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 $\mathbf{P} = [A_1s_1 \ A_2s_2 \cdots A_Ms_S]$
- $s_m$: 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
  \[
  b_m(k) \in \{b^{[1]} = +1+j, \ b^{[2]} = -1+j, \ b^{[3]} = -1-j, \ b^{[4]} = +1-j\}, \ 1 \leq m \leq S
  \]
Denote $N_b = 4^S$ legitimate sequences of $b(k)$ as $b_q$, $1 \leq q \leq N_b$

Noiseless channel state $\bar{x}(k)$ takes values from set

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

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

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

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

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

$N_{sb} = N_b/4 = 4^{M-1}$, noise power is $2\sigma_n^2$ and all priors $\beta_q$ are equal

$p^{[m,i]}(x(k))$ can be interpreted as generalised beampatterns
Optimal Bayesian Detector

- **Optimal detection** strategy is
  
  \[
  \hat{b}_i(k) = b^{[m^*]} \quad \text{with} \quad m^* = \arg \max_{1 \leq m \leq 4} p^{[m,i]}(x(k))
  \]

- Define complex-valued Bayesian decision variable
  
  \[
  y_{Bay,i}(k) \triangleq b^{[1]} \cdot p^{[1,i]}(x(k)) + b^{[2]} \cdot p^{[2,i]}(x(k)) + b^{[3]} \cdot p^{[3,i]}(x(k)) + b^{[4]} \cdot p^{[4,i]}(x(k))
  \]

- Optimal Bayesian detection is: \( \hat{b}_i(k) = \text{sgn}(y_{Bay,i}(k)) \), where
  
  \[
  \text{sgn}(y) = \begin{cases} 
  b^{[1]} = +1 + j, & y_R \geq 0 \text{ and } y_I \geq 0, \\
  b^{[2]} = -1 + j, & y_R < 0 \text{ and } y_I \geq 0, \\
  b^{[3]} = -1 - j, & y_R < 0 \text{ and } y_I < 0, \\
  b^{[4]} = +1 - j, & y_R \geq 0 \text{ and } y_I < 0,
  \end{cases}
  \]
Symmetry of Bayesian Solution

- Four state subsets satisfy following **symmetric** properties
  \[
  \mathcal{X}^{[2,i]} = +j \cdot \mathcal{X}^{[1,i]}, \quad \mathcal{X}^{[3,i]} = -1 \cdot \mathcal{X}^{[1,i]}, \quad \mathcal{X}^{[4,i]} = -j \cdot \mathcal{X}^{[1,i]}
  \]

- Thus **Bayesian solution** becomes, for \( \bar{x}^{[1,i]}_{q} \in \mathcal{X}^{[1,i]} \),
  \[
  y_{\text{Bay},i}(k) = \sum_{q=1}^{N_{sb}} \left\{ b^{[1]} \beta \cdot e^{-\frac{\|x(k) - \bar{x}^{[1,i]}_{q}\|^2}{2\sigma_n^2}} + b^{[2]} \beta \cdot e^{-\frac{\|x(k) - j \cdot \bar{x}^{[1,i]}_{q}\|^2}{2\sigma_n^2}} 
  
  + b^{[3]} \beta \cdot e^{-\frac{\|x(k) + \bar{x}^{[1,i]}_{q}\|^2}{2\sigma_n^2}} + b^{[4]} \beta \cdot e^{-\frac{\|x(k) + j \cdot \bar{x}^{[1,i]}_{q}\|^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

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

- In view of known symmetric underlying signal space,

\[ \phi_q(x) = b[1] \cdot \phi(\|x - c_q\|/\rho) + b[2] \cdot \phi(\|x - j \cdot c_q\|/\rho) \\
+ b[3] \cdot \phi(\|x + c_q\|/\rho) + b[4] \cdot \phi(\|x + j \cdot c_q\|/\rho) \]

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

Given $\rho^2$, use $c_q = x(q), \ 1 \leq q \leq M = K$, define modelling residual $\varepsilon(q) = d(q) - y(q) \Rightarrow$ over training set $D_K$

$$d = \Phi \theta + \varepsilon$$

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

Complex-valued regression matrix

$$\Phi = [\phi_1 \ \phi_2 \cdots \phi_M] \in \mathbb{C}^{K \times M}$$

with column vectors $\phi_q = [\phi_q(x(1)) \ \phi_q(x(2)) \cdots \phi_q(x(K))]^T, \ 1 \leq q \leq M$

Goal: select subset model containing $M_{\text{spa}} \ (\ll M)$ significant RBF nodes

- RBF variance $\rho^2$: determined via cross validation
- Model size: terminate selection when $M_{\text{spa}} = N_{sb}$
Orthogonal Decomposition

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

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

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

$$\Omega = [\omega_1 \omega_2 \cdots \omega_M] = \begin{bmatrix}
\omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,M} \\
\omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,M} \\
\vdots & \vdots & \ddots & \vdots \\
\omega_{K,1} & \omega_{K,2} & \cdots & \omega_{K,M}
\end{bmatrix}$$

- **Equivalent model**

$$d = \Omega \gamma + \varepsilon$$

with complex-valued weight vector $\gamma = [\gamma_1 \gamma_2 \cdots \gamma_M]^T = A \theta$
Multi-Class Fisher Ratio

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

$$X^{[q]} \triangleq \{x(k) \in X : d(k) = b^{[q]}\}, \ 1 \leq q \leq MC$$

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

- Mean and variance of samples belonging to class $X^{[q]}$ in direction $\omega_l$

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

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

- Fisher ratio of class separation between $X^{[p]}$ and $X^{[q]}$ in direction $\omega_l$

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

Ratio of interclass difference to intraclass spread
OFS Based on FRCSM

- **Average** Fisher ratio of class separation in direction $\omega_l$

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

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 $\omega_p$, $l \leq p \leq M$

- Procedure is terminated with a **sparse** classifier of $M_{spa} = N_{sb}$ 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 $N_{sb} = 64$, $M_{spa} = 64$ terms were selected using OFS based on FRCSM
Simulation Results

(a) User-one bit error rate performance comparison, (b) Influence of RBF variance $\rho^2$ 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 \textit{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 excellent test bit error rate performance