Blind Deconvolution and Blind Source Separation

Shane M. Haas

Laboratory for Information and Decision Systems

Massachusetts Institute of Technology

I – Introduction

II – The Bussgang Algorithm for Blind Deconvolution

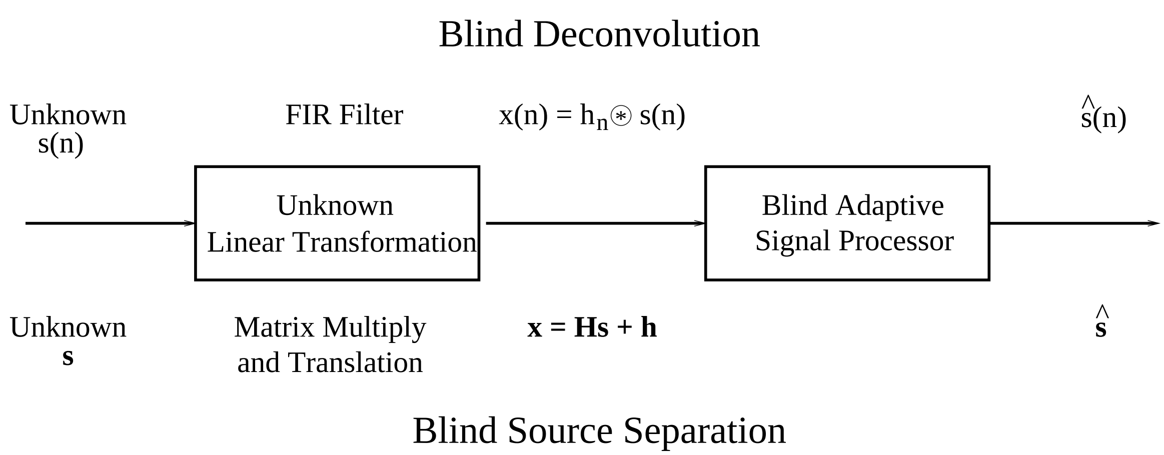
III – An Information-Maximization Algorithm for Blind Source Separation

IV – Demonstrations

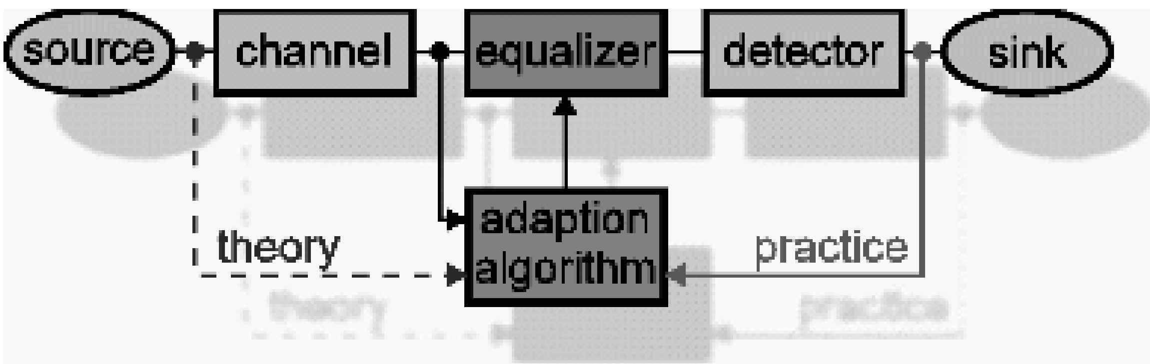
V – Conclusions

I − Introduction

- **THE PROBLEM:** Unraveling an unknown linear operation

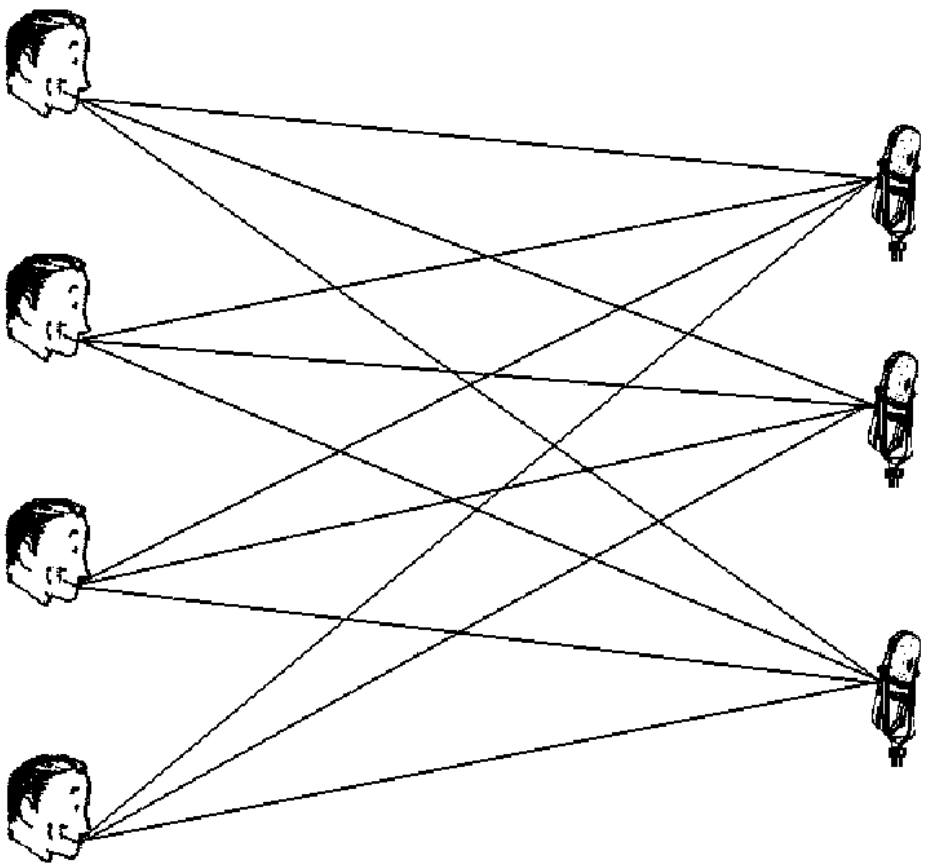


- The Channel Equalization Problem*



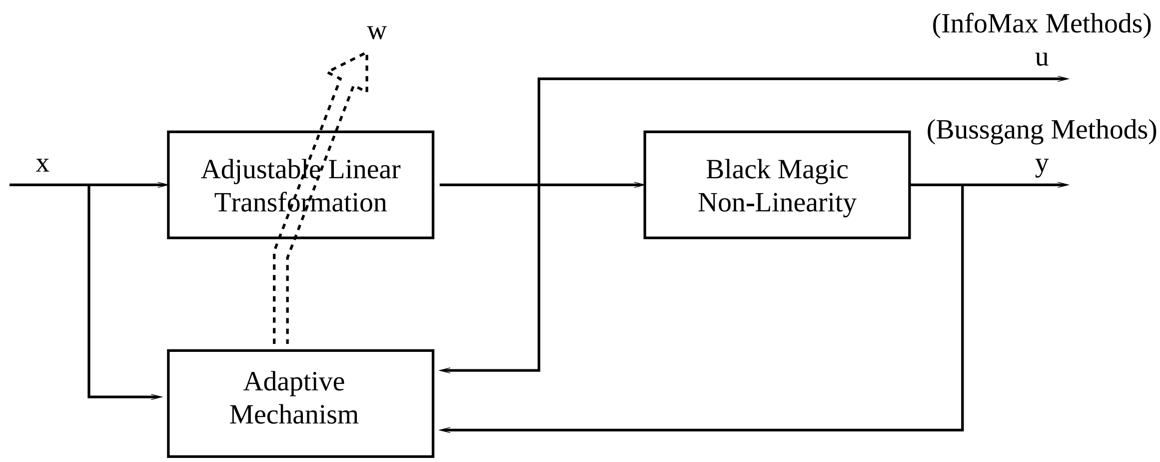
*Graphic from Institute for Communications Engineering Webpage

• The Cocktail Party Problem*



*Graphic from Daniel Rowe's Webpage

- **A SOLUTION:** Blind adaptive signal processing

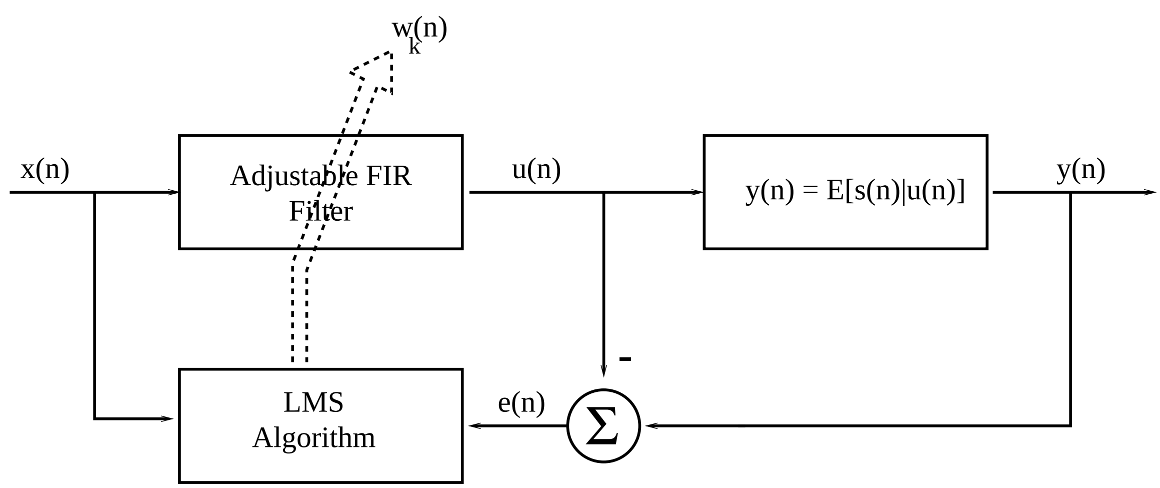


- Two Questions:
  - How to choose the non-linearity?
  - What should the adaptive mechanism optimize?

- The Bussgang Algorithm for Blind Deconvolution (Haykin, 1991)
  - Non-Linearity: Bayes Estimator
  - Adaptive Mechanism: LMS Algorithm

- Information-Maximization Method for Blind Source Separation (Bell & Sejnowski, 1995)
  - Non-Linearity: Logistic Function
  - Adaptive Mechanism: InfoMax Algorithm

II – Bussgang Algorithm for Blind Deconvolution



- **The Non-Linearity:** Bayes Estimator

$$\begin{aligned} u(n) &= \sum_k w_k(n)x(n-k) \\ &= \sum_k a_k x(n-k) + \sum_k \left(w_k(n) - a_k\right) x(n-k) \\ &= s(n) + v(n) \end{aligned}$$

where

$$\begin{aligned} v(n) &= \sum_k (w_k(n) - a_k) x(n-k) \\ a_n &= \text{Ideal Equalizing Filter} \end{aligned}$$

- Assume that $v(n)$ is:
  - zero-mean,
  - Gaussian,
  - white, and is
  - statistically independent of the source sequence $s(n)$.
- Assume that $s(n)$ is iid uniform with zero mean and unit variance

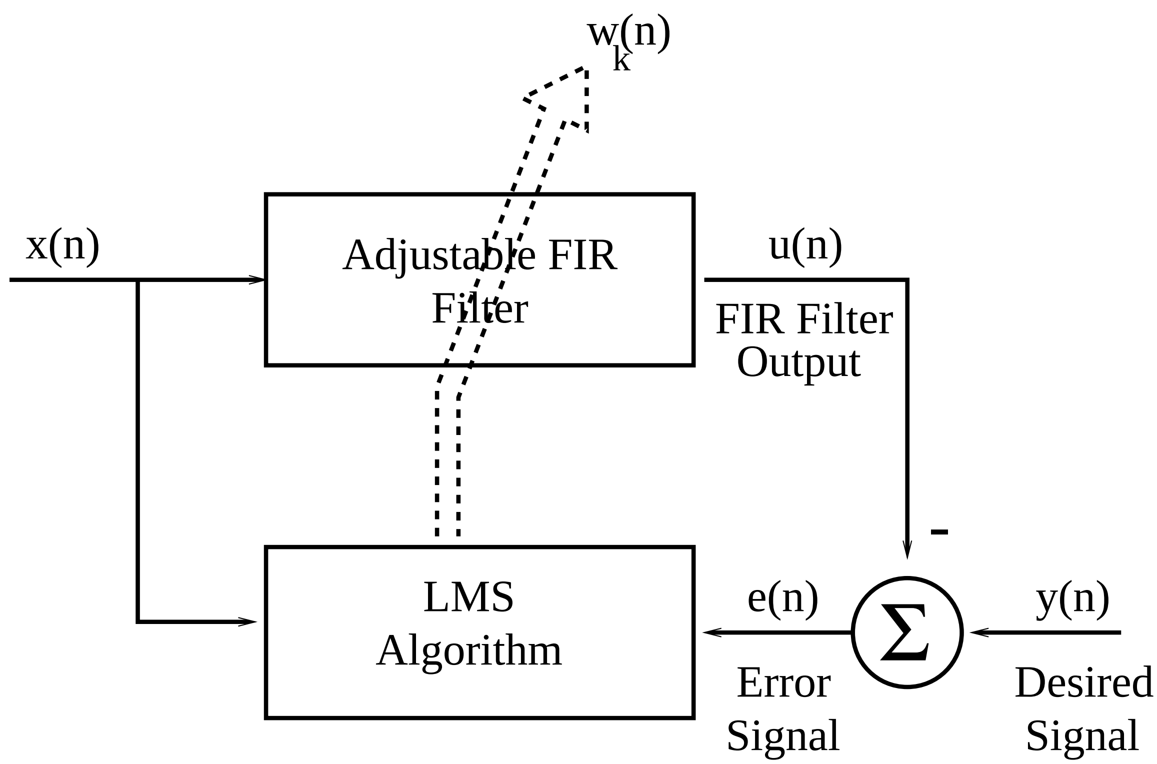
- Under these assumptions:

$$\begin{aligned} y(n) &= E\left[s(n)|u(n)\right] \\ &= \frac{1}{c_0}u(n) + \\ &\quad \frac{\sigma}{c_0}\frac{Z(u(n)+c_0/\sqrt{3}) - Z(u(n)-c_0/\sqrt{3})}{Q(u(n)-c_0/\sqrt{3}) - Q(u(n)+c_0/\sqrt{3})} \end{aligned}$$

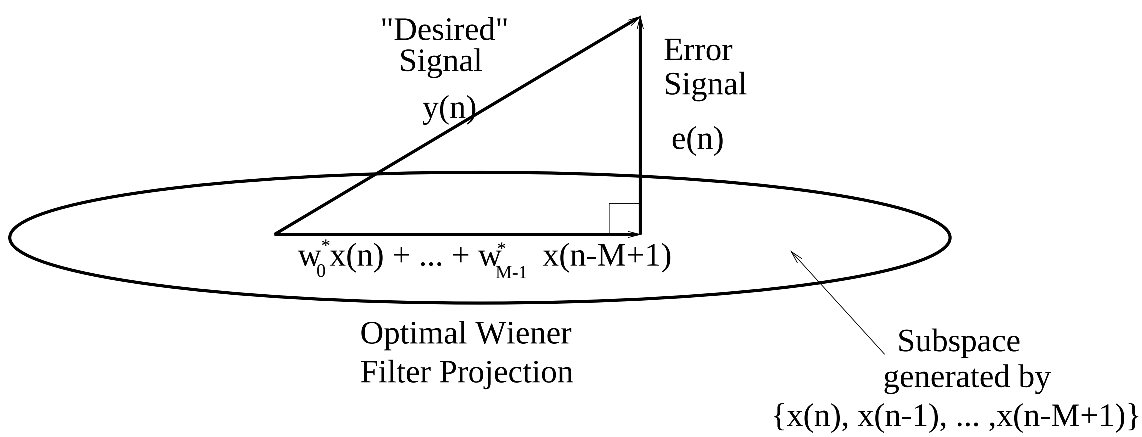
where σ^2 is the variance of $v(n)$, $c_0 = \sqrt{1-\sigma^2}$ is a normalizing factor, and

$$\begin{aligned} Z(u) &= \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \\ Q(u) &= \frac{1}{\sqrt{2\pi}}\int_u^\infty e^{-\frac{z^2}{2}}dz \end{aligned}$$

- **The Adaptive Mechanism:** The Least-Mean Squares (LMS) Algorithm
  - Adjust weights to minimize mean squared error:



- Iterative approximation to optimal Wiener filter:



- Wiener-Hopf Equation:

$$
\begin{aligned}
E[\mathbf{x}(n)e^*(n)] &= E[\mathbf{x}(n)(y(n) - \mathbf{w}^H\mathbf{x}(n))^*] \\
&= \mathbf{p} - \mathbf{R}\mathbf{w} \\
&= 0 \\
\Longleftrightarrow \mathbf{R}\mathbf{w} &= \mathbf{p}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{w} &= (w_0, w_1, \ldots, w_{M-1})^T \\
\mathbf{x}(n) &= (x(n), x(n-1), \ldots, x(n-M+1))^T \\
\mathbf{p} &= E[\mathbf{x}(n)y^*(n)] \\
\mathbf{R} &= E[\mathbf{x}(n)\mathbf{x}^H(n)]
\end{aligned}
$$

- The Mean Squared Error

$$
\begin{aligned}
J(n) &= E\left[|e(n)|^2\right] \\
&= E\left[|y(n) - u(n)|^2\right] \\
&= E\left[|y(n) - \mathbf{w}^H(n)\mathbf{x}(n)|^2\right] \\
&= E[|y(n)|^2] - \mathbf{w}^H(n)\mathbf{p} \\
&\quad -p^H\mathbf{w}(n) + \mathbf{w}^H(n)\mathbf{R}\mathbf{w}(n)
\end{aligned}
$$

- The Gradient (Assume $y(n)$ is independent of $\mathbf{w}$)

$$
\frac{\partial J(n)}{\partial \mathbf{w}(n)} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n)
$$

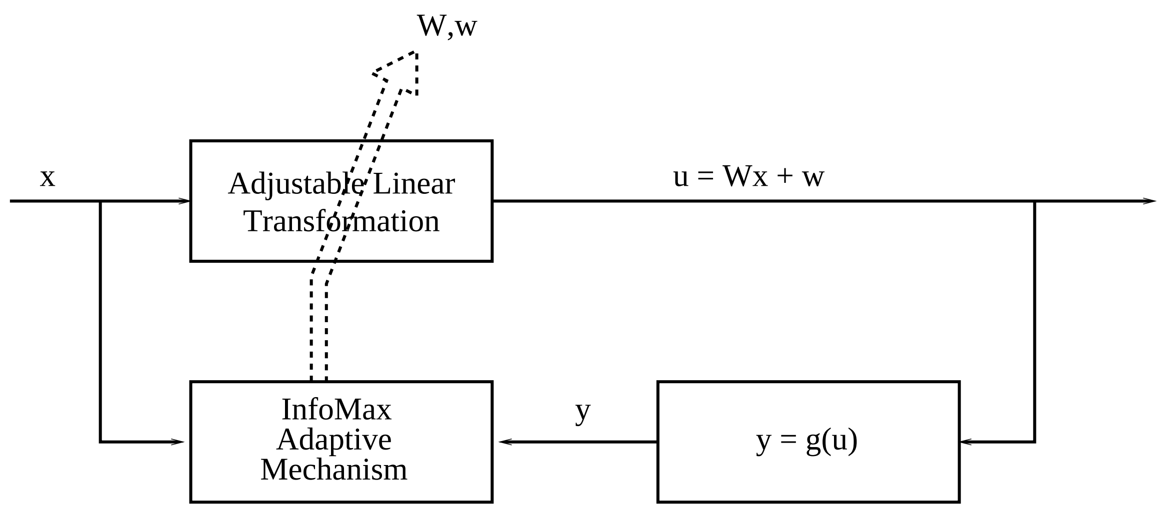
- Steepest-Descent Algorithm

$$\begin{aligned}\Delta \mathbf{w} &= \mathbf{w}(n+1) - \mathbf{w}(n) \\ &\propto -\frac{\partial J(n)}{\partial \mathbf{w}(n)} \\ &\propto \mathbf{p} - \mathbf{R}\mathbf{w}(n)\end{aligned}$$

- The LMS Algorithm ($\mathbf{p} \approx \mathbf{x}(n)y^*(n)$ and $\mathbf{R} \approx \mathbf{x}(n)\mathbf{x}^H(n)$)

$$\begin{aligned}\Delta \mathbf{w} &\propto \mathbf{x}(n)y^*(n) - \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}(n) \\ &= \mathbf{x}(n)(y(n) - \mathbf{w}^H(n)\mathbf{x}(n))^* \\ &= \mathbf{x}(n)e^*(n)\end{aligned}$$

III – An Information-Maximization Approach to Blind Source Separation



- **The Non-Linearity:** Logistic Function

$$
\begin{aligned}
y_i &= \frac{1}{1 + e^{-u_i}} \\
\mathbf{u} &= \mathbf{W}\mathbf{x} + \mathbf{w}
\end{aligned}
$$

- **The Adaptive Mechanism:** InfoMax Algorithm

  - Maximize the differential entropy of the non-linear output

  - The pdf of the non-linear output:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}(\mathbf{x})}{|J|}$$

    where J is the Jacobian

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_N} \\ \vdots & & \vdots \\ \frac{\partial y_N}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_N} \end{vmatrix}$$

– The differential entropy of the non-linear output:

$$\begin{aligned} h(\mathbf{Y}) &= -E[\log f_{\mathbf{Y}}(\mathbf{y})] \\ &= E[\log |J|] - E[\log f_{\mathbf{X}}(\mathbf{x})] \end{aligned}$$

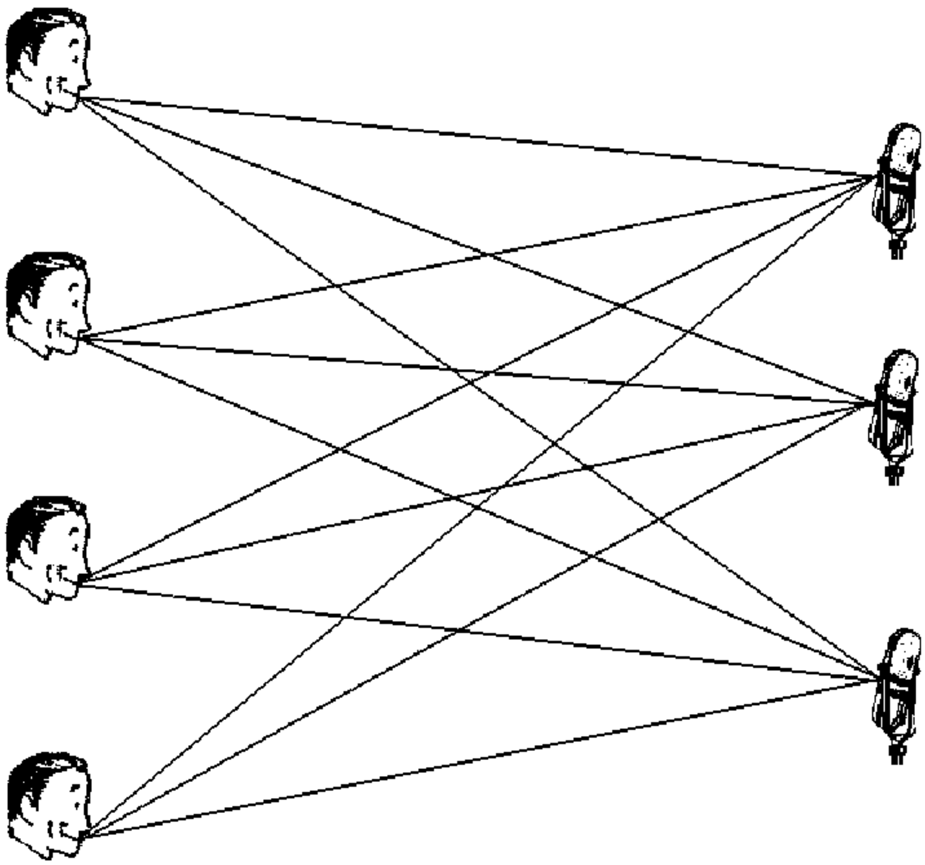
– Adjust weights in recursive manner:

$$\begin{aligned} \boldsymbol{\Delta}\mathbf{W} &\propto \frac{\partial h(\mathbf{Y})}{\partial \mathbf{W}} \approx \frac{\partial \log |J|}{\partial \mathbf{W}} \\ \boldsymbol{\Delta}\mathbf{w} &\propto \frac{\partial h(\mathbf{Y})}{\partial \mathbf{w}} \approx \frac{\partial \log |J|}{\partial \mathbf{w}} \end{aligned}$$

– Several pages of derivatives later...

$$\begin{aligned} \Delta \mathbf{W} &\propto [\mathbf{W}^T]^{-1} + (1 - 2\mathbf{y})\mathbf{x}^T \\ \Delta \mathbf{w} &\propto 1 - 2\mathbf{y} \end{aligned}$$

IV – Demonstrations



V − Conclusions

- Introduced blind deconvolution and blind source separation problem
- Bussgang algorithm for blind deconvolution
- Information-maximization algorithm for blind source separation
- Demonstrations