WCCI 2008 Presentation

Complex-Valued Symmetric Radial Basis Function Classifier for Quadrature Phase Shift Keying Beamforming Systems

S. Chen, C.J. Harris and L. Hanzo

School of Electronics and Computer Science

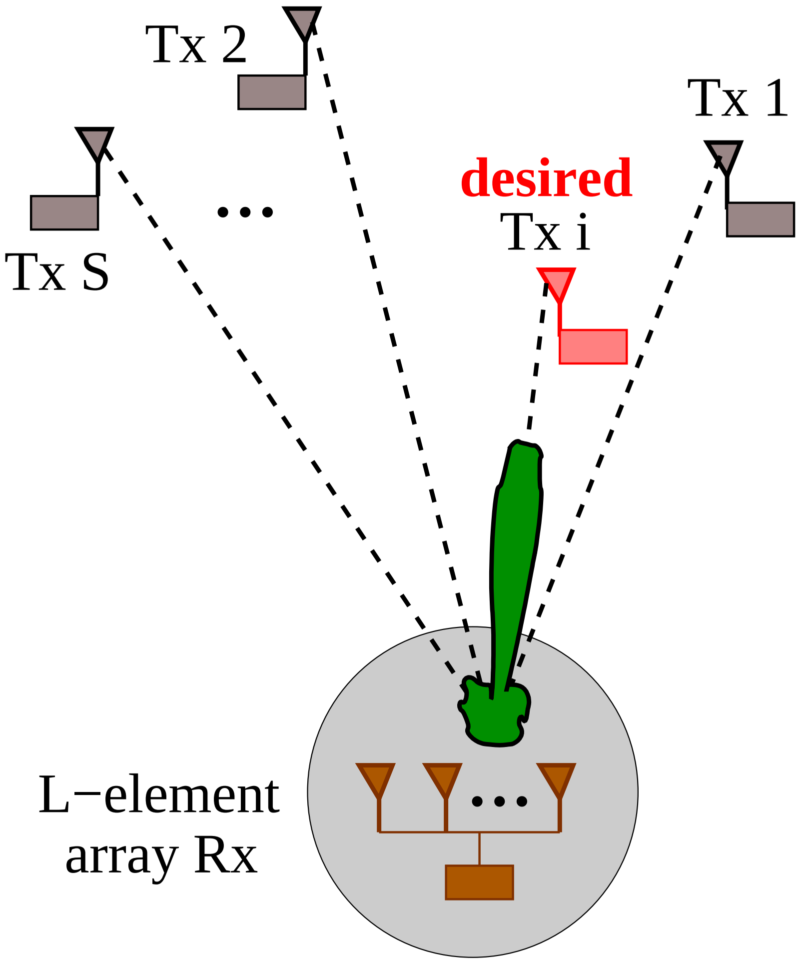
University of Southampton

Southampton SO17 1BJ, UK

Outline

- Existing **linear** beamforming techniques, and motivations for **nonlinear** beamforming
- Signal model and optimal Bayesian detection with an inherent **symmetry** property for QPSK beamforming
- **Complex-valued** symmetric radial basis function classifier by incorporating *a priori* knowledge
- **Multi-class** Fisher ratio of **class separability** measure based **orthogonal forward selection**
- Simulation investigation, and performance comparison

Motivations



- Classical beamforming is **linear** with a **beampattern** interpretation of beamformer's weight vector
  - maximise response at desired user **direction** and place nulls at interferers' directions, **must** $L \geq S$
  - similar to **zero-forcing** equalisation, and suffers from **noise enhancement**
- Best linear beamforming is **minimum bit error rate** (L-MBER)
  - significantly enhance achievable system BER and user capacity

Motivations (continue)

- Beamforming can be viewed as **classification**, which classifies received channel-impaired signal into most-likely transmitted symbol point
- In comparison with linear beamforming, **nonlinear** detection offers
  - significantly better BER performance and much larger user capacity, at cost of higher complexity
- With **posterior** or **conditional probabilities** as **generalised beam-pattern** interpretation
  - This nonlinear detection can be viewed as **nonlinear beamforming**
- A practical case for **complex-valued** radial basis function network
  - A strong motivation for **grey-box** RBF classifier: the art of incorporating *a priori* knowledge

Signal Model

- S single-transmit-antenna users transmit on same carrier, receiver is equipped with L-element **antenna array**, channels are non-dispersive
- Received signal vector $\mathbf{x}(k) = [x_1(k)\ x_2(k) \cdots x_L(k)]^T$ is

$$\mathbf{x}(k) = \mathbf{P}\,\mathbf{b}(k) + \mathbf{n}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k)$$

- $\mathbf{n}(k) = [n_1(k)\ n_2(k) \cdots n_L(k)]^T$ is noise vector, and **system matrix**

$$\mathbf{P} = [A_1\mathbf{s}_1\ A_2\mathbf{s}_2 \cdots A_M\mathbf{s}_S]$$

- $\mathbf{s}_m$: **steering vector** of source m, A_m: m-th non-dispersive channel tap
- User i is **desired** user, and transmitted symbol vector $\mathbf{b}(k) = [b_1(k)\ b_2(k) \cdots b_S(k)]^T$ with QPSK symbol set

$$b_m(k) \in \{b^{[1]} = +1+j,\ b^{[2]} = -1+j, b^{[3]} = -1-j,\ b^{[4]} = +1-j\}, 1 \le m \le S$$

Signal Space

- Denote $N_b = 4^S$ **legitimate sequences** of $\mathbf{b}(k)$ as $\mathbf{b}_q$, $1 \leq q \leq N_b$
- Noiseless **channel state** $\bar{\mathbf{x}}(k)$ takes values from set

$$\bar{\mathbf{x}}(k) \in \mathcal{X} = \{\bar{\mathbf{x}}_q = \mathbf{P}\,\mathbf{b}_q, 1 \leq q \leq N_b\}$$

which can be divided into **four subsets** conditioned on $b_i(k) = b^{[m]}$

$$\mathcal{X}^{[m,i]} \triangleq \{\bar{\mathbf{x}}_q^{[m,i]} \in \mathcal{X}, 1 \leq q \leq N_{sb} : b_i(k) = b^{[m]}\}, \ 1 \leq m \leq 4$$

- **Conditional probabilities** of receiving $\mathbf{x}(k)$ given $b_i(k) = b^{[m]}$ are

$$p^{[m,i]}(\mathbf{x}(k)) = \sum_{q=1}^{N_{sb}} \beta_q e^{-\frac{\|\mathbf{x}(k) - \bar{\mathbf{x}}_q^{[m,i]}\|^2}{2\sigma_n^2}}, \ 1 \leq m \leq 4$$

$N_{sb} = N_b/4 = 4^{M-1}$, noise power is $2\sigma_n^2$ and all priors β_q are equal
- $p^{[m,i]}(\mathbf{x}(k))$ can be interpreted as **generalised beampatterns**

Optimal Bayesian Detector

- **Optimal detection** strategy is

$$\hat{b}_i(k) = b^{[m^*]} \quad \text{with} \quad m^* = \arg \max_{1 \le m \le 4} p^{[m,i]}(\mathbf{x}(k))$$

- Define complex-valued Bayesian **decision variable**

$$y_{\mathrm{Bay},i}(k) \triangleq b^{[1]} \cdot p^{[1,i]}(\mathbf{x}(k)) + b^{[2]} \cdot p^{[2,i]}(\mathbf{x}(k)) + b^{[3]} \cdot p^{[3,i]}(\mathbf{x}(k)) + b^{[4]} \cdot p^{[4,i]}(\mathbf{x}(k))$$

- Optimal **Bayesian** detection is: $\hat{b}_i(k) = \mathrm{sgn}(y_{\mathrm{Bay},i}(k))$, where

$$\mathrm{sgn}(y) = \begin{cases} b^{[1]} = +1 + j, & y_R \ge 0 \text{ and } y_I \ge 0, \\ b^{[2]} = -1 + j, & y_R < 0 \text{ and } y_I \ge 0, \\ b^{[3]} = -1 - j, & y_R < 0 \text{ and } y_I < 0, \\ b^{[4]} = +1 - j, & y_R \ge 0 \text{ and } y_I < 0, \end{cases}$$

Symmetry of Bayesian Solution

- Four state subsets satisfy following **symmetric** properties

$$\mathcal{X}^{[2,i]} = +j \cdot \mathcal{X}^{[1,i]}, \ \mathcal{X}^{[3,i]} = -1 \cdot \mathcal{X}^{[1,i]}, \ \mathcal{X}^{[4,i]} = -j \cdot \mathcal{X}^{[1,i]}$$

- Thus **Bayesian solution** becomes, for $\bar{\mathbf{x}}_q^{[1,i]} \in \mathcal{X}^{[1,i]}$,

$$y_{\mathrm{Bay},i}(k) = \sum_{q=1}^{N_{sb}} \left\{ b^{[1]}\beta \cdot e^{-\frac{\|\mathbf{x}(k)-\bar{\mathbf{x}}_q^{[1,i]}\|^2}{2\sigma_n^2}} + b^{[2]}\beta \cdot e^{-\frac{\|\mathbf{x}(k)-j\cdot\bar{\mathbf{x}}_q^{[1,i]}\|^2}{2\sigma_n^2}} \right.$$

$$\left. +b^{[3]}\beta \cdot e^{-\frac{\|\mathbf{x}(k)+\bar{\mathbf{x}}_q^{[1,i]}\|^2}{2\sigma_n^2}} + b^{[4]}\beta \cdot e^{-\frac{\|\mathbf{x}(k)+j\cdot\bar{\mathbf{x}}_q^{[1,i]}\|^2}{2\sigma_n^2}} \right\}$$

- If system **channel matrix** $\mathbf{P}$ can be estimated, as in **uplink**, subset $\mathcal{X}^{[1,i]}$ can be calculated and Bayesian solution is specified
- In **downlink**, receiver only has access to desired user's training data, estimating $\mathbf{P}$ is difficult, and other adaptive means has to be adopted

Symmetric RBF Network

- Consider **complex-valued radial basis function** network

$$y(k) = \sum_{q=1}^{M} \theta_q \phi_q(\mathbf{x}(k))$$

θ_q: complex-valued **weight**, $\phi_q(\mathbf{x}(k))$: complex-valued **RBF node**

- In view of known symmetric underlying signal space,

$$\begin{aligned} \phi_q(\mathbf{x}) \quad = \quad & b^{[1]} \cdot \varphi(\|\mathbf{x} - \mathbf{c}_q\|/\rho) + b^{[2]} \cdot \varphi(\|\mathbf{x} - j \cdot \mathbf{c}_q\|/\rho) \\ & + b^{[3]} \cdot \varphi(\|\mathbf{x} + \mathbf{c}_q\|/\rho) + b^{[4]} \cdot \varphi(\|\mathbf{x} + j \cdot \mathbf{c}_q\|/\rho) \end{aligned}$$

$\varphi(\bullet)$: real-valued **basis function**, $\mathbf{c}_q$: RBF **centre**, ρ^2: **RBF variance**

- **Task**: construct a **sparse** CV-SRBF classifier when given a block of training data $D_K = \{\mathbf{x}(k), d(k) = b_i(k)\}_{k=1}^{K}$

Training Model

- Given ρ^2, use $\mathbf{c}_q = \mathbf{x}(q)$, $1 \le q \le M = K$, define **modelling residual** $\varepsilon(q) = d(q) - y(q) \Rightarrow$ over **training set** D_K

$$\mathbf{d} = \mathbf{\Phi}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

$$\mathbf{d} = [d(1)\ d(2) \cdots d(K)]^T,\ \boldsymbol{\varepsilon} = [\varepsilon(1)\ \varepsilon(2) \cdots \varepsilon(K)]^T,\ \boldsymbol{\theta} = [\theta_1\ \theta_2 \cdots \theta_M]^T$$

- **Complex-valued regression matrix**

$$\mathbf{\Phi} = [\boldsymbol{\phi}_1\ \boldsymbol{\phi}_2 \cdots \boldsymbol{\phi}_M] \in \mathcal{C}^{K\times M}$$

with **column** vectors $\boldsymbol{\phi}_q = [\phi_q(\mathbf{x}(1))\ \phi_q(\mathbf{x}(2)) \cdots \phi_q(\mathbf{x}(K))]^T$, $1 \le q \le M$

- Goal: select subset model containing $M_{\rm spa}$ ($\ll M$) significant RBF nodes
  - RBF variance ρ^2: determined via **cross validation**
  - **Model size**: terminate selection when $M_{\rm spa} = N_{sb}$

Orthogonal Decomposition

- **Orthogonal decomposition** of $\mathbf{\Phi}$: $\mathbf{\Phi} = \mathbf{\Omega}\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 1 & \alpha_{1,2} & \cdots & \alpha_{1,M} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \alpha_{M-1,M} \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

with **complex-valued** $\alpha_{q,l}$, $1 \leq q < l \leq M$, and **orthogonal matrix**

$$\mathbf{\Omega} = [\boldsymbol{\omega}_1 \ \boldsymbol{\omega}_2 \cdots \boldsymbol{\omega}_M] = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,M} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,M} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{K,1} & \omega_{K,2} & \cdots & \omega_{K,M} \end{bmatrix}$$

- Equivalent model

$$\mathbf{d} = \mathbf{\Omega}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

with complex-valued weight vector $\boldsymbol{\gamma} = [\gamma_1 \ \gamma_2 \cdots \gamma_M]^T = \mathbf{A}\boldsymbol{\theta}$

Multi-Class Fisher Ratio

- Divide training data $\mathbf{X} = \{\mathbf{x}(k)\}_{k=1}^{K}$ into $M_C = 4$ **classes**

$$\mathbf{X}^{[q]} \triangleq \{\mathbf{x}(k) \in \mathbf{X} : d(k) = b^{[q]}\}, \; 1 \le q \le M_C$$

Number of samples in $\mathbf{X}^{[q]}$ is $K^{[q]}$ with $\sum_{q=1}^{M_C} K^{[q]} = K$

- **Mean** and **variance** of samples belonging to **class** $\mathbf{X}^{[q]}$ in **direction** $\boldsymbol{\omega}_l$

$$m_{q,l} = \frac{1}{K^{[q]}} \sum_{k=1}^{K} \delta\left(d(k) - b^{[q]}\right) \omega_{k,l}, \;\; \sigma_{q,l}^2 = \frac{1}{K^{[q]}} \sum_{k=1}^{K} \delta\left(d(k) - b^{[q]}\right) \left(\omega_{k,l} - m_{q,l}\right)^2$$

where $\delta(x) = 1$ for $x = 0 + j0$ and $\delta(x) = 0$ for $x \neq 0 + j0$

- **Fisher ratio** of **class separation** between $\mathbf{X}^{[p]}$ and $\mathbf{X}^{[q]}$ in direction $\boldsymbol{\omega}_l$

$$F_{p,q,l} = \left(m_{p,l} - m_{q,l}\right)^2 / \left(\sigma_{p,l}^2 + \sigma_{q,l}^2\right)$$

Ratio of **interclass difference** to **intraclass spread**

OFS Based on FRCSM

- **Average** Fisher ratio of class separation in direction $\boldsymbol{\omega}_l$

$$F_l = \frac{2}{(M_C - 1)M_C} \sum_{p=1}^{M_C-1} \sum_{q=p+1}^{M_C} F_{p,q,l}$$

  Fisher ratio provides a good **class separability** measure

- **Orthogonal decomposition** makes computation of Fisher ratio of class separation measure very efficient

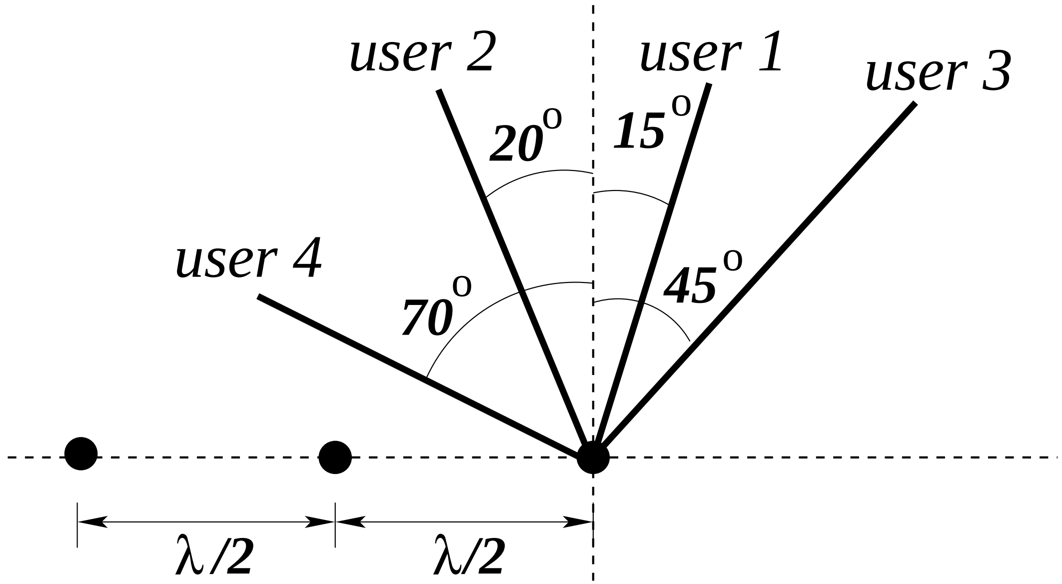
- Based on FRCSM, significant RBF nodes is selected in an OFS procedure

- At l-th stage of **orthogonal forward selection** procedure
  - A node is chosen as l-th term in selected CV-SRBF classifier if it produces **largest** F_l among candidates $\boldsymbol{\omega}_p$, $l \le p \le M$

- Procedure is terminated with a **sparse** classifier of $M_{\text{spa}} = N_{sb}$ terms

Simulation Set Up

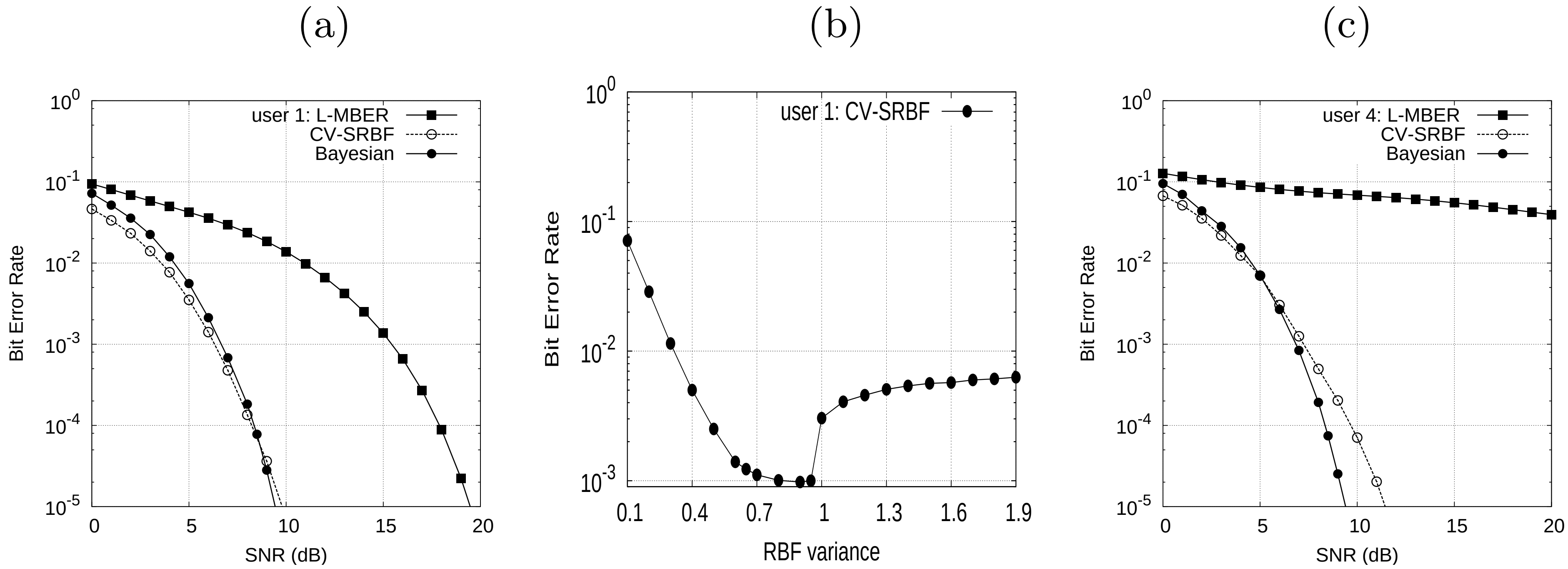
- Three-element antenna array having half wavelength spacing to support four QPSK users
- Angular locations of four users as illustrated
- Simulated channel conditions were $A_i = 1 + j0$, $1 \leq i \leq 4$
- All four users had an equal signal power



- Given each SNR, $K = 600$ training data were generated to train CV-SRBF classifier
- Since number of signal states $N_{sb} = 64$, $M_{\text{spa}} = 64$ terms were selected using OFS based on FRCSM

Simulation Results

(a) User-one bit error rate performance comparison, (b) Influence of RBF variance ρ^2 on bit error rate performance of user-one CV-SRBF classifier given SNR$=6$ dB, and (c) User-four bit error rate performance comparison



Conclusions

- We propose complex-valued symmetric radial basis function classifier for QPSK nonlinear beamforming
- Grey-box model by incorporating *a priori* knowledge
- Orthogonal forward selection based on multi-class Fisher ratio of class separability measure
- Select sparse CV-SRBF classifier from training data efficiently with excellet test bit error rate performance