A Fast Learning Fully Complex-valued Relaxation Network (FCRN)

S. Suresh\textsuperscript{1} R. Savitha\textsuperscript{1} N. Sundararajan\textsuperscript{2}

\textsuperscript{1}School of Computer Engineering, Nanyang Technological University, Singapore
Email: ssundaram@ntu.edu.sg; savi0001@ntu.edu.sg
Website: http://www3.ntu.edu.sg/home/ssundaram/

\textsuperscript{2}Sri Jaya Chamarajendra College of Engineering (SJCE), Mysore, India.
Email: ensundara@ntu.edu.sg

Special Session: Complex-valued Neural Networks, International Joint Conference on Neural Networks, 2011.
Motivation: Complex-Valued Neural Networks (CVNN)

- Neural networks are capable of learning complex input-output relationships.
Neural networks are capable of learning complex input-output relationships.

Signals in applications like adaptive array signal processing, medical image processing, etc., are naturally complex-valued.
Neural networks are capable of learning complex input-output relationships.

Signals in applications like adaptive array signal processing, medical image processing, etc., are naturally complex-valued.

Representing in Complex domain preserves their physical characteristics.
Neural networks are capable of learning complex input-output relationships.

Signals in applications like adaptive array signal processing, medical image processing, etc., are naturally complex-valued.

Representing in Complex domain preserves their physical characteristics.

Hence, developing complex-valued neural networks capable of accurate magnitude and phase approximation is important.
Challenges

- Conflict between the desired properties of an activation function and Liouville’s theorem
  - **Desired properties of an activation function**: Has to be bounded and entire (differentiable at every point in the plane).
  - **Liouville’s theorem**: An entire and bounded complex-valued function is a constant.
Introduction

Complex-valued Neural Networks

Challenges

- Conflict between the desired properties of an activation function and Liouville’s theorem
  - **Desired properties of an activation function**: Has to be bounded and entire (differentiable at every point in the plane).
  - **Liouville’s theorem**: An entire and bounded complex-valued function is a constant.

Relaxed desired properties of a complex-valued activation function:

Multi-layer perceptron networks

- Split complex-valued multi-layer perceptron network (SC-MLP)
Multi-layer perceptron networks

- **Split complex-valued multi-layer perceptron network (SC-MLP)**
  

- **Fully complex-valued multilayer perceptron (FC-MLP)**
  
## Multi-layer perceptron networks

- **Split complex-valued multi-layer perceptron network (SC-MLP)**

- **Fully complex-valued multilayer perceptron (FC-MLP)**

- **Improved complex-valued multilayer perceptron network (IC-MLP)**
  - (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)
Multi-layer perceptron networks

- Split complex-valued multi-layer perceptron network (SC-MLP)  

- Fully complex-valued multilayer perceptron network (FC-MLP)  

- Improved complex-valued multilayer perceptron network (IC-MLP)  
  (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)
Multi-layer perceptron networks

- Improved complex-valued multilayer perceptron network (IC-MLP) (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

Radial basis function networks

- Complex-valued radial basis function network (CRBF) (Chen et. al., *Eurasip Signal Processing Journal*, vol. 35 (1), 1991.)
A Fully Complex-valued Relaxation Network

Introduction

Survey of Complex-valued neural network algorithms in literature

**Multi-layer perceptron networks**

- Improved complex-valued multilayer perceptron network (IC-MLP) (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

**Radial basis function networks**

- Complex-valued radial basis function network (CRBF) (Chen et. al., *Eurasip Signal Processing Journal*, vol. 35 (1), 1991.)
- Fully complex-valued radial basis function network (FC-RBF) (Savitha et. al., *International Journal of Neural Systems*, vol. 19 (4), 2009.)
### Multi-layer perceptron networks

- **Split complex-valued multi-layer perceptron network (SC-MLP)**
  

- **Fully complex-valued multilayer perceptron (FC-MLP)**
  

- **Improved complex-valued multilayer perceptron network (IC-MLP)**
  
  (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

### Radial basis function networks

- **Complex-valued radial basis function network (CRBF)**
  
  (Chen et. al., *Eurasip Signal Processing Journal*, vol. 35 (1), 1991.)

- **Fully complex-valued radial basis function network (FC-RBF)**
  
  (Savitha et. al., *International Journal of Neural Systems*, vol. 19 (4), 2009.)

- **Self-regulatory fully complex-valued radial basis function network (SR-FCRBF)**
  
  (Savitha et. al., *International Joint Conference on Neural Networks (IJCNN 2010)*, 2010.)
Sequential learning algorithms

- Complex-valued minimal resource allocation network (CMRAN)
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)
A Fully Complex-valued Relaxation Network

Introduction

Survey of Complex-valued neural network algorithms in literature

Sequential learning algorithms

- **Complex-valued minimal resource allocation network (CMRAN)**
  
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)

- **Complex-valued growing and pruning network (CGAP-RBF)**
  

Issues in existing complex-valued neural network algorithms

- **Convergence:** Slow convergence due to gradient descent based learning. (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

- **Error function:** Mean squared error function that is only an explicit representation of magnitude.

- **Activation functions** in SC-MLP and CRBF are not fully complex-valued: Inaccurate approximation of phase.
Survey of Complex-valued neural network algorithms in literature

Sequential learning algorithms

- Complex-valued minimal resource allocation network (CMRAN)
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)

- Complex-valued growing and pruning network (CGAP-RBF)

- Complex-valued self-regulatory resource allocation network (CSRAN)
  (Savitha et. al., *IEEE Transactions on Neural Networks*, vol. 22 (7), 2011.)

Issues in existing complex-valued neural network algorithms

- **Convergence:** Slow convergence due to gradient descent based learning.
  (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

- **Error function:** Mean squared error function that is only an explicit representation of magnitude.

- **Activation functions** in SC-MLP and CRBF are not fully complex-valued: Inaccurate approximation of phase.
Introduction

Survey of Complex-valued neural network algorithms in literature

Sequential learning algorithms

- Complex-valued minimal resource allocation network (CMRAN) (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)


- Complex-valued self-regulatory resource allocation network (CSRAN) (Savitha et. al., *IEEE Transactions on Neural Networks*, vol. 22 (7), 2011.)

Issues in existing complex-valued neural network algorithms

- **Convergence:** Slow convergence due to gradient descent based learning.
  (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

- **Error function:** Mean squared error function that is only an explicit representation of magnitude.

- **Activation functions** in SC-MLP and CRBF are not fully complex-valued: Inaccurate approximation of phase.
Introduction

Survey of Complex-valued neural network algorithms in literature

Sequential learning algorithms

- **Complex-valued minimal resource allocation network (CMRAN)**
  
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)

- **Complex-valued growing and pruning network (CGAP-RBF)**
  

- **Complex-valued self-regulatory resource allocation network (CSRAN)**
  
  (Savitha et. al., *IEEE Transactions on Neural Networks*, vol. 22 (7), 2011.)

Issues in existing complex-valued neural network algorithms

- **Convergence:**
  
  - Slow convergence due to gradient descent based learning.
  - Dependent on the initialization due to singularity in activation function.  
    (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)
**Sequential learning algorithms**

- Complex-valued minimal resource allocation network (CMRAN)  
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)

- Complex-valued growing and pruning network (CGAP-RBF)  

- Complex-valued self-regulatory resource allocation network (CSRAN)  
  (Savitha et. al., *IEEE Transactions on Neural Networks*, vol. 22 (7), 2011.)

**Issues in existing complex-valued neural network algorithms**

- **Convergence:**
  - Slow convergence due to gradient descent based learning.
  - Dependent on the initialization due to singularity in activation function.  
    (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

- **Error function:** Mean squared error function that is only an explicit representation of magnitude.
A Fully Complex-valued Relaxation Network

Introduction

Survey of Complex-valued neural network algorithms in literature

Sequential learning algorithms

- Complex-valued minimal resource allocation network (CMRAN)
  (Deng et. al., *IEEE Transactions on Neural Networks*, vol. 13 (3), 2002.)

- Complex-valued growing and pruning network (CGAP-RBF) (Li et. al.,

- Complex-valued self-regulatory resource allocation network (CSRAN) (Savitha et. al., *IEEE Transactions on Neural Networks*, vol. 22 (7), 2011.)

Issues in existing complex-valued neural network algorithms

- **Convergence:**
  - Slow convergence due to gradient descent based learning.
  - Dependent on the initialization due to singularity in activation function. (Savitha et. al., *Neurocomputing*, vol. 72 (16-18), 2009.)

- **Error function:** Mean squared error function that is only an explicit representation of magnitude.

- **Activation functions** in SC-MLP and CRBF are not fully complex-valued: Inaccurate approximation of phase
**Definition: Relaxation**

From a given initial condition, the system always returns to a minimum energy state (Yu and Tsai, *Pattern Recognition*, vol. 25 (2), 1992).
Principle of relaxation in FCRN

- **Initial conditions:**
  - Given training data set \( \{(z^1, y^1), \ldots, (z^t, y^t), \ldots, (z^N, y^N)\} \) with \( z^t \in \mathbb{C}^m \) is the \( m \)-dimensional input and \( y^t \in \mathbb{C}^n \) are the targets,
  - Number of hidden neurons \( K \),
  - Random constant complex-valued hidden neuron centers \( v_k \in \mathbb{C}^m; k = 1, \ldots, K \), and
  - Random constant complex-valued hidden neuron scaling factors \( u_k \in \mathbb{C}^m \).
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

**Principle of relaxation in FCRN**

- **Initial conditions:**
  - Given training data set \( \{(z_1^t, y_1^t), \ldots, (z_t^t, y_t^t), \ldots, (z_N^N, y_N^N)\} \), with \( z_t \in \mathbb{C}^m \) is the \( m \)-dimensional input and \( y_t \in \mathbb{C}^n \) are the targets,
  - Number of hidden neurons \( K \),
  - Random constant complex-valued hidden neuron centers \( v_k \in \mathbb{C}^m; \ k = 1, \ldots, K \), and
  - Random constant complex-valued hidden neuron scaling factors \( u_k \in \mathbb{C}^m \)

- **Reaching the minimum energy state:** Estimate the optimal output weights \( (W^* \in \mathbb{C}^{n \times K}) \) such that the total energy reaches its minimum.

\[
W^* := \arg \min_{W \in \mathbb{C}^{n \times K}} J(W) \tag{1}
\]

obtained using

\[
\frac{\partial J(W)}{\partial w_{lp}} = 0; \quad l = 1, \ldots, n; \quad p = 1, \ldots, K \tag{2}
\]
Key Features of Fully Complex-valued Relaxation Network (FCRN)

- Fast learning algorithm based on relaxation principles
Key Features of Fully Complex-valued Relaxation Network (FCRN)

- Fast learning algorithm based on relaxation principles
- Using a fully complex-valued activation function at the hidden/output layers
Key Features of Fully Complex-valued Relaxation Network (FCRN)

- Fast learning algorithm based on relaxation principles
- Using a fully complex-valued activation function at the hidden/output layers
- Using an energy function that is representative of both magnitude and phase of the error
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

Key Features of Fully Complex-valued Relaxation Network (FCRN)

- Fast learning algorithm based on relaxation principles
- Using a fully complex-valued activation function at the hidden/output layers
- Using an energy function that is representative of both magnitude and phase of the error
- Estimates output weights corresponding to the minimum energy point of the energy function
Key Features of Fully Complex-valued Relaxation Network (FCRN)

- Fast learning algorithm based on relaxation principles
- Using a fully complex-valued activation function at the hidden/output layers
- Using an energy function that is representative of both magnitude and phase of the error
- Estimates output weights corresponding to the minimum energy point of the energy function
- Estimated output weights are unique and optimum
Architecture of FCRN

\[ h'_t = \text{sech}(v_j^T(z^t - u_j)) \]

\[ z^t \]

\[ z_j^t \]

\[ z_m^t \]
Architecture of FCRN

FCRN Outputs:

- **Hidden layer response:**

  \[ h_j^t = \text{sech} \left( u_j^T (z^t - v_j) \right); \]

  \( j = 1, \ldots, K; \)

  \( u_j \in \mathbb{C}^m, v_j \in \mathbb{C}^m \)
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

FCRN Architecture

Architecture of FCRN

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_n^t$</td>
<td>Layer output</td>
</tr>
<tr>
<td>$w_{1l}$</td>
<td>Weight</td>
</tr>
<tr>
<td>$z_m$</td>
<td>Input</td>
</tr>
<tr>
<td>$v_{1l}$</td>
<td>Bias</td>
</tr>
<tr>
<td>$w_{nK}$</td>
<td>Weight</td>
</tr>
<tr>
<td>$h_{K-1}$</td>
<td>Hidden layer output</td>
</tr>
<tr>
<td>$h_K$</td>
<td>Hidden layer output</td>
</tr>
<tr>
<td>$h_1^t$</td>
<td>Hidden layer output</td>
</tr>
<tr>
<td>$h_2$</td>
<td>Hidden layer output</td>
</tr>
</tbody>
</table>

FCRN Outputs:

- **Hidden layer response:**

\[
  h_j^t = \text{sech} \left( u_j^T (z^t - v_j) \right);
\]

\[ j = 1, \ldots, K; \quad u_j \in \mathbb{C}^m, \ v_j \in \mathbb{C}^m \]

- **Network output:**

\[
  \hat{y}_l^t = \exp ( w_{lj} h_j^t ); \quad l = 1, \ldots, n
\]

\[
  w_l = [w_{l1}, \ldots w_{lj}, \ldots, w_{lK}]^T \in \mathbb{C}^n
\]
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

Nonlinear Energy Function

Energy Function:

\[ J_t = \sum_{l=1}^{n} (\ln(\hat{y}_t^l) - \ln(y_t^l)) (\ln(\hat{y}_t^l) - \ln(y_t^l)) \] (4)

In the polar co-ordinate system, \( \hat{y}_t^l = \hat{r}_t^l \exp(i\hat{\phi}_t^l) \); \( y_t^l = r_t^l \exp(i\phi_t^l) \)

where, \( \hat{r}_t^l \): Mag. of \( \hat{y}_t^l \); \( \hat{\phi}_t^l \): Phase of \( \hat{y}_t^l \); \( r_t^l \): Mag. of \( y_t^l \); \( \phi_t^l \): Phase of \( y_t^l \)
Energy Function:

\[ J_t = \sum_{l=1}^{n} (\ln(\hat{y}_t^l) - \ln(y_t^l)) (\ln(\hat{y}_t^l) - \ln(y_t^l)) \]  \hspace{1cm} (4)

In the polar co-ordinate system, \( \hat{y}_t^l = \hat{r}_t^l \exp(i\phi_t^l) \); \( y_t^l = r_t^l \exp(i\phi_t^l) \)

where, \( \hat{r}_t^l \): Mag. of \( \hat{y}_t^l \); \( \phi_t^l \): Phase of \( \hat{y}_t^l \); \( r_t^l \): Mag. of \( y_t^l \); \( \phi_t^l \): Phase of \( y_t^l \)

Using the polar co-ordinate definition of outputs/targets, energy function of each sample:

\[ J_t = \sum_{l=1}^{n} \left( \ln \left( \frac{\hat{r}_t^l}{r_t^l} \right)^2 + \left( \frac{\hat{r}_t^l}{r_t^l} \right)^2 \right) \]
For $N$ training samples, the overall energy is

\[ J(W) = \frac{1}{2} \sum_{t=1}^{N} J_t; \ N : \text{Number of training samples} \]

\[ = \frac{1}{2} \sum_{t=1}^{N} \sum_{l=1}^{n} \left( \ln \left( \frac{r_l^t}{\hat{r}_l^t} \right)^2 + (\hat{\phi}_l^t - \phi_l^t)^2 \right) \]  

(5)
A Fully Complex-valued Relaxation Network
Fully Complex-valued Relaxation Network (FCRN)
Nonlinear Energy Function

For $N$ training samples, the overall energy is

$$J(W) = \frac{1}{2} \sum_{t=1}^{N} J_t; \quad N : \text{Number of training samples}$$

$$= \frac{1}{2} \sum_{t=1}^{N} \sum_{l=1}^{n} \left( \ln \left( \frac{r_t}{\hat{r}_t^l} \right)^2 + (\hat{\phi}_t^l - \phi_t^l)^2 \right)$$  \hspace{1cm} (5)$$

Energy function in Eq. (5)

- Represents both the magnitude and phase of complex-valued error
- $J_t$ tends to 0, when $\hat{y}_t^l \rightarrow y_t^l$
- Is second order continuously differentiable with respect to the network parameters
Substituting for $\hat{y}^t_i$ from Eq. (3) in Eq. (4),

$$J(W) = \frac{1}{2} \sum_{t=1}^{N} \sum_{l=1}^{n} \left( \ln (y^t_i) - \sum_{j=1}^{K} \ln (\exp (w_{lj} h^t_j)) \right)$$

$$\left( \ln (\overline{y}^t_i) - \sum_{j=1}^{K} \ln (\exp (w_{lj} h^t_j)) \right)$$

(6)
A Fully Complex-valued Relaxation Network
Fully Complex-valued Relaxation Network (FCRN)
Projection based learning algorithm for FCRN

Substituting for $\hat{y}_i^t$ from Eq. (3) in Eq. (4),

$$J(W) = \frac{1}{2} \sum_{t=1}^{N} \sum_{l=1}^{n} \left( \ln(y_i^t) - \sum_{j=1}^{K} \ln \left( \exp \left( w_{ij} h_j^t \right) \right) \right) \left( \ln(y_i^t) - \sum_{j=1}^{K} \ln \left( \exp \left( w_{ij} h_j^t \right) \right) \right)$$

Simplifying Eq. (6), we have,

$$J(W) = \frac{1}{2} \sum_{t=1}^{N} \sum_{l=1}^{n} \left( \ln(y_i^t) - \sum_{j=1}^{K} (w_{ij} h_j^t) \right) \left( \ln(y_i^t) - \sum_{j=1}^{K} (w_{ij} h_j^t) \right)$$

Minimum energy point of $J(W)$ is obtained from $\frac{\partial J(W)}{\partial w_{ip}} = 0$
Using the Wirtinger calculus and the commutative property of the Complex conjugate operator,

\[
\frac{\partial J(W)}{\partial w_{lp}} = \sum_{t=1}^{N} h_p^t \left[ \sum_{k=1}^{K} \overline{w}_{lk} \overline{h}_k^t - \ln (\overline{y}^t_l) \right] \tag{7}
\]

The minimum energy of the energy function \( J(W) \) can be obtained by equating Eq. (7) to zero:

\[
\frac{\partial J(W)}{\partial w_{lp}} = 0 \implies \sum_{t=1}^{N} h_p^t \left[ \sum_{k=1}^{K} \overline{w}_{lk} \overline{h}_k^t - \ln (\overline{y}^t_l) \right] = 0 \tag{8}
\]

Rearranging Eq. (8), we have,

\[
\sum_{k=1}^{K} \overline{w}_{lk} \sum_{t=1}^{N} h_p^t \overline{h}_k^t = \sum_{t=1}^{N} \ln \left( \overline{y}^t_l \right) h_p^t
\]
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

Projection based learning algorithm for FCRN

\[ \sum_{k=1}^{K} w_{lk} \sum_{t=1}^{N} h_{p}^{t} \overline{h}_{k}^{t} = \sum_{t=1}^{N} \ln \left( \overline{y}_{i}^{t} \right) h_{p}^{t} \]

This can be written as:

\[ \sum_{k=1}^{K} w_{lk} A_{pk} = B_{lp}; \ p = 1, \cdots, K; \ l = 1, \cdots, n \] (9)

where, **Projection matrix** \( A_{pk} \in \mathbb{C}^{K \times K} = \sum_{t=1}^{N} h_{p}^{t} \overline{h}_{k}^{t} \); \( p, k = 1, \cdots, K \)

and, **Output matrix** \( B_{lp} \in \mathbb{C}^{n \times K} = \sum_{t=1}^{N} \ln \overline{y}_{i}^{t} h_{p}^{t} \); \( l = 1, \cdots, n \)
Re-writing Eq. (9)

\[ \sum_{k=1}^{K} \bar{w}_{lk} A_{pk} = B_{lp}; \ p = 1, \ldots, K; \ l = 1, \ldots, n \]

in matrix form,

\[ \bar{W}A = B \]

Applying the commutative law of multiplication of complex-valued conjugates, the closed form solution for \( W \) can be obtained using

\[ W^* = \bar{B} A^{-1} \]

if and only if \( A \) is invertible.
Proposition 1:
The responses of the neurons in the hidden layer are unique. i.e.
∀ \( z^t \), when \( k \neq p \), \( h^t_k \neq h^t_p \); \( k, p = 1, 2 \cdots K, t = 1, \cdots N \).
Proposition 1:

The responses of the neurons in the hidden layer are unique. i.e. \( \forall \mathbf{z}^t, \text{ when } k \neq p, h_k^t \neq h_p^t; k, p = 1, 2 \cdots K, t = 1, \cdots N. \)

Proof:

**Assumption:** For a given \( \mathbf{z}^t \), \( h_p^t = h_k^t; k \neq p \) \hspace{1cm} (10)

Assumption in Eq. (10) valid iff

\[
\mathbf{v}_p^T (\mathbf{z}^t - \mathbf{u}_p) = \mathbf{v}_k^T (\mathbf{z}^t - \mathbf{u}_k)
\]

\( u_{kj} \) and \( u_{pj} \), \( v_{kj} \) and \( v_{pj} \): uncorrelated random constants

\( \implies u_k \neq u_p \) and \( v_k \neq v_p \)

Hence, \( h_p^t \neq h_k^t \forall \mathbf{z}^t; t = 1, \cdots, N \)
**Proposition 2:**

The responses of the neurons in the hidden layer are non-zero. *i.e.*

\[ \forall \ z, \ h_k^t \neq 0; \ k = 1, 2 \cdots K. \]
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

Proof: $A$ is invertible

Proposition 2:
The responses of the neurons in the hidden layer are non-zero. i.e
$\forall \ z, h_k^t \neq 0; k = 1, 2 \cdots K.$

Proof:
Assumption: $h_k^t = 0$ (11)

$u_k^T(z^t - v_k) = \infty$
$z \rightarrow \infty, \text{ or } u_k \rightarrow \infty, \text{ or } v_k \rightarrow \infty$

$\|u_{kj}\| < 1; \|v_{kj}\| < 1; \|z_j\| < 1$

$\Rightarrow$ Assumption in Eq. (11) invalid $\forall z$

Hence, $\forall \ z, h_k^t \neq 0; k = 1, 2 \cdots K.$
A Fully Complex-valued Relaxation Network

Fully Complex-valued Relaxation Network (FCRN)

Proof: $\mathbf{A}$ is invertible

From propositions 1 and 2, the projection matrix $\mathbf{A}$:

- Has distinct rows and columns
From propositions 1 and 2, the projection matrix $\mathbf{A}$:

- Has distinct rows and columns
- Has non-zero elements
From propositions 1 and 2, the projection matrix $A$:

- Has distinct rows and columns
- Has non-zero elements
- Therefore, the projection matrix $A$ is invertible.
From propositions 1 and 2, the projection matrix $\mathbf{A}$:
- Has distinct rows and columns
- Has non-zero elements
- Therefore, the projection matrix $\mathbf{A}$ is invertible.
From propositions 1 and 2, the projection matrix $A$: 
- Has distinct rows and columns
- Has non-zero elements
- Therefore, the projection matrix $A$ is invertible.

Therefore, from any given initial condition, FCRN returns to the minimum energy point of $J(W)$, and the closed form solution for $W$ is given by:

$$W^* = \overline{B} \; \overline{A}^{-1}$$  \hspace{1cm} (12)
Given: Training data set: \((z^1, y^1), \ldots (z^t, y^t), \ldots, (z^N, y^N)\) and 
\((u_j, v_j), j = 1, \ldots, K\)

**START**

- Compute the hidden layer responses
  \[ h^t_j = sech(\mathbf{v}^T_k (z^t - u_k)); \ k = 1, \ldots, K \]
- Compute the projection matrix \(A \in \mathbb{C}^{K \times K}\)

\[
A_{pk} = \sum_{t=1}^{N} h^t_p \bar{h}^t_k; \ p, k = 1, \ldots, K
\]

- Compute the output matrix \(B \in \mathbb{C}^{K \times n}\)

\[
B_{lp} = \sum_{t=1}^{N} ln (\bar{y}^t_i) \ h^t_p; \ l = 1, \ldots, n; \ p = 1, \ldots, K
\]

- Estimate the optimum output weights: \(W^* = \mathbf{B} \mathbf{A}^{-1}\)

**END**
Problems used in the study

- Complex-valued function approximation problem (Savitha et. al., Neurocomputing, vol. 72 (16-18), 2009).

Performance measures used for comparison:

Root mean squared magnitude error ($J_{Me}$)

$$J_{Me} = \sqrt{\frac{1}{N \times n} \sum_{t=1}^{N} \sum_{l=1}^{n} (y_{t}^{l} - \hat{y}_{t}^{l})^2}.$$

Average absolute phase error ($\phi_{e}$)

$$\phi_{e} = \frac{1}{N \times n} \sum_{t=1}^{N} \sum_{l=1}^{n} |\phi_{t}^{l} - \hat{\phi}_{t}^{l}| \times \frac{180}{\pi}.$$
Problems used in the study

- Complex-valued function approximation problem (Savitha et. al., Neurocomputing, vol. 72 (16-18), 2009).

Performance measures used for comparison:

- Root mean squared magnitude error ($J_{Me}$)

$$J_{Me} = \sqrt{\frac{1}{N \times n} \sum_{t=1}^{N} \sum_{l=1}^{n} (y_t^l - \hat{y}_t^l) \cdot (y_t^l - \hat{y}_t^l)}$$

- Average absolute phase error ($\phi_e$)

$$\phi_e = \frac{1}{N \times n} \sum_{t=1}^{N} \sum_{l=1}^{n} |\phi_t^l - \hat{\phi}_t^l| \times \frac{180}{\pi}$$
Complex-valued function approximation problem

\[ f(\mathbf{z}) = \frac{1}{1.5} \left( z_3 + 10z_1 z_4 + \frac{z_2^2}{z_1} \right); \mathbf{z} \in \mathbb{C}^4; \quad 0 < |\mathbf{z}| < 2.5 \]

Table: Performance Comparison: Function Approximation Problem

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>No. of neurons</th>
<th>Training time (sec.)</th>
<th>Training error ( J_{Me} )</th>
<th>( \phi_e )</th>
<th>Testing error ( J_{Me} )</th>
<th>( \phi_e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC-MLP</td>
<td>15</td>
<td>1857</td>
<td>0.029</td>
<td>15.74</td>
<td>0.054</td>
<td>15.6</td>
</tr>
<tr>
<td>C-ELM</td>
<td>15</td>
<td>0.2</td>
<td>0.192</td>
<td>90</td>
<td>0.23</td>
<td>88.2</td>
</tr>
<tr>
<td>CMRAN</td>
<td>14</td>
<td>52</td>
<td>0.026</td>
<td>2.23</td>
<td>0.48</td>
<td>18.7</td>
</tr>
<tr>
<td>C-RBF</td>
<td>15</td>
<td>9686</td>
<td>0.15</td>
<td>51</td>
<td>0.18</td>
<td>52</td>
</tr>
<tr>
<td>FC-RBF</td>
<td>20</td>
<td>1910</td>
<td>0.02</td>
<td>15.9</td>
<td>0.05</td>
<td>15.8</td>
</tr>
<tr>
<td>FCRN</td>
<td>10</td>
<td>0.42</td>
<td>0.03</td>
<td>1.38</td>
<td>0.06</td>
<td>3.22</td>
</tr>
</tbody>
</table>
QAM channel equalization problem

\[ x_t = o_t + 0.1 o_t^2 + 0.05 o_t^3 + \nu_t \]

\[ o_t = (0.34 - 0.27i)s_t + (0.87 + 0.43i)s_{t-1} + (0.34 - 0.21i)s_{t-2} \]

**Table:** Performance Comparison: QAM Channel Equalization

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>No. of neurons</th>
<th>Training time (sec.)</th>
<th>Training error</th>
<th>Testing error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( J_{Me} )</td>
<td>( \phi_e )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( J_{Me} )</td>
<td>( \phi_e )</td>
</tr>
<tr>
<td>FC-MLP</td>
<td>15</td>
<td>3862</td>
<td>0.2</td>
<td>6.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.72</td>
<td>31.1</td>
</tr>
<tr>
<td>C-ELM</td>
<td>15</td>
<td>0.36</td>
<td>0.6</td>
<td>34.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
<td>35.11</td>
</tr>
<tr>
<td>CMRAN</td>
<td>40</td>
<td>555</td>
<td>0.4</td>
<td>17.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.43</td>
<td>21.17</td>
</tr>
<tr>
<td>C-RBF</td>
<td>15</td>
<td>8107</td>
<td>0.6</td>
<td>35.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
<td>39.86</td>
</tr>
<tr>
<td>FC-RBF</td>
<td>15</td>
<td>3840</td>
<td>0.4</td>
<td>31.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.41</td>
<td>14.69</td>
</tr>
<tr>
<td>FCRN</td>
<td>14</td>
<td>0.998</td>
<td>0.3</td>
<td>22.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.35</td>
<td>12.62</td>
</tr>
</tbody>
</table>
Symbol Error Rate (SER) Vs Signal-to-Noise Ratio (SNR) plot

![Graph showing SER vs SNR for different models including CRBF, C-ELM, FC-RBF, FC-MLP, and FCRN. The graph includes a Bayesian boundary and log_{10}(SER) on the y-axis and Signal-to-Noise Ratio (dB) on the x-axis. The graph illustrates the performance evaluation of FCRN.]
Conclusions

- **Accurate magnitude and phase approximation**: Use of the logarithmic energy function that represents magnitude and phase of error explicitly.
  - Complex-valued Function Approximation Problem: About 70% improvement in phase approximation
  - QAM Channel Equalization Problem: About 15% improvement in phase approximation
- **Compact network structure**: Requires fewer neurons compared to other complex-valued learning algorithms
- Requires **Minimum computational effort** due to the relaxation learning.

Directions for future work

- **FCRN Classifier**: Develop a complex-valued FCRN classifier.
- **Sequential Mode**: Train FCRN for large data sets using a sequential learning algorithm
- **Applications**: Use FCRN to solve practical applications
Conclusions

- **Accurate magnitude and phase approximation**: Use of the logarithmic energy function that represents magnitude and phase of error explicitly.
  - Complex-valued Function Approximation Problem: About 70% improvement in phase approximation
  - QAM Channel Equalization Problem: About 15% improvement in phase approximation
- **Compact network structure**: Requires fewer neurons compared to other complex-valued learning algorithms
- Requires **Minimum computational effort** due to the relaxation learning.

Directions for future work

- **FCRN Classifier**: Develop a complex-valued FCRN classifier.
- **Sequential Mode**: Train FCRN for large data sets using a sequential learning algorithm
- **Applications**: Use FCRN to solve practical applications