Blind Source Separation Techniques for Multiuser Processing Problems

Charles Casimiro Cavalcante, Dr.
GTEL-UFC, Brazil

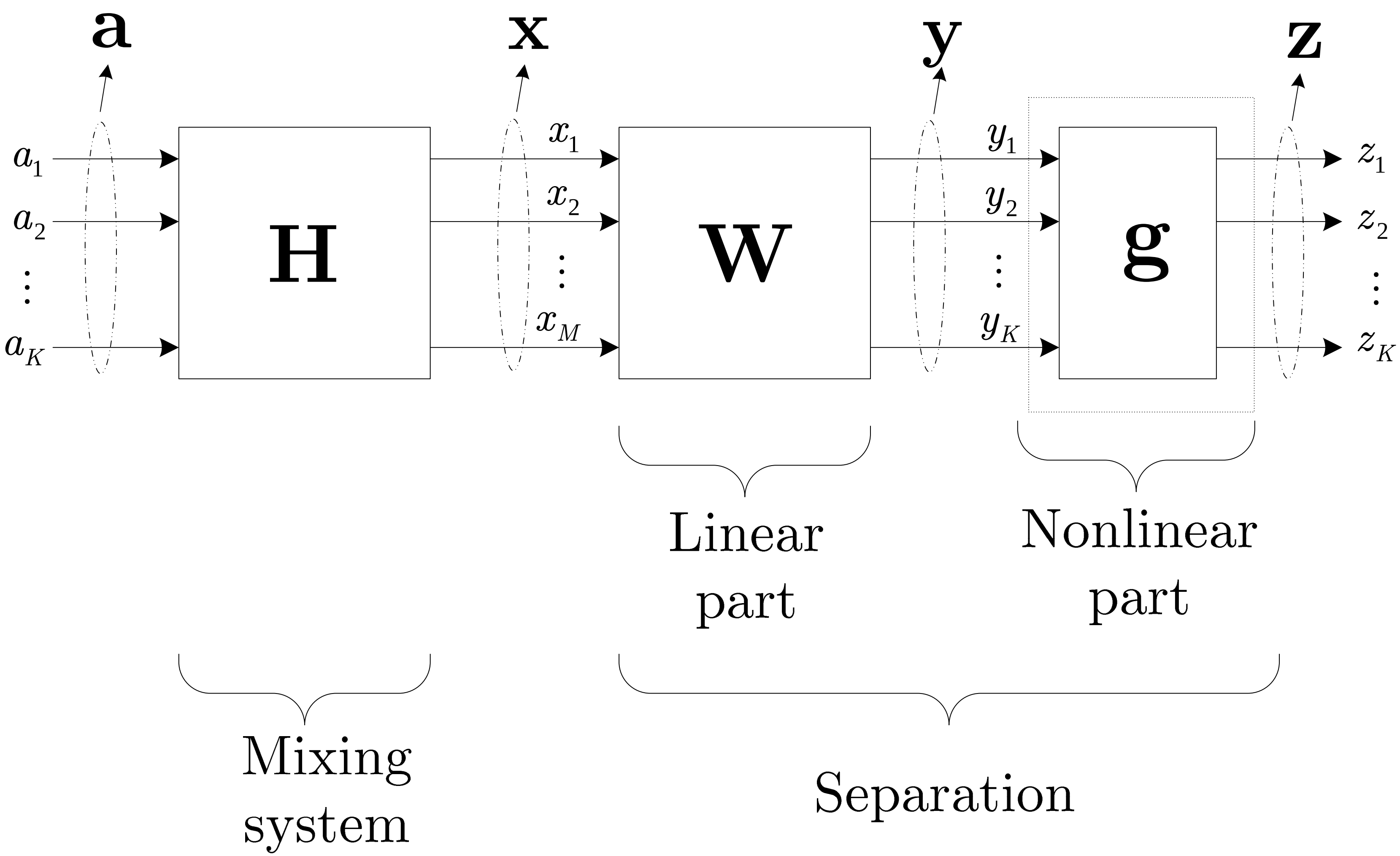
September 11th, 2007

Contents

Blind Source Separation (BSS)

- Generic model
- Supposedly unsolvable until 1985 (Jutten & Hérault)
- Application in several areas
  - Image processing
  - Communications
  - Biomedical processing
  - Pattern recognition

BSS: general model



BSS: model

- Dimensions

$$\mathbf{x}(n) = \mathcal{F}\left(a(n), \mathbf{v}(n), n\right) \longleftarrow \text{Mapping}$$

$$\mathbf{a}(n) = \left[a_1(n) \quad a_2(n) \quad \cdots \quad a_N(n)\right]^T \longleftarrow N \text{ sources}$$

$$\mathbf{v}(n) = \left[v_1(n) \quad v_2(n) \quad \cdots \quad v_K(n)\right]^T \longleftarrow K \text{ noise signals}$$

$$\mathbf{x}(n) = \left[x_1(n) \quad x_2(n) \quad \cdots \quad x_M(n)\right]^T \longleftarrow M \text{ sensors}$$

BSS: model (cont.)

- Typical assumptions
  - $\mathcal{F}$ linear and time-invariant
  - Sources are mutually independent and independent from noise
  - $K = M$
  - $M \geq N$ (at least as many sensors as sources)

BSS: model (cont.)

- Matricially

$$\mathbf{x}(n) = \mathbf{H}\mathbf{a}(n) + \mathbf{v}(n)$$

  - Goal:

$$\mathbf{y}(n) = \mathbf{W}^H \mathbf{x}(n) = \hat{\mathbf{a}}(n)$$

  - Noisy case: not possible perfect recovering
  - Including noise:

$$\mathbf{H}^{\dagger} = \begin{bmatrix} \mathbf{H} & \mathbf{I}_M \end{bmatrix}$$

$$\mathbf{a}^{\dagger}(n) = \begin{bmatrix} \mathbf{a}^T(n) & \mathbf{v}^T(n) \end{bmatrix}^T$$

BSS techniques

- InfoMax/MaxEnt
  - Entropy maximization
- Negentropy measure
  - Make the estimates “as much non-Gaussian as possible”
- Negative Mutual Information
  - Maximization of the negative of mutual information
- *Independent Component Analysis* (ICA)
  - Independent sources $\rightarrow$ independent estimates
- Contrast functions
  - Chosen functions in such a way that when they are maximized the sources are identified/separated

BSS techniques (cont.)

- *Independent Component Analysis*
  - Independent sources
  - Force independence on the system output
  - Measure of independence of the sources
    - Non-Gaussianity
      - Kurtosis
      - Negentropy
      - Mutual information

BSS techniques (cont.)

- Contrast functions
  - Maximized whenever the separaion is achieved
    - $\Psi(p_y)$ is invariant to permutations:
      - $\Psi(p_{Py}) = \Psi(p_y)$ for any permutation matrix **P**
    - $\Psi(py)$ is invariant to scale changes:
      - $\Psi(p_{Dy}) = \Psi(p_y)$ for any diagonal matrix **D**
    - If $\mathbf{y}$ has independent components, then:
      - $\Psi(p_{Wy}) \leq \Psi(p_y)$ for any invertible matrix **W**

BSS techniques (cont.)

- InfoMax/MaxEnt

$$I(\mathbf{z},\mathbf{x}) = \mathcal{H}(\mathbf{z}) - \mathcal{H}(\mathbf{z}|\mathbf{x}) \qquad \mathbf{z} = g(\mathbf{W}\mathbf{x})$$

$$\mathbf{g}(\mathbf{y}) = \begin{bmatrix} g_1(y_1) & g_2(y_2) & \cdots & g_K(y_K) \end{bmatrix}^T$$

  - Choice of the functions
  - Artificial neural networks usage
  - Contrast function

$$\Psi_{\text{InfoMax}}(\mathbf{W}) \overset{\Delta}{=} \ln\left[\left|\det(\mathbf{W})\right|\right] + E\left\{\ln\left[\prod_{i=1}^{K} g'_i(y_i)\right]\right\}$$

BSS techniques (cont.)

- Negentropy

$$N_G(\mathbf{y}) = \mathcal{H}(\mathbf{y}^G) - \mathcal{H}(\mathbf{y})$$

$$\Updownarrow$$

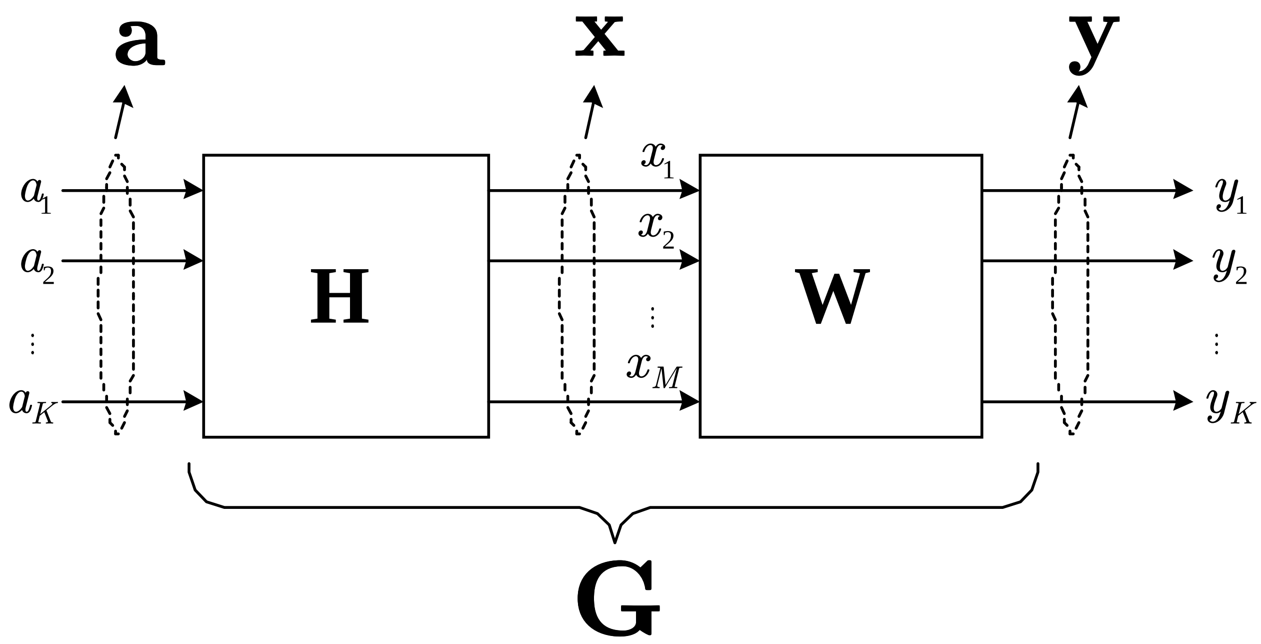
$$N_G(\mathbf{y}) = D\left(p_Y(\mathbf{y}) \| p_Y{}^G(\mathbf{y})\right)$$

Multiuser processing

- Particular case of source separation
  - **Discrete sources**
  - Finite alphabet
  - Same pdf (in general)
- Antenna array is very usual
- Linear processing is possible

Multiuser processing (cont.)

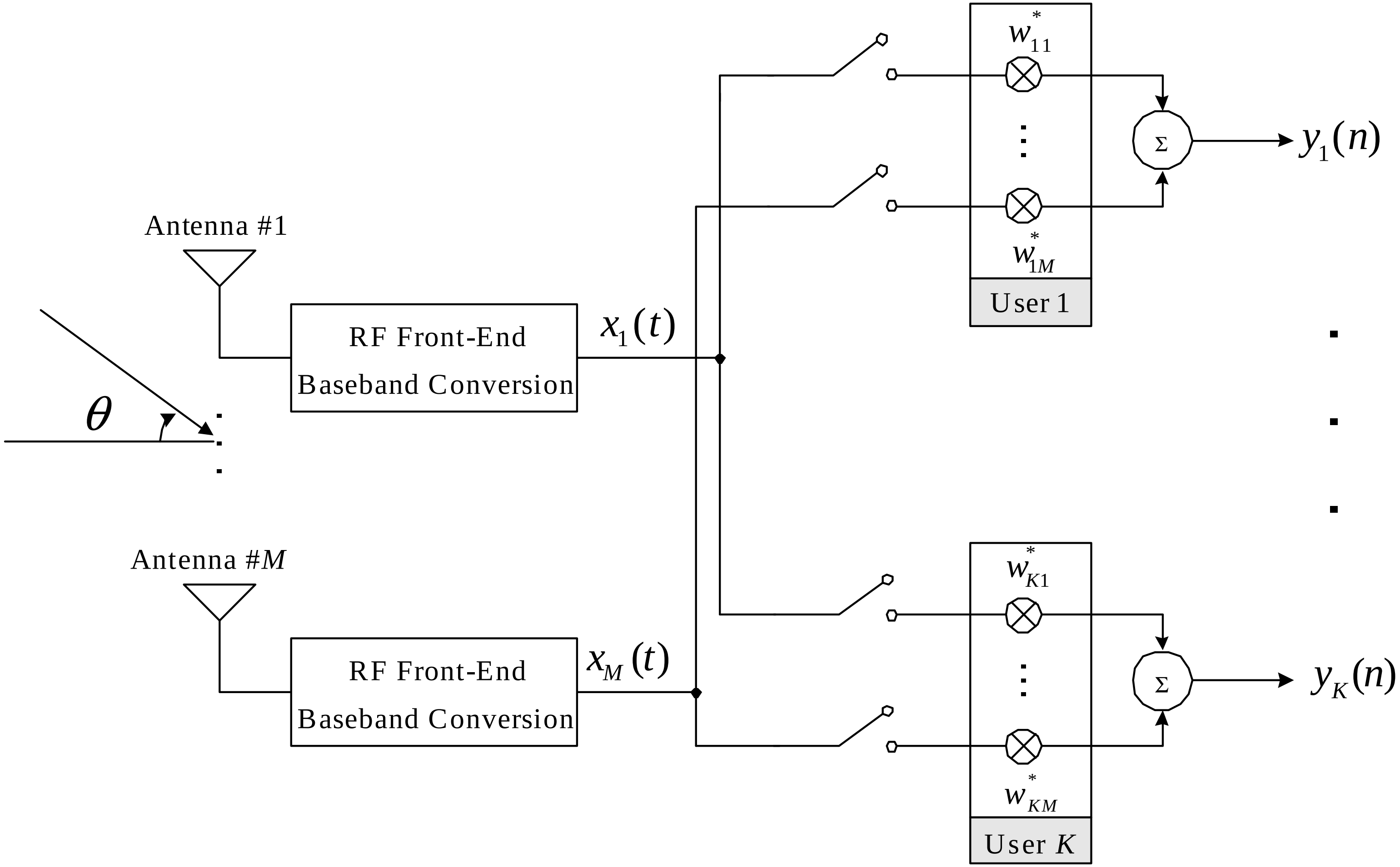
- Resulting model



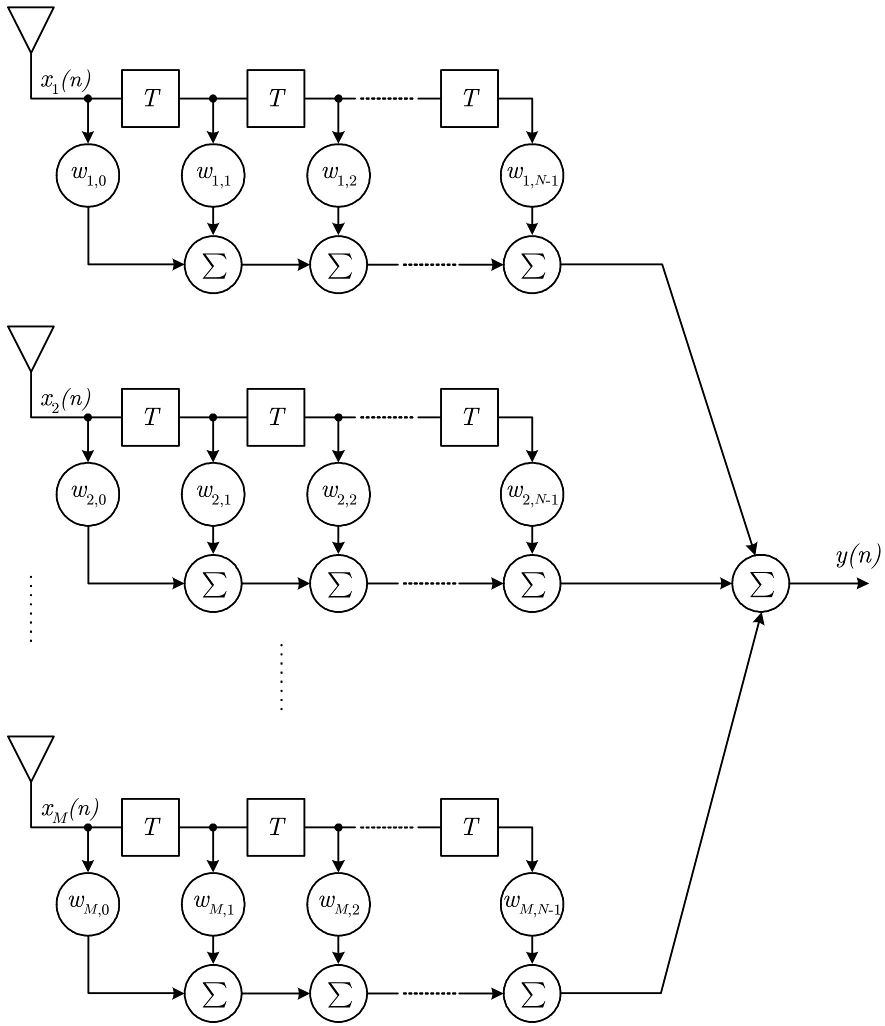
Multiuser processing (cont.)

- Use of multiple access techniques
  - TDMA : time
  - CDMA : code
  - SDMA : space
- Features
  - ISI – *Intersymbol Interference*
  - MAI – *Multiaccess Interference*

Spatial processing



Space-time processing



Multiuser processing (cont.)

- Signal modeling
  - Transmitted signal: $S_k(t) = \sum_{n=-\infty}^{\infty} \sqrt{P_k}\, a_k(n) p(t - nT)$
  - Received signal: $x(t) = \sum_{k=1}^{K} S_k(t - \tau_k) e^{j\phi_k} + v(t)$
  - Received signal at antenna *m*:

$$x_m(t) = \sum_{k=1}^{K} S_k(t - \tau_k) e^{j\phi_k} f_m(\theta_k) + v(t)$$

Multiuser processing (cont.)

- Signal modeling
  - *Steering vector:*

$$\mathbf{f}(\theta_k) = [f_1(\theta_k) \quad f_2(\theta_k) \quad f_3(\theta_k) \quad \cdots \quad f_M(\theta_k)]^T \qquad f_m(\theta_k) = \exp\left(j\,\frac{2\pi(m-1)d\,\mathrm{sen}(\theta_k)}{\lambda} \right)$$

  - Discrete base-band :

$$\mathbf{x}(n) = \sum_{k=1}^{K} \sqrt{P_k} \cdot a_k(n) \cdot \exp(j\phi_k) \cdot \mathbf{f}(\theta_k) + \mathbf{v}(n)$$

Multiuser processing (cont.)

- Signal model
  - *Beamformer*

$$y_k(n) = \mathbf{w}_k^H(n)\mathbf{x}(n)$$

Multiuser processing (cont.)

- Space-time processing
  - Considers multipath propagation
  - Received discrete base-band signal

$$\mathbf{x}(n) = \sum_{k=1}^{K} \sum_{l=0}^{L-1} h_k(l) a_k(n-l) \mathbf{f}\left(\theta_{k,l}\right) + \mathbf{v}(n)$$

  - Space-time filter output

$$y_k(n) = \mathbf{W}_k^H(n)\mathbf{X}(n)$$

$$\mathbf{W}_k(n) = \left[\mathbf{w}_{k0}^T(n) \quad \cdots \quad \mathbf{w}_{k(N-1)}^T(n)\right]^T \qquad \mathbf{X}(n) = \left[\mathbf{x}^T(n) \quad \cdots \quad \mathbf{x}^T(n-N+1)\right]^T$$

Multiuser processing (cont.)

- Channel behavior

$$\mathbf{H}_k = \mathbf{F}\left(\theta_k\right)\mathbf{h}_k$$

$$\mathbf{F}\left(\theta_k\right) = \left[\mathbf{f}\left(\theta_{k,0}\right) \;\; | \;\; \mathbf{f}\left(\theta_{k,1}\right) \;\; | \;\; \cdots \;\; | \;\; \mathbf{f}\left(\theta_{k,L-1}\right)\right]$$

$$\mathbf{h}_k = diag\left(\left[h_k(0) \quad \cdots \quad h_k(L-1)\right]\right)$$

Multiuser processing (cont.)

- Received signal:

$$\boldsymbol{X}(n)=\sum_{k=1}^{K}\mathcal{H}_k\boldsymbol{A}_k(n)+\boldsymbol{V}(n) \qquad \boldsymbol{A}_k(n)=\begin{bmatrix}a_k(n) & \cdots & a_k(n-L-N+2)\end{bmatrix}^T$$

$$\boldsymbol{V}(n)=\begin{bmatrix}\boldsymbol{v}^T(n) & \cdots & \boldsymbol{v}^T(n-N+1)\end{bmatrix}^T$$

$$\mathcal{H}_k=\begin{bmatrix}\mathbf{H}_k & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_k & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_k\end{bmatrix}$$

Supervised strategies

- Need pilot signal
- Two types considered:
  - *Direct Matrix Inversion* (DMI)
    - Wiener solution based
  - *Least Mean Squares* (LMS)
    - Stochastic gradient

Supervised strategies (cont.)

- DMI
  - Wiener solution

$$\mathbf{R}_k = E\left\{\mathbf{x}(n)\mathbf{x}^H(n)\right\} = \sigma_a^2 \sum_{k=1}^{K} \mathbf{f}\left(\theta_k\right)\mathbf{f}^H\left(\theta_k\right) + \sigma_v^2 \mathbf{I}_M$$

$$\mathbf{p}_k = E\left\{a_k^*(n-\delta_k)\mathbf{x}(n)\right\} = \sigma_a^2 \mathbf{f}\left(\theta_k\right)$$

Space

$$\mathbf{R}_k = E\left\{\mathbf{X}(n)\mathbf{X}^H(n)\right\} = \sigma_a^2 \sum_{k=1}^{K} \mathcal{H}_k \mathcal{H}_k^H + \sigma_v^2 \mathbf{I}_{MN}$$

$$\mathbf{p}_k = E\left\{a_k^*(n-\delta_k)\mathbf{X}(n)\right\} = \sigma_a^2 \mathcal{H}_k^{\delta_k}$$

Space-time

Supervised strategies (cont.)

- DMI
  - Adaptive solution

$$\mathbf{w}_k(n) = \hat{\mathbf{R}}_k^{-1}(n)\hat{\mathbf{p}}_k(n)$$

$$\hat{\mathbf{R}}_k(n+1) = \omega\hat{\mathbf{R}}_k(n) + (1-\omega)\mathbf{x}(n)\mathbf{x}^H(n)$$

$$\hat{\mathbf{p}}_k(n+1) = \omega\hat{\mathbf{p}}_k(n) + (1-\omega)a_k^*(n-\delta_k)\mathbf{x}(n)$$

    - High complexity
    - Faster convergence

Supervised strategies (cont.)

- LMS
  - Classical strategy

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu\left[y_k(n) - a_k\left(n-\delta_k\right)\right]^* \mathbf{x}(n)$$

$$\mathbf{W}_k(n+1) = \mathbf{W}_k(n) - \mu\left[y_k(n) - a_k\left(n-\delta_k\right)\right]^* \mathbf{X}(n)$$

- Low complexity
- Slower convergence

Blind strategies

- Need of estimative of desired signal
- Strategies
  - *Multiuser Constant Modulus* (MU-CMA)
  - *Fast* MU-CMA
  - Maximization of *kurtosis* (*MUK*)

Blind strategies (cont.)

- MU-CMA
  - CMA - most known (studied) blind equalization criterion

$$J_{CMA}(\mathbf{w}) = E\left\{\left(\left|y(n)\right|^2 - R_2\right)^2\right\}$$

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) + \mu \cdot y^*(n) \cdot \left[1 - \left|y_k(n)\right|^2\right]\mathbf{x}(n)$$

- Penalizes signals which are away from a constant modulus

Blind strategies (cont.)

- MU-CMA
  - CMA – cannot be directly applied to MUD since it tends to capture users with higher power
  - Requirement – auxiliar criterion
  - Explicit decorrelation criterion
    - Penalizes the correlation among the output signals

Blind strategies (cont.)

- MU-CMA
  - Resulting cost function

$$J_{MU-CMA}(\mathbf{w}_k) = J_{CMA}(\mathbf{w}_k) + \underbrace{\gamma \sum_{i=1}^{K} \sum_{\substack{j=1 \\ j \neq 1}}^{K} \left| r_{ij} \right|^2}_{\text{Decorrelation criterion}}$$

$$\nabla J_{MU-CMA}(\mathbf{w}_k) = E\left\{ y^*(n) \cdot \left[\left| y_k(n) \right|^2 - 1 \right] \mathbf{x}(n) \right\} + \gamma \sum_{\substack{i=1 \\ i \neq k}}^{K} r_{ik} E\left\{ y_i^*(n) \mathbf{x}(n) \right\}$$

Blind strategies (cont.)

- MU-CMA
  - Stochastic approximations

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) + \mu \cdot \left(1 - \left|y(n)\right|^2\right) \cdot y_k^*(n) \cdot \mathbf{x}(n) - \gamma \sum_{\substack{i=1 \\ i \neq k}}^{K} \hat{r}_{ik}(n)\mathbf{p}_i(n)$$

$$\mathbf{R}_{yy}(n+1) = \alpha \mathbf{R}_{yy}(n) + (1-\alpha)\mathbf{y}(n)\mathbf{y}^H(n)$$

$$\mathbf{P}(n+1) = \alpha \mathbf{P}(n) + (1-\alpha)\mathbf{x}(n)\mathbf{y}^H(n)$$

Blind strategies (cont.)

- MU-CMA
  - Space-time

Delay

$$J_{MU-CMA}(\mathbf{w}_k) = J_{CMA}(\mathbf{w}_k) + \gamma \sum_{i=1}^{K} \sum_{\substack{j=1 \\ j \neq i}}^{K} \sum_{d=0}^{D-1} \left| r_{ij}(d) \right|^2$$

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) + \mu \cdot \left(1 - \left| y(n) \right|^2\right) \cdot y_k^*(n) \cdot \mathbf{x}(n) - \gamma \sum_{\substack{i=1 \\ i \neq k}}^{K} \sum_{d=0}^{D-1} \hat{r}_{ik,d}(n) \mathbf{p}_{i,d}(n)$$

$$\mathbf{R}_d(n+1) = \alpha \mathbf{R}_d(n) + (1-\alpha) \mathbf{y}(n) \mathbf{y}^H(n-d)$$

$$\mathbf{P}_d(n+1) = \alpha \mathbf{P}_d(n) + (1-\alpha) \mathbf{X}(n) \mathbf{y}^H(n-d)$$

Blind strategies (cont.)

- MU-CMA
  - Characteristics
    - Slow convergence
    - Decorrelation factor limits the residual error

Blind strategies (cont.)

- FMU-CMA (*Fast* MU-CMA)
  - Recursive version of MU-CMA
  - Improve convergence speed
  - Adaptive decorrelation factor
    - Improve residual error

Blind strategies (cont.)

- FMU-CMA (*Fast* MU-CMA)

$$\mathbf{w}_k(n) = \mathbf{R}_k^{-1}(n)\mathbf{d}(n)$$

$$\mathbf{R}_k(n+1) = \omega\mathbf{R}_k(n) + (1-\omega)\left|y_k(n)\right|^2 \mathbf{x}(n)\mathbf{x}^H(n)$$

$$\mathbf{d}_k(k+1) = \omega\mathbf{d}_k(n) + (1-\omega)y_k^*(n)\mathbf{x}(n) - \gamma\sum_{\substack{i=1\\ i\neq k}}^{K}\hat{r}_{ik}(n)\mathbf{p}_i(n)$$

$$\overline{r}_k(n) = \frac{1}{K-1}\sum_{\substack{i=1\\ i\neq k}}^{K}\left|\hat{r}_{ik}(n)\right|^2$$

$$\gamma(n) = \tanh\left(\left|\overline{r}_k(n)\right|\right)$$

Blind strategies (cont.)

- MUK (new idea!)

$$\begin{cases} \max_{\mathbf{G}} \; J_{\text{MUK}}(\mathbf{G}) = \sum_{j=1}^{K} \left| \kappa\left(y_j\right) \right| \\ \text{subject to: } \mathbf{G}^H \mathbf{G} = \mathbf{I} \end{cases}$$

1. Obtain $\mathbf{W}^e(n+1)$ as

$$\mathbf{W}^e(n+1) = \mathbf{W}(n) + \mu \cdot \text{sign}\left(\kappa_a\right) \mathbf{x}^*(n) \mathcal{Z}(n)$$

$$\mathcal{Z}(n) = \left[\left|y_1(n)\right|^2 y_1(n) \quad \cdots \quad \left|y_K(n)\right|^2 y_K(n) \right]$$

2. Obtain $\mathbf{w}_1(n+1) = \dfrac{\mathbf{w}_1^e(n+1)}{\left\| \mathbf{w}_1^e(n+1) \right\|}$

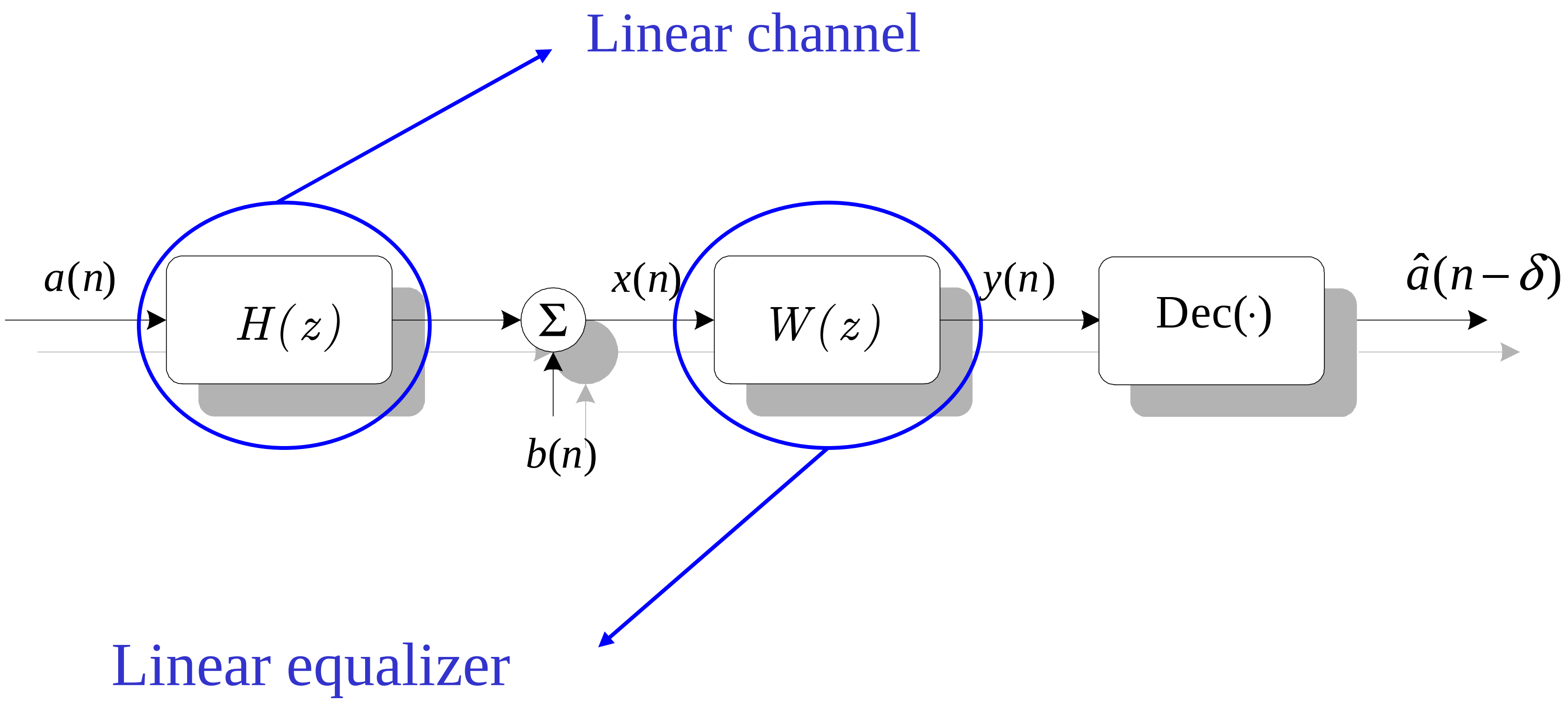
3. For $j = 2{:}K$

$$\mathbf{w}_j(n+1) = \frac{\mathbf{w}_j^e(n+1) - \sum_{l=1}^{j-1} \left[\mathbf{w}_l^H(n+1) \mathbf{w}_l^e(n+1) \right] \mathbf{w}_l(n+1)}{\left\| \mathbf{w}_j^e(n+1) - \sum_{l=1}^{j-1} \left[\mathbf{w}_l^H(n+1) \mathbf{w}_l^e(n+1) \right] \mathbf{w}_l(n+1) \right\|}$$

Fitting pdf Criterion

- Inicial proposition for SISO systems (single user equalization)
- Characterization of the probability density function of the output of an ideal equalizer
- Parameric estimation
- Criterion: estimation/entropy
  - Kullback-Leibler Divergence(KLD)
  - Contrast function
- Classification approach
  - Relation between the FP and CM

FP criterion: model



FP criterion (cont.)

- fdp of the output ideal equalizer

Thus:

$$\begin{aligned} y(n) &= \left(\mathbf{H}^T \mathbf{a}(n) + \mathbf{b}(n)\right) \mathbf{w}_{\text{ideal}} \\ &= \mathbf{a}^T(n) \underbrace{\mathbf{H}\mathbf{w}_{ideal}}_{\mathbf{g}_{\text{ideal}}} + \mathbf{b}^T(n)\mathbf{w}_{\text{ideal}} \\ &= a(n-\delta) + \vartheta(n) \end{aligned}$$

And

$$p_{Y,\text{ideal}}(y) = \frac{1}{\sqrt{2\pi\sigma_\vartheta^2}} \cdot \frac{1}{S} \sum_{i=1}^{S} exp\left(-\frac{\left|y - a_i\right|^2}{2\sigma_\vartheta^2}\right)$$

FP criterion (cont.)

- Criterion: measure of similarity of functions

$$\int_{-\infty}^{\infty} f(x)g(x)dx$$

- Functions to be compared
  - fdp of the signal at the ideal equalizer
  - Parametric model

$$\Phi(y) = \frac{1}{S}\sum_{i=1}^{S} exp\left(-\frac{\left|\mathbf{x}^T(n)\mathbf{w}(n) - a_i\right|^2}{2\sigma_r^2}\right)$$

FP criterion (cont.)

- How to measure it?
  - Kullback-Leibler Divergence

$$D\left(p(y)\|\Phi(y)\right) = \int_{-\infty}^{\infty} p(y)\cdot ln\left(\frac{p(y)}{\Phi(y)}\right)dy$$

- Cost function:
  - KLD term $\Phi(y)$-dependent

$$D\left(p(y)\|\Phi(y)\right) = \underbrace{\int_{-\infty}^{\infty} p(y)\cdot ln\left(p(y)\right)dy}_{\text{Entropy of } y} - \underbrace{\int_{-\infty}^{\infty} p(y)\cdot ln\left(\Phi(y)\right)dy}_{\text{Cost function}}$$

FP criterion (cont.)

- Cost function:

$$J(\mathbf{w}) = -\int_{-\infty}^{\infty} p(y) \cdot ln(\Phi(y))\, dy$$
$$= -E\{ln(\Phi(y))\}$$
$$= -E\left\{ln\left[\sum_{i=1}^{S} exp\left(-\frac{|y(n) - a_i|^2}{2\sigma_r^2}\right)\right]\right\}$$

- ***Fitting pdf Criterion*** **(FPC) -** J_{FPC}

FP criterion (cont.)

- Stochastic gradient:

$$\nabla J_{\text{FPC}}(\mathbf{w}) = \frac{\sum_{i=1}^{S} exp\left(-\frac{|y-a_i|^2}{2\sigma_r^2}\right)(y-a_i)}{\sigma_r^2 \cdot \sum_{i=1}^{S} exp\left(-\frac{|y-a_i|^2}{2\sigma_r^2}\right)}\mathbf{x}^*$$

- ***Fitting pdf Algorithm* (FPA)**

Multiuser criterion

- Generalization for MIMO case
- Idea: use K filters for separation of K users
- Problem:
  - Possible replication of one user!
- Alternative:
  - Explicit decorrelation

MU-FP criterion

- Utilization of an aditional criterion to ensure recovering/identification of all users
- Resulting criterion: ***MU-FPC***

$$J_{MU-FP}\left(\mathbf{w}_k\right) = J_{FP}\left(\mathbf{w}_k\right) + \gamma \sum_{i=1}^{K} \sum_{\substack{j=1 \\ j\neq 1}}^{K} \left|r_{ij}\right|^2 .$$

Equalization ($J_{FP}(\mathbf{w}_k)$)

Decorrelation among the users ($\gamma \sum \sum |r_{ij}|^2$)

Stochastic algorithm

- **MU-FPA**: ***Multi-User Fitting pdf Algorithm***

$$\mathbf{R}_{yy}(n+1) = \alpha\mathbf{R}_{yy}(n) + (1-\alpha)\mathbf{y}(n)\mathbf{y}^{H}(n)$$
$$\mathbf{P}(n+1) = \alpha\mathbf{P}(n) + (1-\alpha)\mathbf{x}(n)\mathbf{y}^{H}(n)$$

→ Estimation of the matrices of cross-correlation and autocorrelation

$$\nabla J_{\text{FPC}}(\mathbf{w}) = \frac{\sum_{i=1}^{S} exp\left(-\frac{|y-a_i|^2}{2\sigma_r^2}\right)(y-a_i)}{\sigma_r^2 \cdot \sum_{i=1}^{S} exp\left(-\frac{|y-a_i|^2}{2\sigma_r^2}\right)}\mathbf{x}^*$$

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu \cdot \nabla J_{FPC}(\mathbf{w}_k) - \gamma \sum_{\substack{i=1 \\ i \neq k}}^{K} \hat{r}_{ik}(n)\mathbf{p}_i(n)$$

→ ***MU-FPA***

Features

- Recovering of the phase of transmitted
- Low complexity: LMS-type
- Degree of freedom: σ_r^2
- Limitations:
  - Dependence of the factor γ
  - *Trade-off* – steady-state error $\times$ lost users

Alternative strategy

- Adaptive decorrelation factor

$$\bar{r}_k(n) = \frac{1}{K-1}\sum_{\substack{i=1\\ i\neq k}}^{K}\left|\hat{r}_{ik}(n)\right|^2$$

$$\gamma(n) = tanh\left(\left|\bar{r}_k(n)\right|\right)$$

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu\cdot\nabla J_{FPC}(\mathbf{w}_k) - \gamma(n)\sum_{\substack{i=1\\ i\neq k}}^{K}\hat{r}_{ik}(n)\,\mathbf{p}_i(n)$$

Results MU-FPA: SDMA

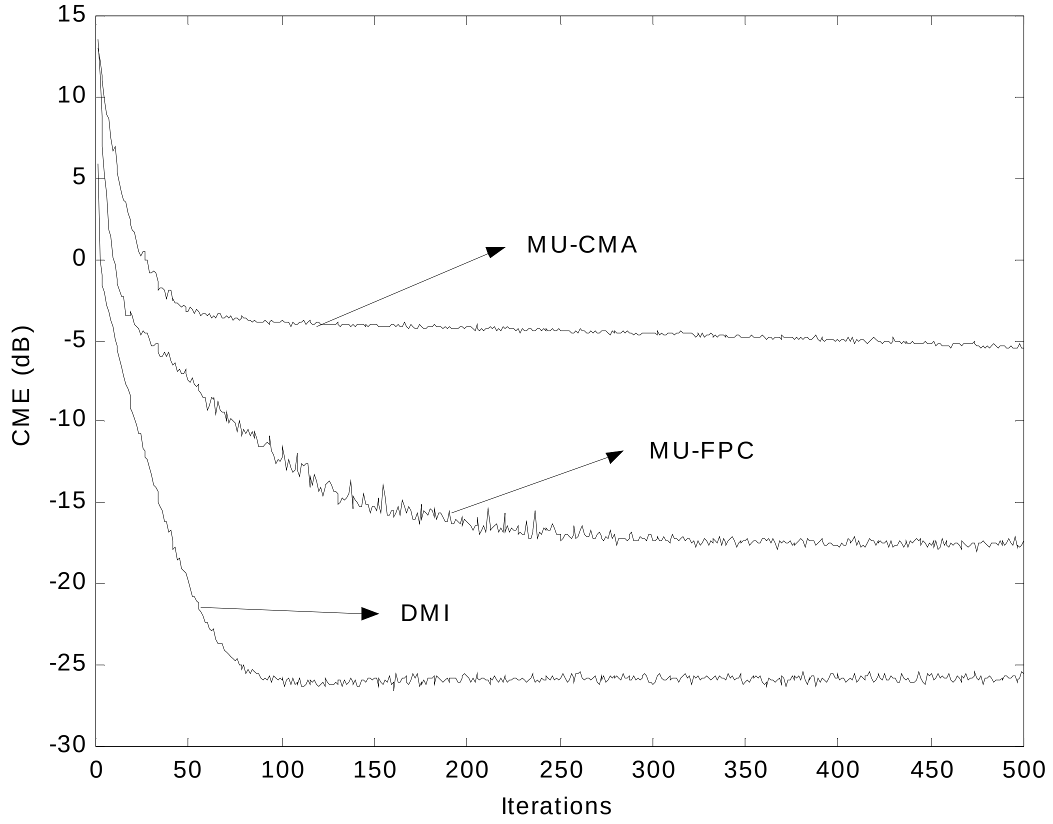
- Instantaneous mixture: SDMA system
  - 4 QPSK users
  - 8 antennas
  - SNR = 20 dB
  - Merit figures

User #	**DOA (degrees)**
1	1
2	-52
3	29
4	76

$$\text{CME}_k(n) = \left(\left|y_k(n)\right|^2 - 1\right)^2$$

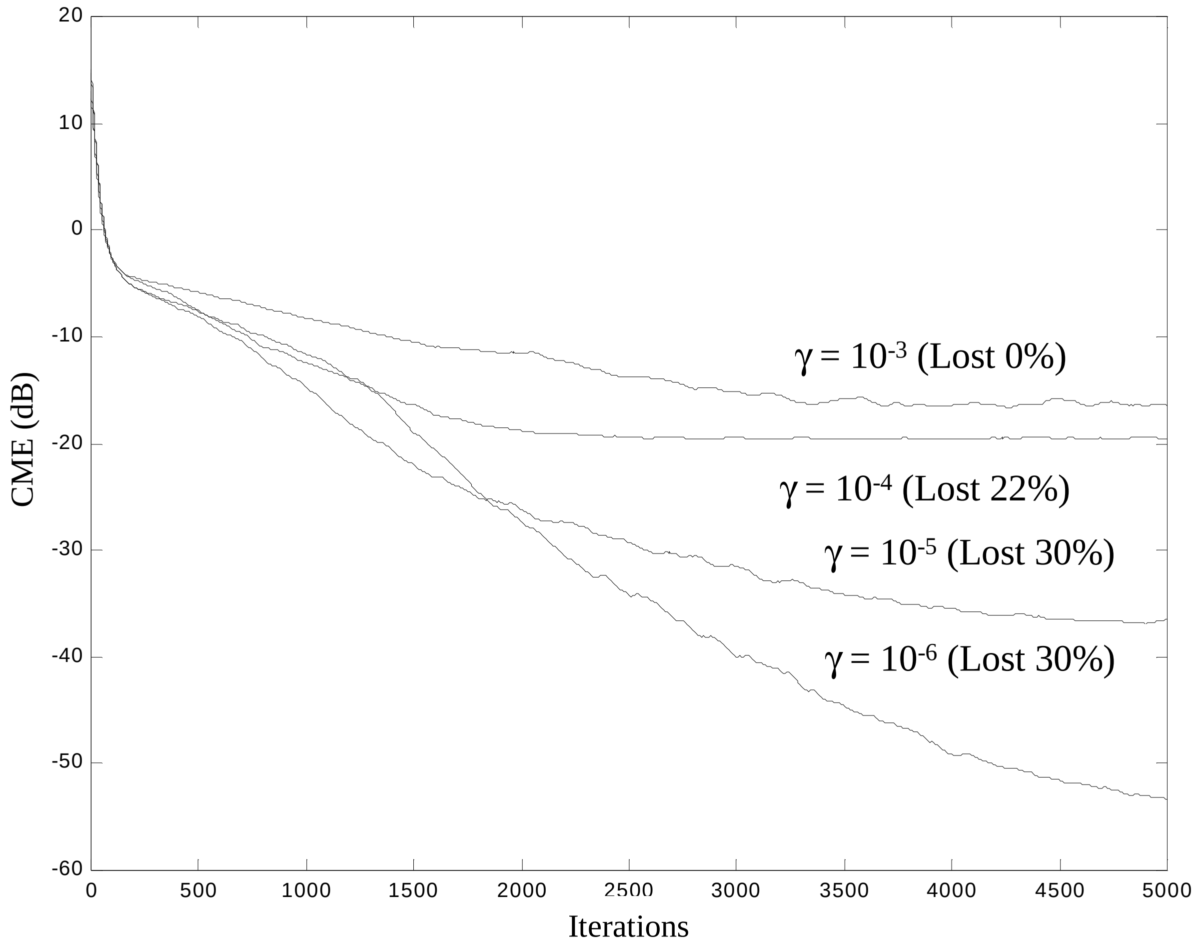
$$g_k(\theta) = \left|\mathbf{w}_k^H \mathbf{h}(\theta)\right|, \quad \theta \in (0, 2\pi)$$

Results MU-FPA: SDMA (cont.)



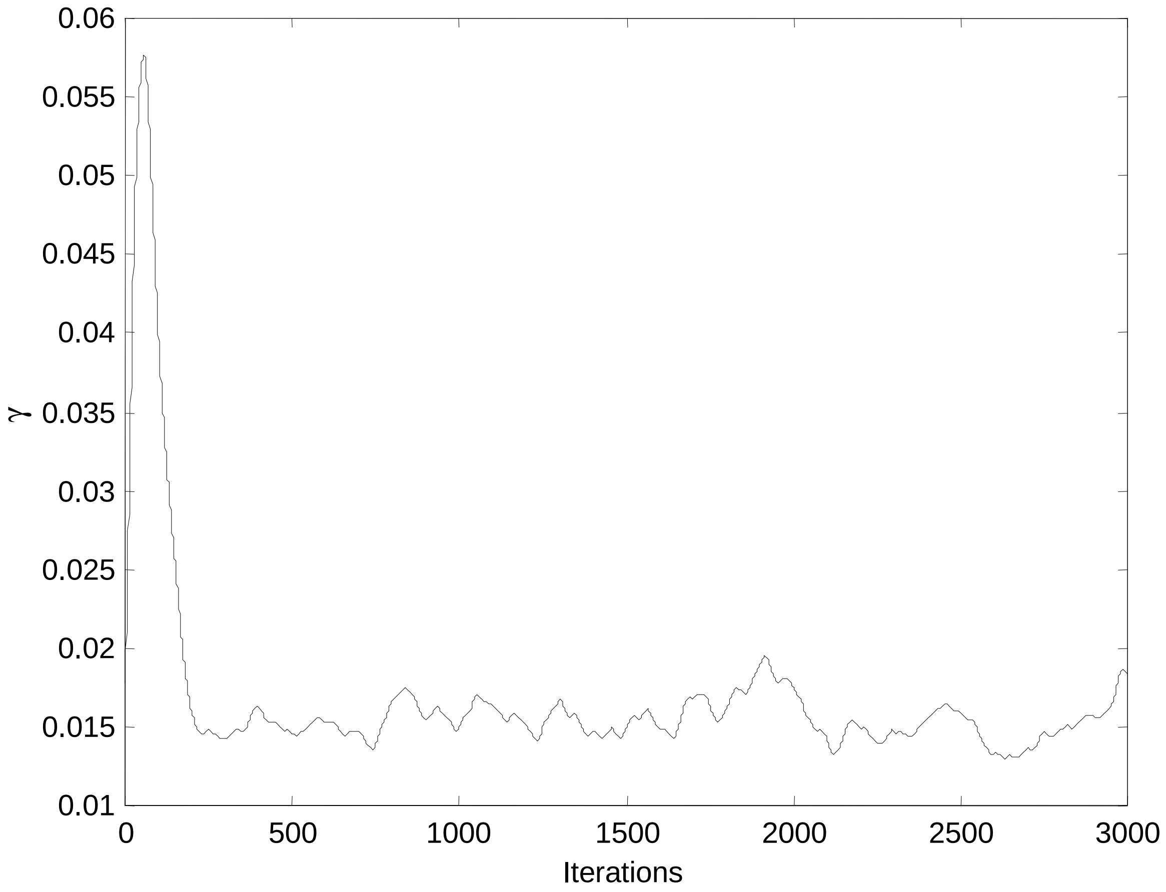
- Faster convergence
- Robustness

Results MU-FPA: SDMA (cont.)



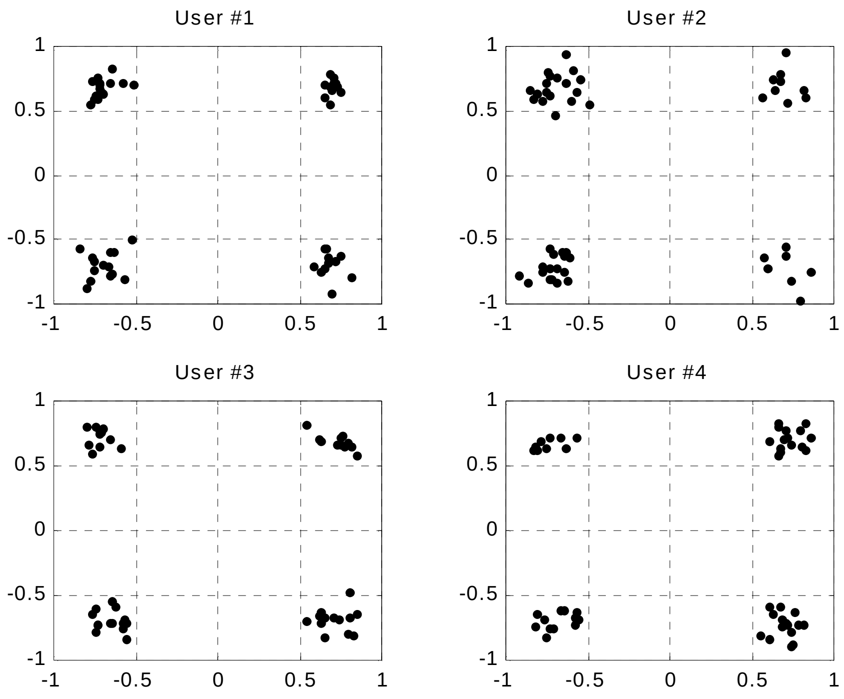
Probability of erroneous capture of users

Results MU-FPA: SDMA (cont.)

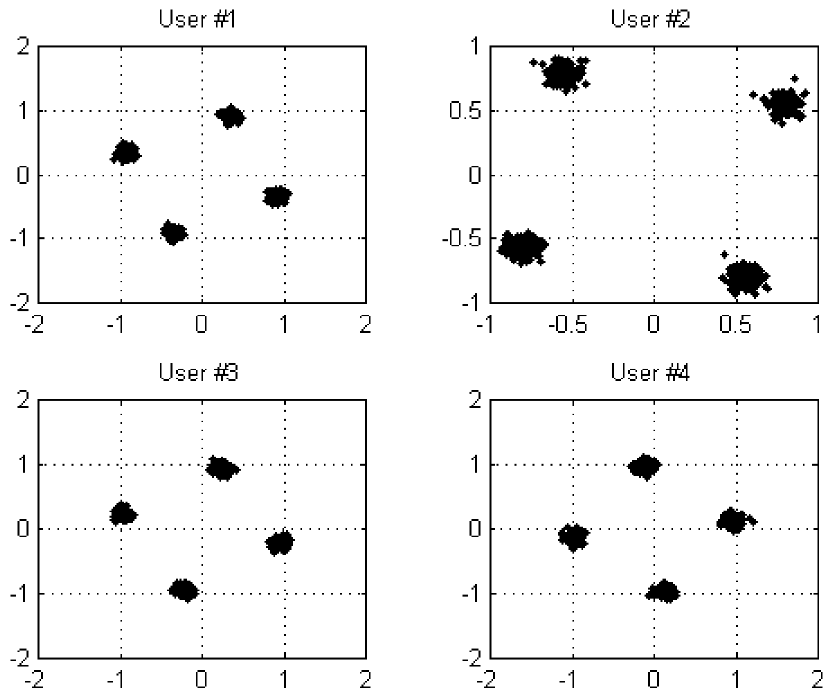


Temporal evolution of the decorrelation factor

Results MU-FPA: SDMA (cont.)



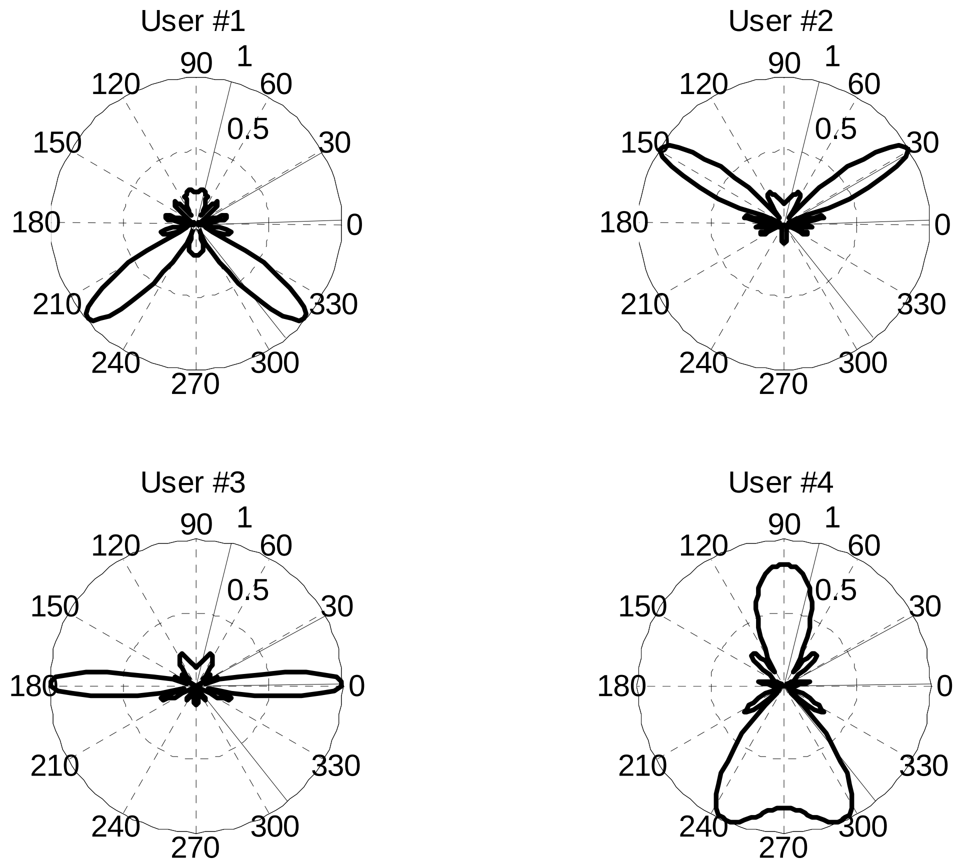
MU-FPA



MU-CMA

Results MU-FPA: SDMA (cont.)

User #	DOA (degrees)
1	1
2	-52
3	29
4	76



Results MU-FPA: *wireless*

- Space-time processing (model)

$$\mathbf{x}(n) = \sum_{k=1}^{K} \sum_{l=0}^{L-1} h_k(l) a_k(n-l) \mathbf{h}\left(\theta_{k,l}\right) + \mathbf{v}(n)$$

Received signal

$$y_k(n) = \mathbf{W}_k^H(n)\mathbf{X}(n)$$

Signal at the receiver output

$$\mathbf{W}_k(n) = \left[\mathbf{w}_{k0}^T(n) \quad \cdots \quad \mathbf{w}_{k(N-1)}^T(n)\right]^T$$

$$\mathbf{X}(n) = \left[\mathbf{x}^T(n) \quad \cdots \quad \mathbf{x}^T(n-N+1)\right]^T$$

Matricial form

Results MU-FPA: *wireless* (cont.)

- Space-time processing (model)

$$\mathbf{H}_k = \mathbf{F}(\theta_k)\mathbf{h}_k$$

$$\mathbf{F}(\theta_k) = \left[\mathbf{f}(\theta_{k,0}) \quad | \quad \mathbf{f}(\theta_{k,1}) \quad | \quad \cdots \quad | \quad \mathbf{f}(\theta_{k,L-1})\right]_{M\times L}$$

$$\mathbf{h}_k = \mathrm{diag}\left(\left[h_k(0) \quad \cdots \quad h_k(L-1)\right]\right)$$

Channel model

$$\mathbf{X}(n) = \sum_{k=1}^{K} \mathcal{H}_k \mathbf{A}_k(n) + \mathbf{V}(n)$$

$$\mathbf{A}_k(n) = \left[a_k(n) \quad \cdots \quad a_k(n-L-N+2)\right]^T$$

$$\mathbf{V}(n) = \left[\mathbf{v}^T(n) \quad | \quad \cdots \quad | \quad \mathbf{v}^T(n-N+1)\right]^T$$

$$\mathcal{H}_k = \begin{bmatrix} \mathbf{H}_k & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_k & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_k \end{bmatrix}_{MN\times(N+L-1)}$$

Results MU-FPA: *wireless* (cont.)

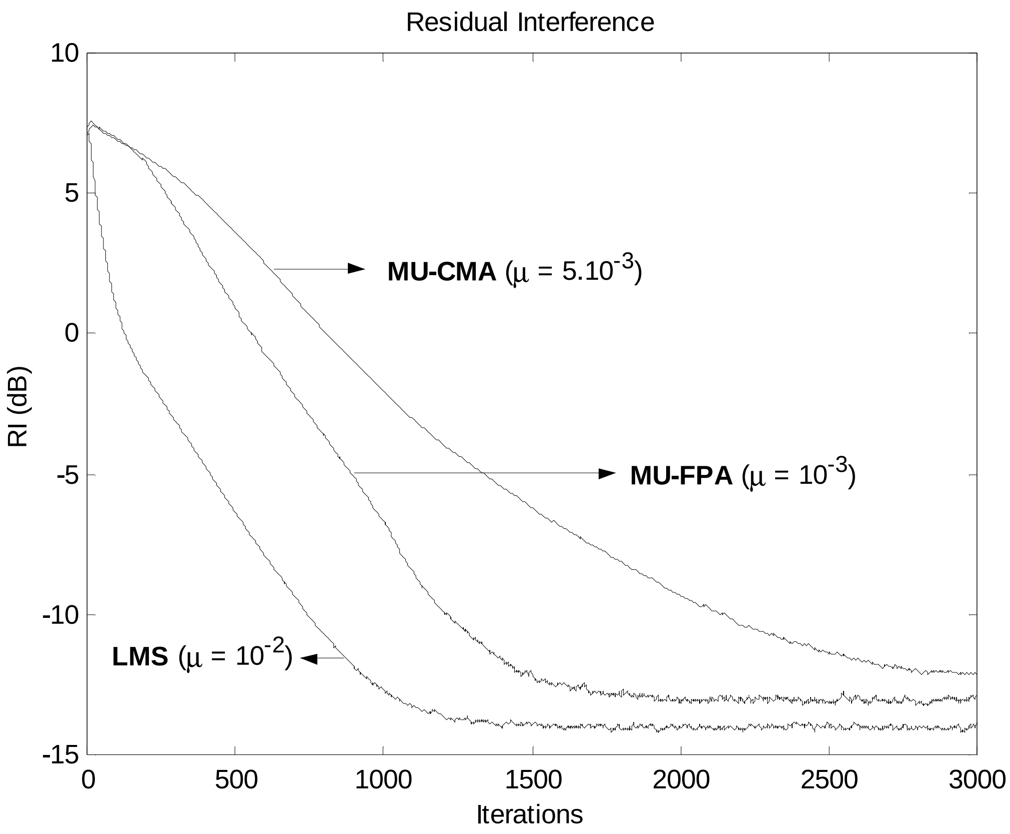
- 2 users (2 multipaths each)
- 4 antennas
- L = 2 (# of filter coefficients)
- D = 2 (# of considered delays)
- SNR = 30 dB
- Metric

$$RI_k(n) = \frac{\left| \sum_i \left|g_{k,i}(n)\right| - \max_i \left|g_{k,i}(n)\right| \right|}{\max_i \left|g_{k,i}(n)\right|}$$

Results MU-FPA: *wireless* (cont.)

User	**Path 1**		**Path 2**	
	Delay (x*T*)	DOA (rad)	Delay (x*T*)	DOA (rad)
#1	0.1	2π/5	1.1	-π/3
#2	0.4	π/7	1.2	-π/6

Results MU-FPA: *wireless*



- Faster
- Robustness

Constrained criterion

- Base on Shalvi-Weinstein theorem: generalization for MIMO case;
- *Conditions:*

C1. $a_k(n)$ is i.i.d. and zero mean $(l = 1, \ldots, K)$;

C2. $a_k(n)$ and $a_q(n)$ are statistically independent for $k \neq q$ and has the same pdf;

C3. $\left|\kappa\left(y_l(n)\right)\right| = \kappa_a \qquad (k = 1, \ldots, K)$

C4. $E\left\{\left|y_l(n)\right|^2\right\} = \sigma_a^2 \qquad (k = 1, \ldots, K)$

C5. $E\left\{y_k(n)\, y_q^*(n)\right\} = 0 \qquad k \neq q$

Constrained criterion (cont.)

- Can be split into two parts
  - *Equalization*
  - *Separation*
- Equalization
  - FP criterion
- Separation
  - Constraint about the system global response matrix: Gram-Schimdt orthogonalization

Constrained criterion (cont.)

- ***Multi-User Constrained Fitting pdf Algorithm (MU-CFPA)***

$$\begin{cases} \max_{\mathbf{W}} \; J_{\mathrm{FP}}(\mathbf{W}) = \sum_{j=1}^{K} D\left(p_{Y,\mathrm{ideal}}(y) \,\middle\|\, \Phi(y_j)\right) \\ \text{sujeito a: } \mathbf{G}^H\mathbf{G} = \mathbf{I} \end{cases}$$

- Features
  - pdf estimation by means of a parametric model
  - Can be understood as some kind of *Parzen estimation!*

MU-CFPA

MU-CFPA

1. Inicialize $\mathbf{W}(0)$
2. For $n > 0$
3. Compute $\mathbf{W}^e$

$$\mathbf{W}^e(n+1) = \mathbf{W}(n) - \mu \cdot \nabla J_{FP}(\mathbf{W}(n))$$

4. Normalize $\mathbf{w}_1$

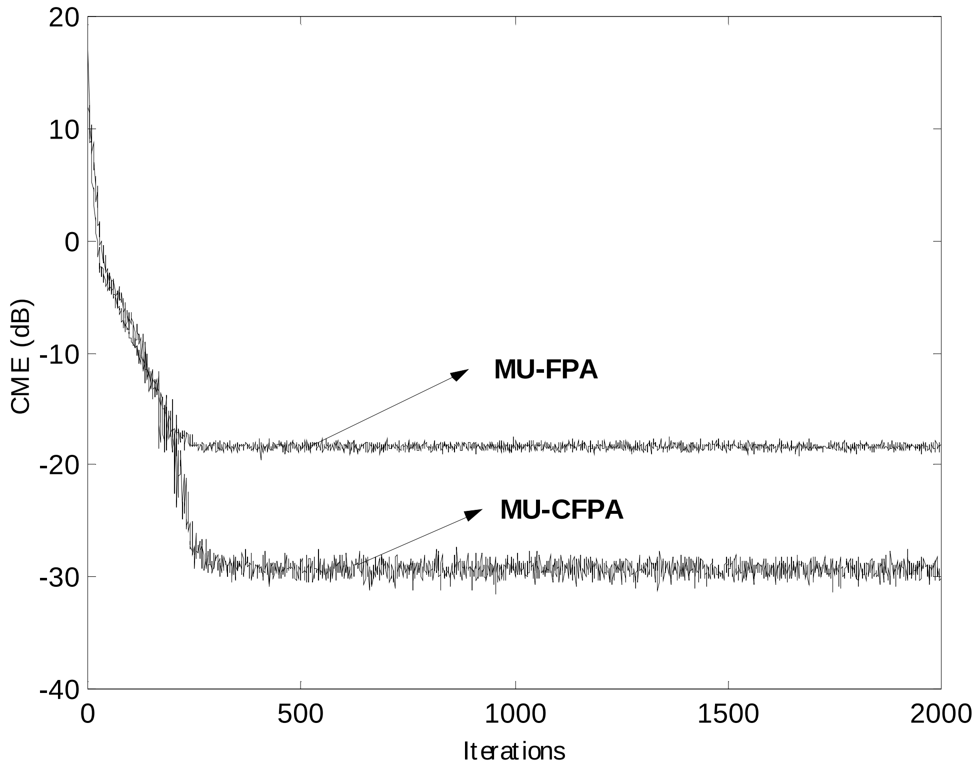
$$\mathbf{w}_1(n+1) = \frac{\mathbf{w}_1^e(n+1)}{\left\| \mathbf{w}_1^e(n+1) \right\|}$$

5. For $j = 2{:}K$
6. Compute $\mathbf{w}_j$

$$\mathbf{w}_j(n+1) = \frac{\mathbf{w}_j^e(n+1) - \sum_{l=1}^{j-1} \left[\mathbf{w}_l^H(n+1)\, \mathbf{w}_l^e(n+1) \right] \mathbf{w}_l(n+1)}{\left\| \mathbf{w}_j^e(n+1) - \sum_{l=1}^{j-1} \left[\mathbf{w}_l^H(n+1)\, \mathbf{w}_l^e(n+1) \right] \mathbf{w}_l(n+1) \right\|}$$

Results MU-CFPA

- Comparation with MU-FPA



- Lower steady-state error
- Gain of 10 dB

Results MU-CFPA x MUK

- *Setup* of simulation
  - $K = 2$ and $M = 2$
  - QPSK symbols
  - SNR = 30 dB
  - Initializations:

$$\mu_{\text{MUK}} = 2\times 10^{-3}$$
$$\mu_{\text{MU-CFPA}} = 10^{-2}$$
$$\sigma_r^2 = 0.1$$
$$\mathbf{W}(0) = \mathbf{W}^e(0) = \mathbf{I}$$

Mixture matrix

$$\mathbf{H} = \begin{bmatrix} 0.701 + j0.172 & 0.629 + j0.286 \\ -0.274 - j0.634 & 0.159 + j0.704 \end{bmatrix}$$

- Metric: Constant Modulus Error (CME)

$$\text{CME}_k(n) = \left(|y_k|^2 - 1\right)^2$$

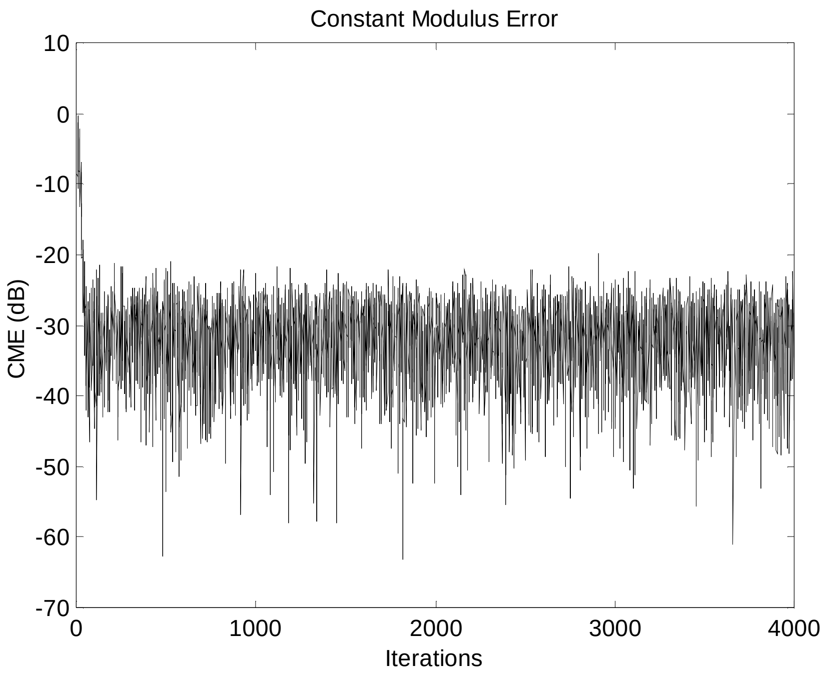
Results MU-CFPA x MUK (cont.)

- After convergence

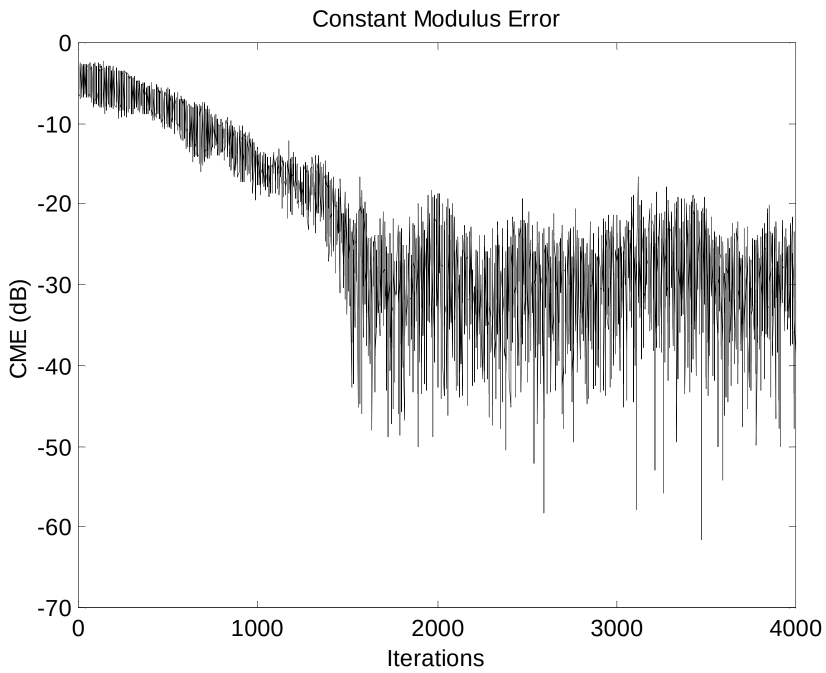
$$\mathbf{G}_{\text{MU-CFPA}} = \begin{bmatrix} -0.0015 - j0.0024 & -0.9990 + j0.0067 \\ 0.9992 - j0.0046 & -0.0017 + j0.0026 \end{bmatrix}$$

$$\mathbf{G}_{\text{MUK}} = \begin{bmatrix} -0.0143 + j0.0087 & -0.9369 - j0.3463 \\ 0.3295 + j0.9431 & -0.0091 + j0.0139 \end{bmatrix}$$

Results MU-CFPA x MUK (cont.)



MU-CFPA



MUK

Higher-Order Statistics (HOS)

- Which HOS are important for BSS?
  - At least one.
- Do we gain if using more?
  - Yes, in adaptive versions!
- How to verify that?
  - pdf estimation of the desired signal

pdf approximations

- Gram-Charlier expansion

$$p_Y(y) = \alpha(y)\left(1 + \sum_{i=3}^{\infty} c_i P_i(y)\right)$$

- Edgeworth expansion $i = (0),(3),(4,6),(5,7,9)$

$$p_Y(y) = \alpha(y)\left(1 + \frac{c_3}{3!}P_3(y) + \frac{c_4}{4!}P_4(y) + \frac{10c_3^2}{6!}P_6(y) + \frac{c_5}{5!}P_5(y) + \frac{35c_3c_4}{7!}P_7(y) + \frac{280c_3^3}{9!}P_9(y) + \frac{c_6}{6!}P_6(y) + \cdots\right)$$

Hermite polynomials

- Expression dependent of Gaussian pdf

$$\frac{\partial^i \alpha(y)}{\partial y^i} = (-1)^i \alpha(y) P_i(y)$$

$$P_0(y) = 1$$
$$P_1(y) = y$$
$$P_2(y) = y^2 - 1$$
$$P_3(y) = y^3 - 3y$$
$$P_4(y) = y^4 - 6y^2 + 3$$
$$P_5(y) = y^5 - 10y^3 + 15y$$
$$P_6(y) = y^6 - 15y^4 + 45y^2 - 15$$

$$P_{i+1}(y) = yP_i(y) - iP_{i-1}(y)$$

HOS influence

- Adaptive estimation of the signal from the time samples
- Pdf of the signal in the beginning of the adaptation has higher dependence of the higher order cumulants: faster convergence
- Thus, criteria that consider more HOS present faster convergence
- Complexity increases with the increase of the number of HOS

New pdf expansion

- Expansion on orthornormal series
- Use of a ***reference density***
- Edgeworth and Gram-Charlier: around Gaussian distribution
  - Hermite polynomials
- Edgeworth: decrescent order of the coefficients – useful for truncation
- The reference should be ***"similar"*** to the desired density

New pdf expansion (cont.)

- Representation

$$p_Y(y) = p_0(y) \cdot \sum_{k=1}^{\infty} C_k \cdot \mathfrak{h}_k(y)$$

Orthonormal serie

Coefficients of the serie

$$\Rightarrow \quad \mathfrak{h}_k(y) = (-1)^k \cdot \frac{1}{p_0(y)} \cdot \frac{d^k p_0(y)}{dy^k}$$

New pdf expansion (cont.)

- Expansion about Gaussian mixture densities
- Approximation of pdfs observed in digital communications
- Consideration of a different reference density
- Use of **Parzen estimation**: “*sum of Gaussians is able to approximate a wide range of densities*”
- Density reference given by summation of S Gaussian kernels

New pdf expansion (cont.)

$$p_{GM}(y) = \sum_{i=1}^{S} \frac{1}{\sqrt{2\pi\sigma_{\vartheta}^2}} \exp\left[-\frac{|y - a_i|}{2\sigma_{\vartheta}^2}\right] \cdot \Pr(a_i)$$

$$\mathfrak{c}_{k+1}(y) = S \cdot y \cdot \mathfrak{c}_k(y) - k \cdot \mathfrak{c}_{k-1}(y)$$

Recursion rule

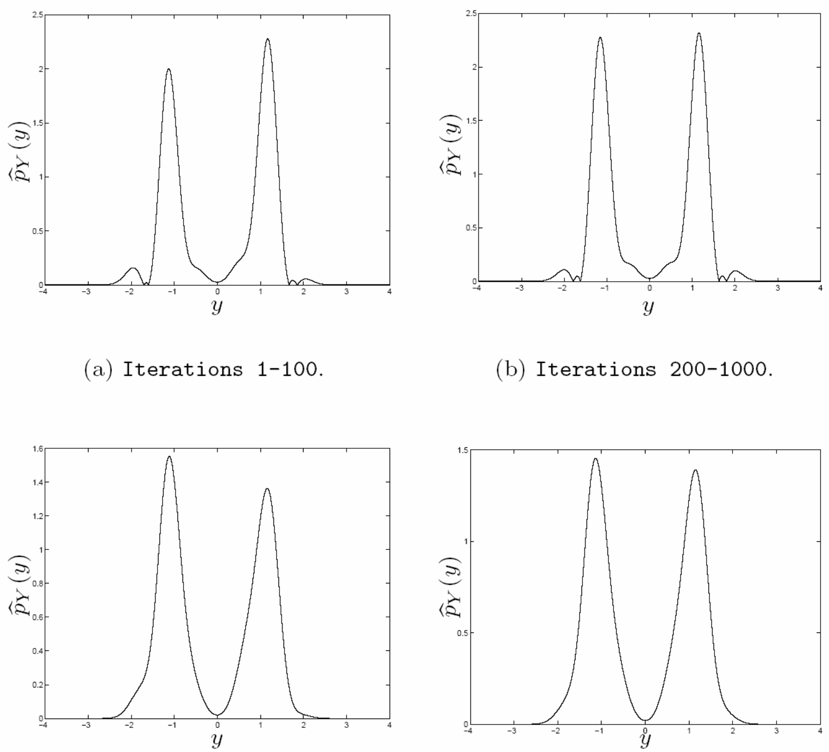
$$p_Y(y) = p_{GM}(y) \cdot \sum_{k=1}^{\infty} C_k \cdot \mathfrak{c}_k(y)$$

New pdf expansion

New pdf expansion (cont.)

- Evaluation of blind algorithms
- Verification of influence of HOS
- Different strategies
  - Kurtosis
  - *All* HOS

Adaptive pdf estimation



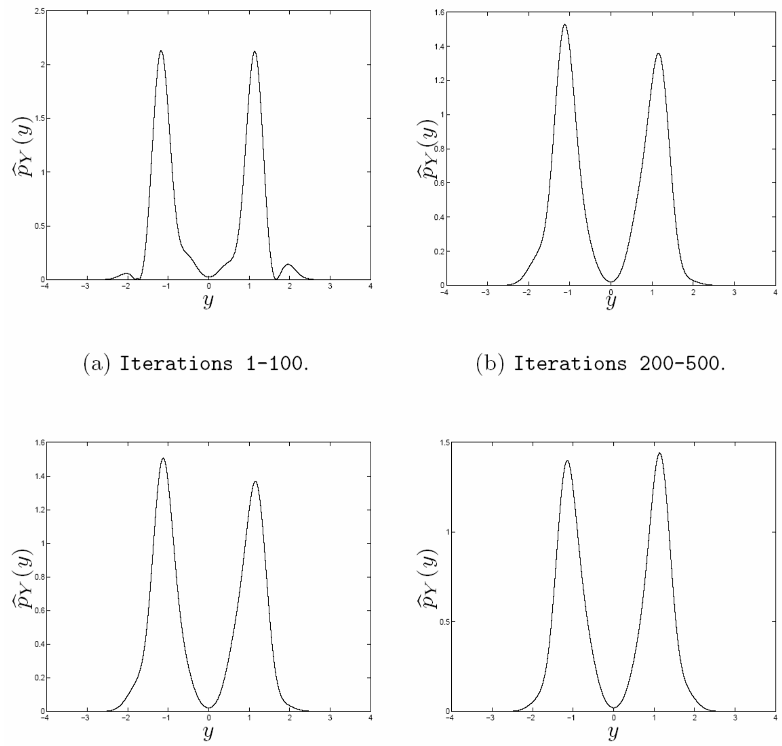
(a) Iterations 1-100.

(b) Iterations 200-1000.

(c) Iterations 1500-2000.

(d) Iterations 2500-4000.

MUK



(a) Iterations 1-100.

(b) Iterations 200-500.

(c) Iterations 500-1000.

(d) Iterations 2500-4000.

MU-CFPA

Use of HOS for blind criteria

- Use the pdf estimation to drive the design of blind criteria in adaptive processing
  - Complexity
  - Convergence
  - Robustness

Application: *S-ALOHA*

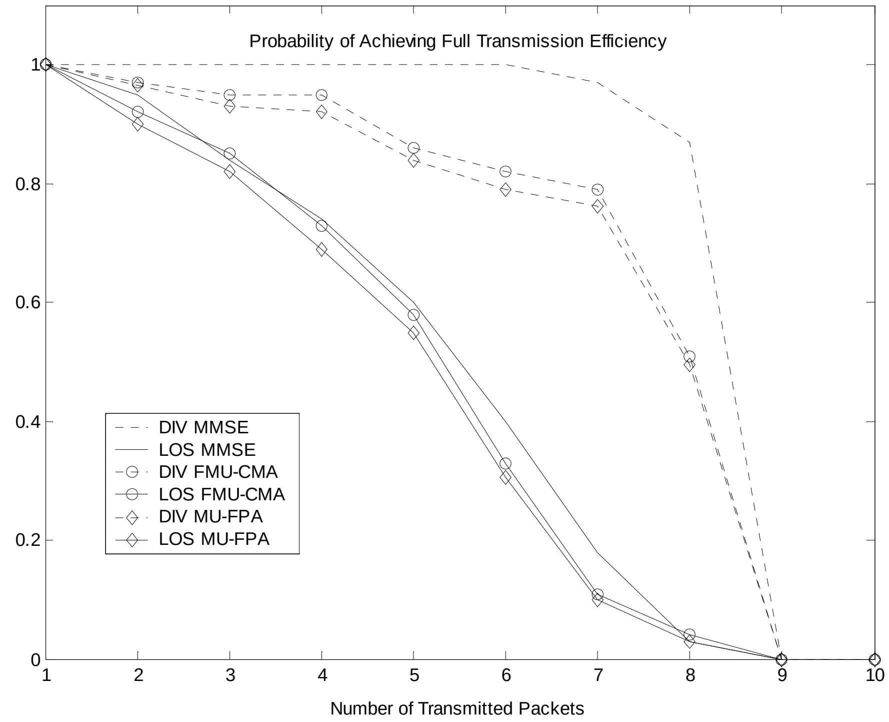
- Packet collision resolution in *Slotted-ALOHA*
- Similar to user separation

Application: *S-ALOHA* (cont.)

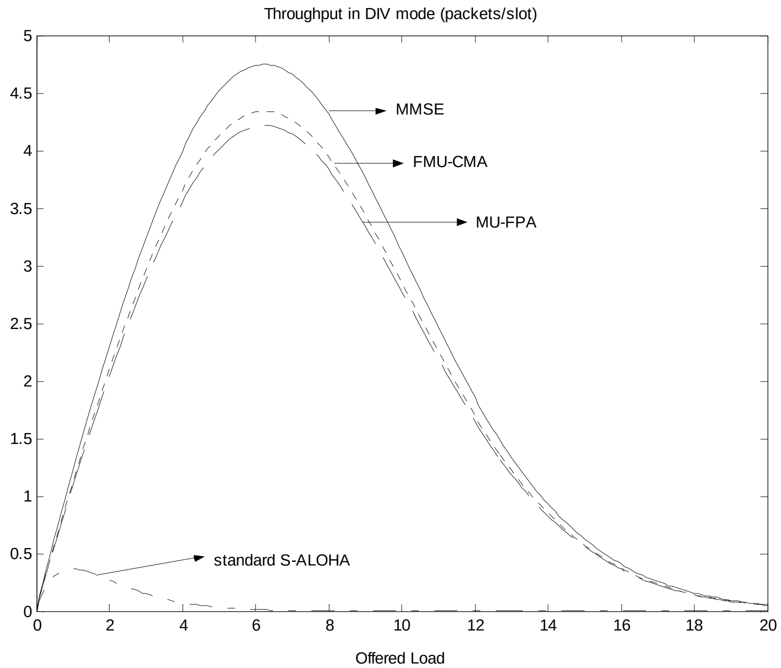
- Parameters
  - 8 antenna elements in the array
  - 500 QPSK symbols
  - SNR = 20 dB
  - 180° cell
  - Scenarios: Line-Of-Sight (LOS) and Diversity (DIV)
  - 200 independent simulations for estimation of success probabilities

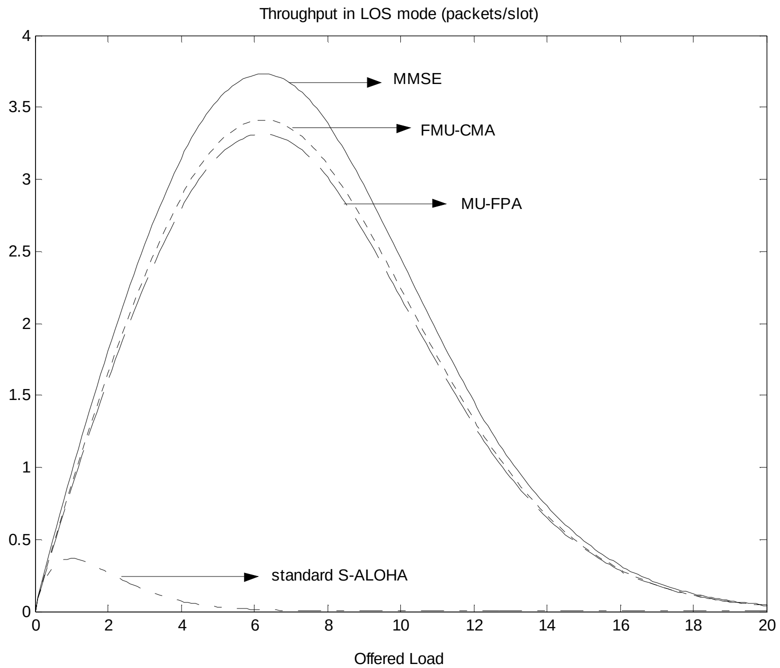
Application: *S-ALOHA* (cont.)

- ***Full efficiency:*** $P(K|K, M)$



Application: *S-ALOHA* (cont.)





DIV **LOS**

Conclusions

- Source separation techniques: powerful tool for communication system analysis
- Algorithms for multiuser processing
- Analysis of the HOS impact on adaptive processing

Perspectives

- Comparison with algorithms that minimize the error probability: is the same of making the variables independent?
- Under-determined problems
- Tensorial methods for application in space-time processing

Perspectives (cont.)

- Differential geometry
  - Optimization in non-Euclidean space
  - Manifolds are pdf spaces (Riemann spaces)
  - *Information geometry* (Amari's contributions)
  - Constrained filtering

Info

IX Signal Processing Advances in Wireless Communications (SPAWC2008)

Recife, Brazil

July, 3rd-6th

Deadline (tentative): January 28th, 2008

URL: www.spawc2008.org

Contact

- **Charles Casimiro Cavalcante**, Dr.

Email: charles@gtel.ufc.br

URL: www.gtel.ufc.br/~charles