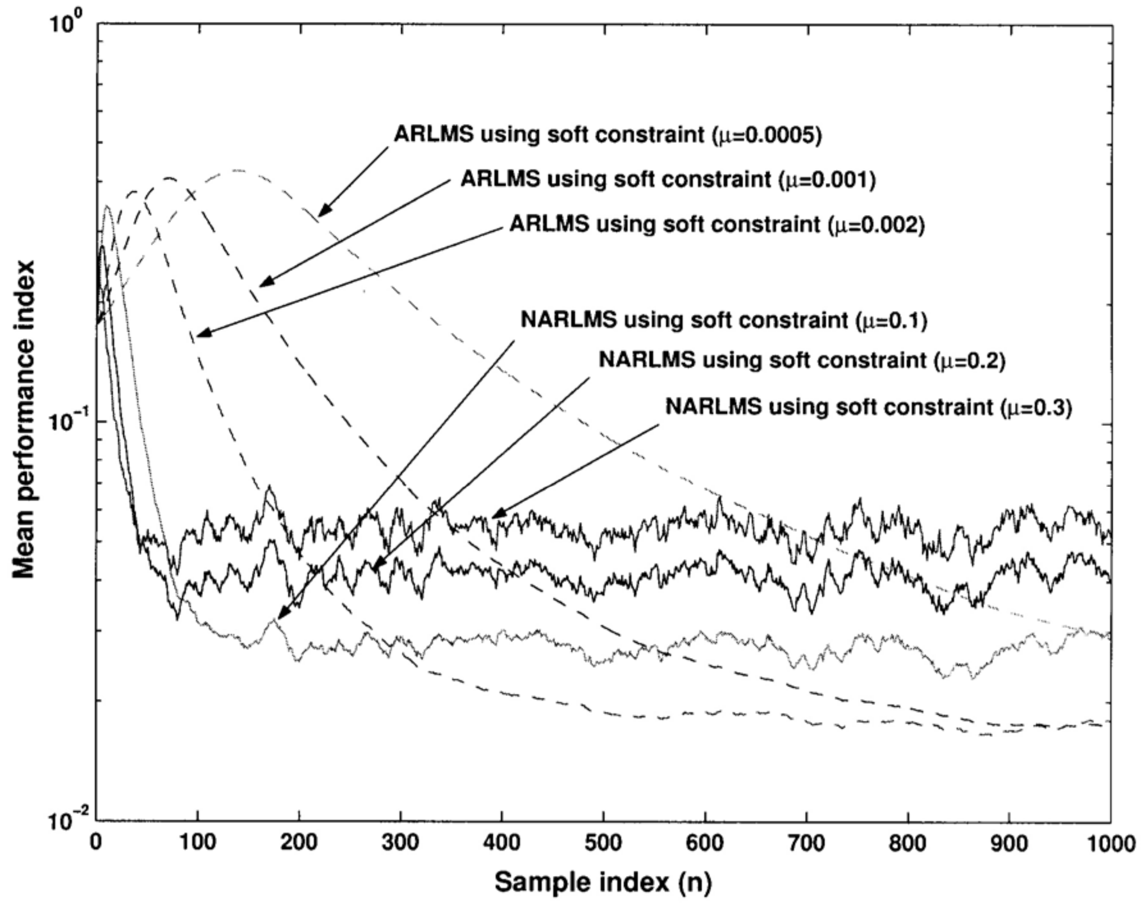


- Assumptions
  - A is an *M*x*N* matrix of rank N and $M \leq N$
  - The signal to be extracted was generated by a nonwhite source signal generated by an AR process

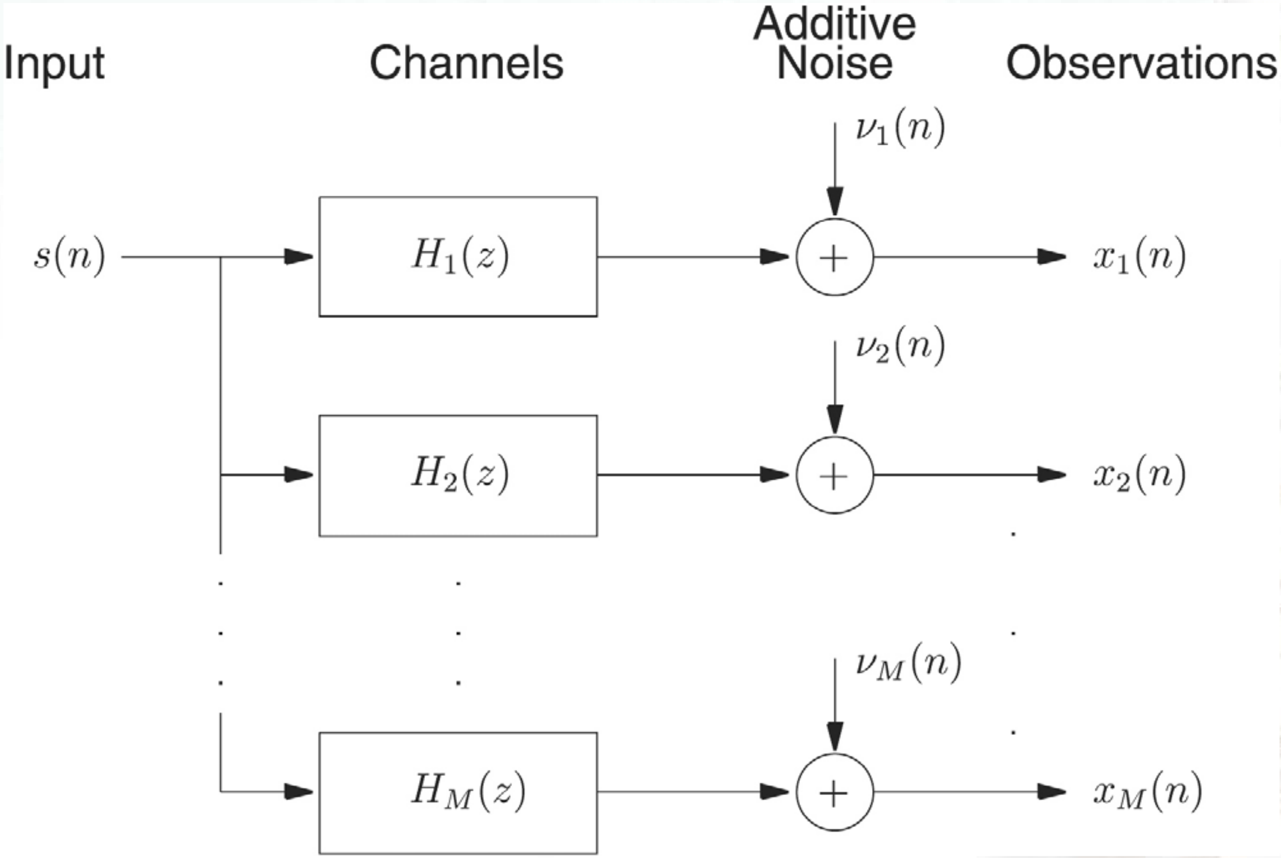
Results

worse

better



Blind Multichannel System Identification

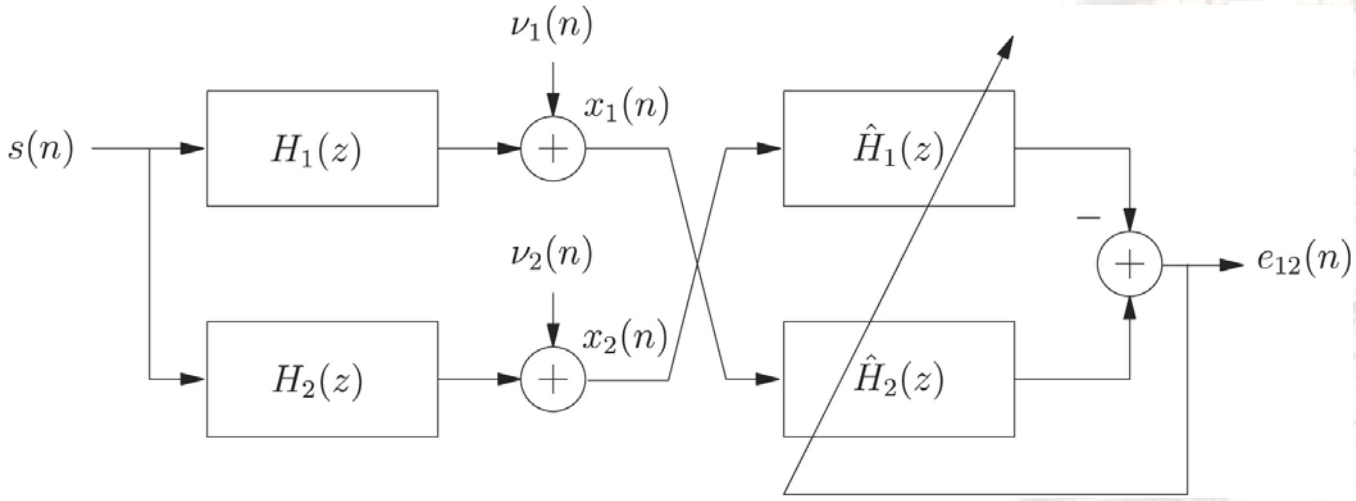


- The aim is to estimate the channels from the observations alone.

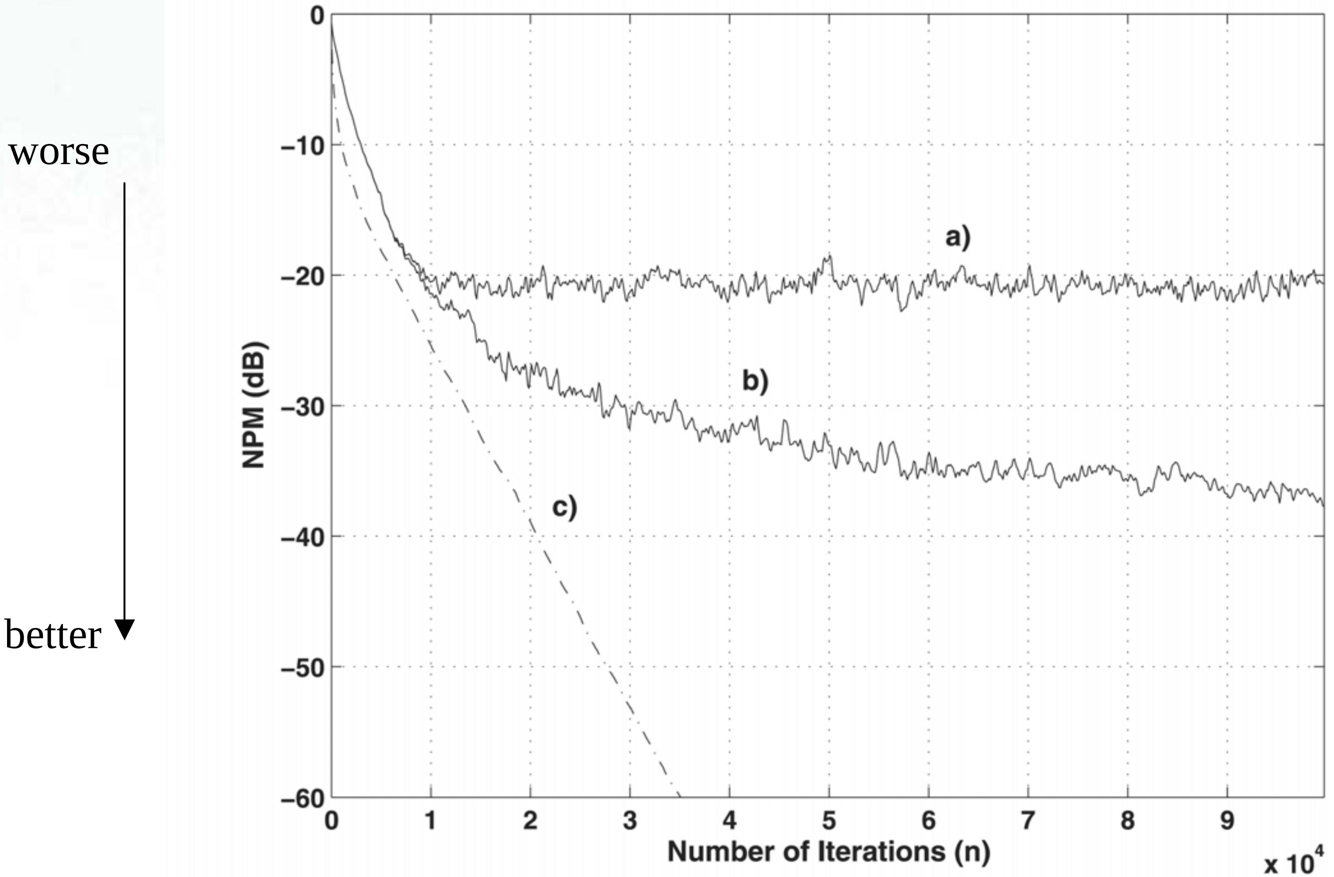
Approach

- The only fact that is known is that all the observations originate from the same source. An error function can therefore be written:

$$e(n) = \mathbf{x}_i^T(n)\mathbf{h}_j - \mathbf{x}_j^T(n)\mathbf{h}_i \qquad i, j = 1, 2, \ldots, M$$



Performance Evaluation



Variable step-size LMS-type adaptation

a) state-of-the-art; b) optimal variable step-size; c) theoretical performance bound

Inversion

- Inversion aims to recover the source signal
  - Convolve the observation with the inverse system estimate

- Considerations
  - Non-minimum phase systems result in unstable inverses
    - Zeros outside the unit circle in z give rise to unstable poles when inverted
  - Multi-channel systems can be exactly inverted using the MINT method
  - Note that exact inverse of a system estimate is usually worthless
    - Need the system estimate to be perfect
    - Current research into approximate inverse techniques

Summary and Conclusions

- Adaptive system identification is a powerful technique
- Supervised algorithms require a reference signal
- Unsupervised algorithms require other information in lieu of the reference
  - Assumptions on the nature of the signal or channel
  - Multichannel observations
- Open research questions in unsupervised multichannel blind methods are many
  - Step-size
  - Noise robustness
  - Order estimation

References

Haykin, S., *Adaptive Filter Theory*, 4th Ed. Prentice Hall, 2002.

Morgan, D.R.; Benesty, J.; Sondhi, M.M., "On the evaluation of estimated impulse responses," *Signal Processing Letters, IEEE* , vol.5, no.7 pp.174-176, Jul 1998

Amari, S., Cichocki, A. and Yang, H. H., "Blind signal separation and extraction: neural and information-theoretic approaches," in *Unsupervised Adaptive Filter, Vol. I: Blind Source Separation*, S. Haykin, Ed., pp. 63-138, John Wiley, 2000.

Gaubitch, N. D., Hasan, Md. K. and Naylor, P. A., "Generalized Optimal Step-Size for Blind Multichannel LMS System Identification," *Signal Processing Letters, IEEE* , to appear.

Miyoshi, M.; Kaneda, Y., "Inverse filtering of room acoustics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol.36, no.2 pp.145-152, Feb 1988.