Mutistability Analysis of Neural Networks with Applications

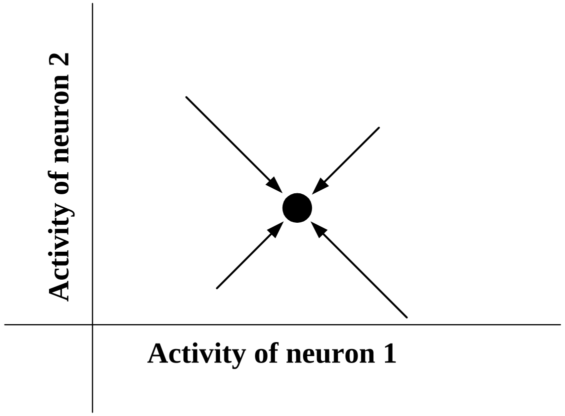
章毅
电子科技大学计算机科学与工程学院
计算智能实验室
http://cilab.uestc.edu.cn

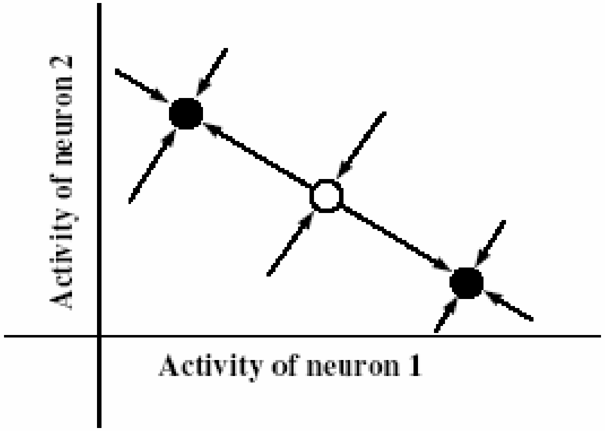
Multistability

- Multistability Analysis in Recurrent Neural Networks
- Multistability Analysis in Learning Algorithms
- Multistability Analysis with Applications

Concepts

- Monostability
  - A dynamic system has only one equilibrium point.
- Multistability
  - A dynamic system has more than one equilibrium points.
  - Stable and unstable equilibrium points can co-exist.

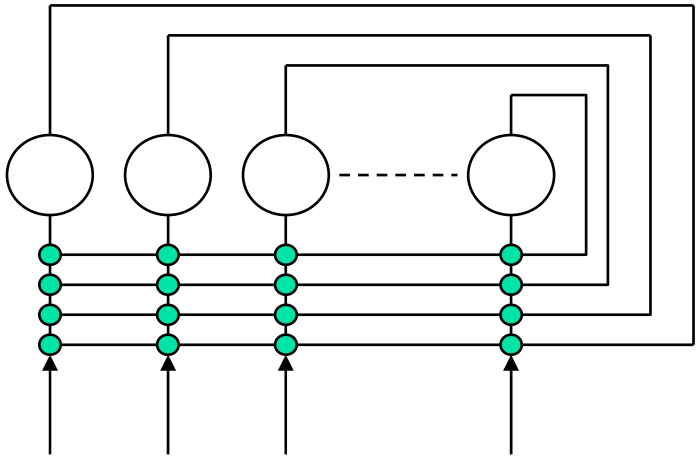




Recurrent Neural Networks

- Multistability is closely related to RNNs.
- Recurrent feedback loops pervade the synaptic connectivity of the brain.

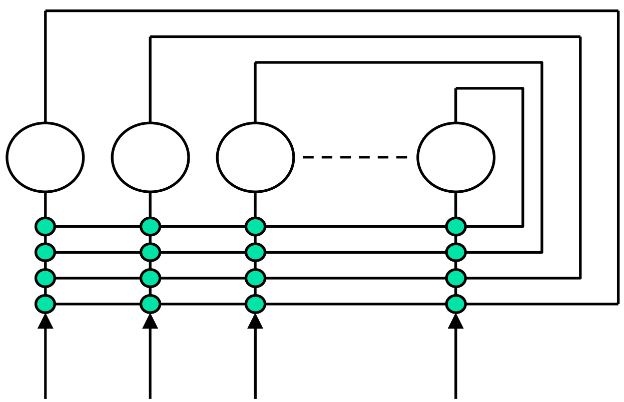
----Amit, 1995



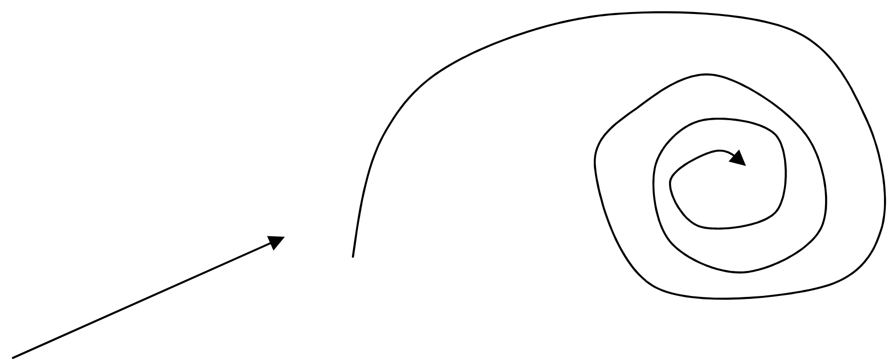
$$\frac{dx(t)}{dt} = -x(t) + f\left(wx(t) + b\right)$$

$$x(k+1) = f\left(wx(k) + b\right)$$

Recurrent Neural Networks



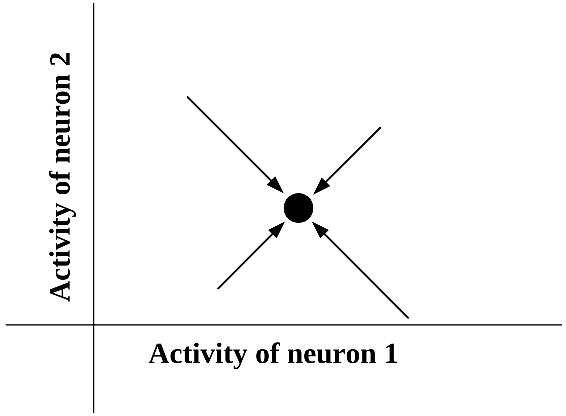
$$\frac{dx(t)}{dt} = -x(t) + f\left(wx(t) + b\right)$$



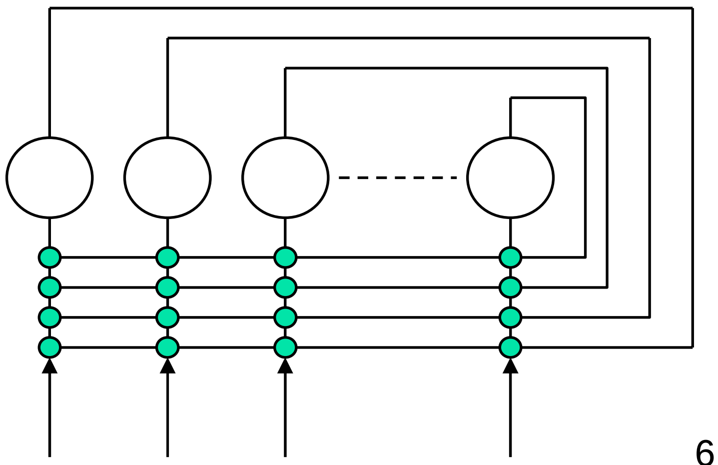
Network computing

On Monostability of RNNs

- A RNN has only one equilibrium point.
- Problem: whether or not the equilibrium point is a global attractor?
- Main method: Lyapunov second method
- Applications: optimizations

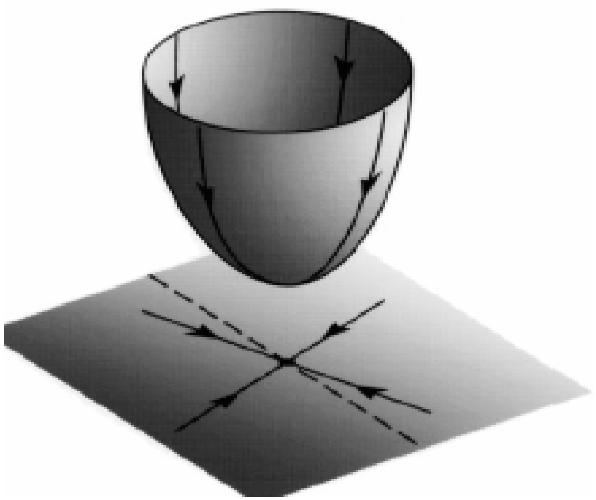


$$\frac{dx(t)}{dt} = -x(t) + f\left(wx(t) + b\right)$$



On Monostability of RNNs

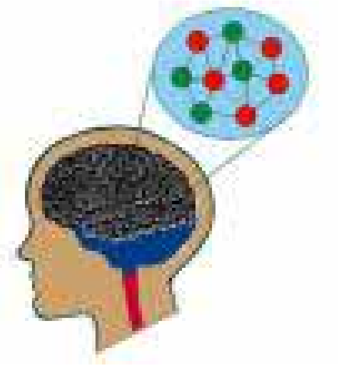
- There are a lot of publications.
- My view: most of these publications are not very interesting.
  - No more new methods.
  - Parallel generalization of existing method from mathematics.
  - Most results are not new from the point of mathematic.
  - Applications are restrictive.
  - Not strongly motivated by brain RNNs.



Seung 1996

Multistability in Recurrent NNs

- Recurrent feedback loops pervade the synaptic connectivity of the brain. One possible role of these feedback loops is to endow neural networks with multiple stable states, or dynamical attractors.



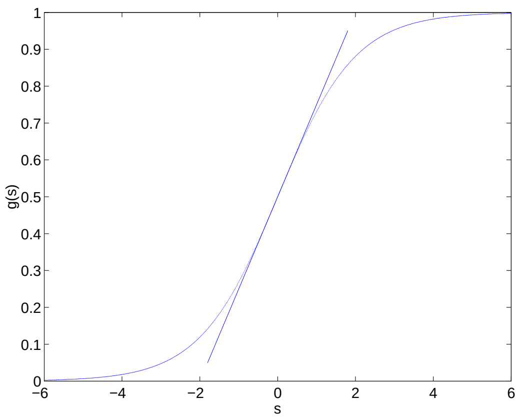
------------- Amit, 1995

Multistability Analysis Methods

- Problem: dynamical behaviors of a system with multiple equilibrium points.
  - Boundedness
  - Continuous/discrete attractors
  - Convergence of trajectories
  - ………..
- Methods:
  - Energy method
  - Invariant set principle
  - Cauchy convergence principle
  - ………..

Multistability in Hopfield NNs

$$\begin{cases} C_i \dfrac{du_i(t)}{dt} = -\dfrac{u_i(t)}{R_i} + \displaystyle\sum_{j=1}^{n} T_{ij} v_i(t) + I_i \\ v_i(t) = g_i(u_i(t)), (i = 1, \cdots, n) \end{cases}$$



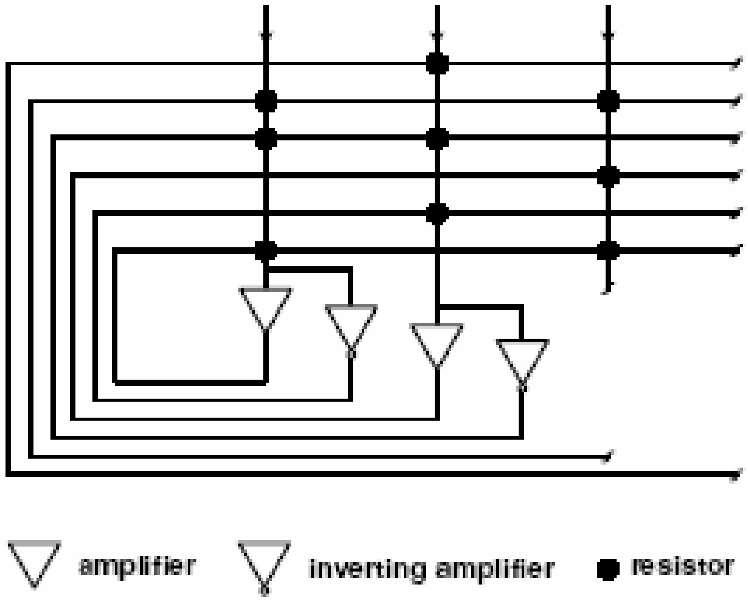
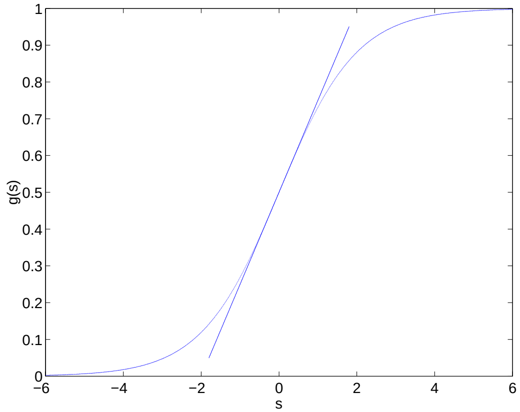


Figure 2.1. Electric circuit of Hopfield RNNs.

Multistability in Hopfield NNs

- Each trajectory converges to an equilibrium.
- Hopfield neural networks do not have any continuous attractors.
- Equilibrium points are isolated.

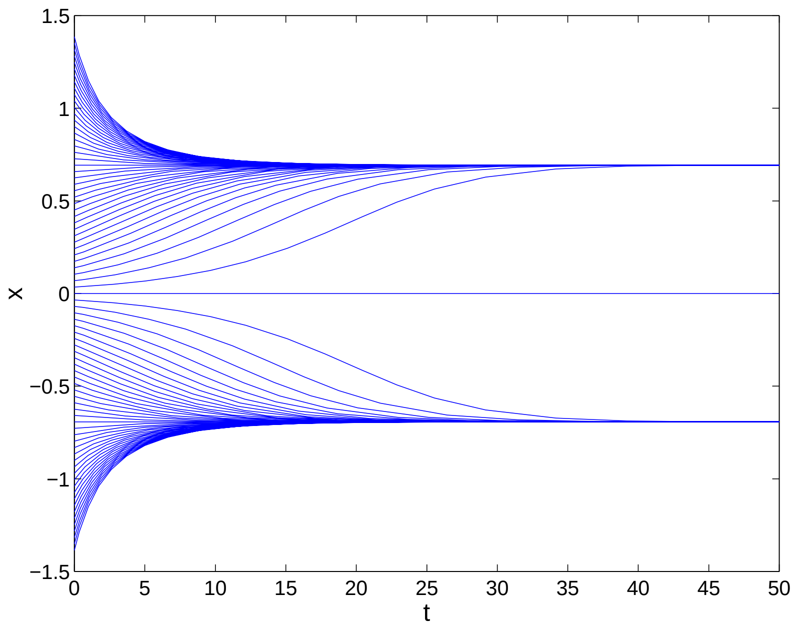
$$\begin{cases} C_i \dfrac{du_i(t)}{dt} = -\dfrac{u_i(t)}{R_i} + \displaystyle\sum_{j=1}^{n} T_{ij} v_i(t) + I_i \\ v_i(t) = g_i(u_i(t)), (i = 1, \cdots, n) \end{cases}$$



Multistability in Hopfield NNs

$$\dot{x}(t) = -\frac{3}{5\ln 2}x(t) + g(x(t)), \quad t \geq 0$$

$$g(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}}, \quad s \in R.$$



Multistability in Oculomotor Control

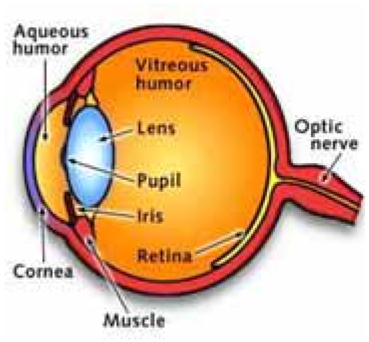
- H. S. Seung, How the brain keeps the eyes still, Proc. Natl. Acad. Sci. USA, vol. 93, pp. 13339-13344, 1996

$$v_i = v_i^0 + k_i E$$

v_i ----- the firing rate

v_i^0 ----- the firing rate at central gaze $E = 0$

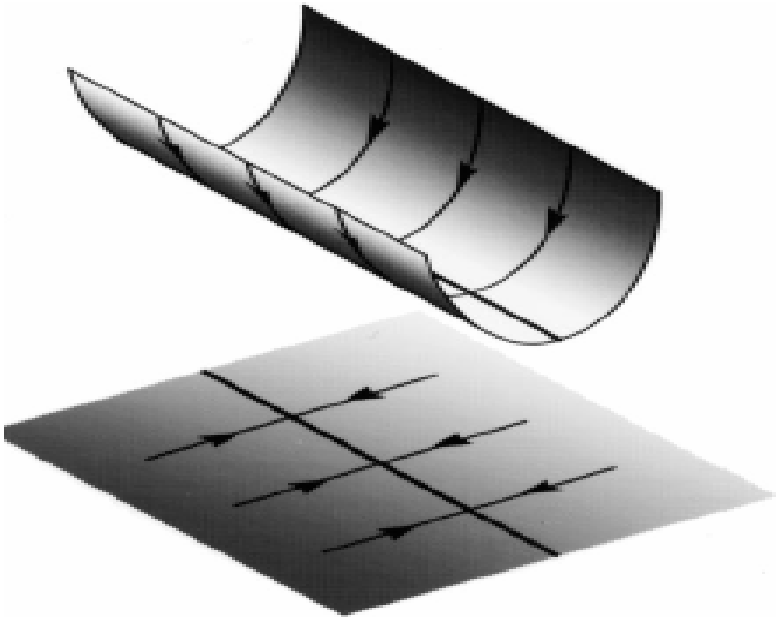
k_i ----- the position sensitivity



Line Attractor

$$x_i = k_i \left(E - E_i \right)$$

$$\tau \frac{dx_i\left(t\right)}{dt} + x_i\left(t\right) = \sum_{j=1}^{n} w_{ij} x_j + b_i$$

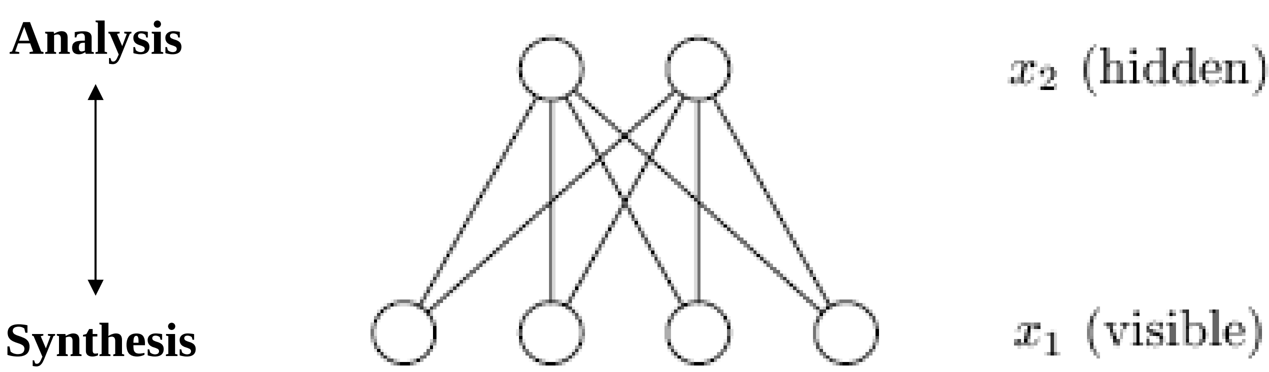


Seung 1996

Attractor RNNs

- H. S. Seung, Pattern analysis and synthesis in attractor neural networks

$$
\begin{aligned}
\dot{x}_1 + x_1 &= W_{12} x_2 \ , \\
\dot{x}_2 + x_2 &= W_{21} x_1 \ .
\end{aligned}
$$



Attractor RNNs

$$\dot{x_1} + x_1 = W_{12} x_2$$
$$\dot{x_2} + x_2 = W_{21} x_1$$

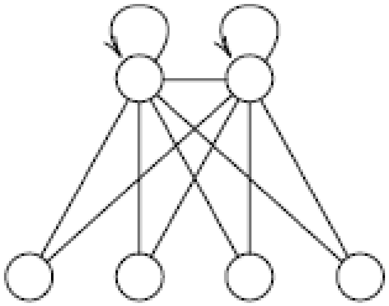
$$W_{21} W_{12} = I$$
$$W_{12} = W_{21}^T$$

$$\begin{aligned} E &= \frac{1}{2}|x_1|^2 + \frac{1}{2}|x_2|^2 - x_1^T W_{12} x_2 \\ &= \frac{1}{2}|x_1 - W_{12} x_2|^2 \end{aligned}$$

Memory The set of zero energy states

$$Z = \{(x_1, x_2) : x_1 = W_{12} x_2\}$$

Attractor RNNs



$$\begin{aligned} \dot{x}_1 + x_1 &= [W_{12} x_2]^+ , \\ \dot{x}_2 + x_2 &= [W_{21} x_1 + W_{22} x_2]^+ . \end{aligned}$$

$$\begin{aligned} W_{12} &= W_{21}^T , \\ W_{12} &\geq 0 , \\ W_{22} &= I - W_{21} W_{12} . \end{aligned}$$

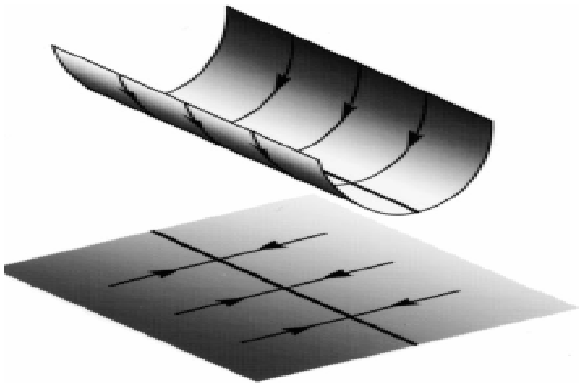
$$E = \frac{1}{2} |x_1 - W_{12} x_2|^2$$

Memory The n_2-dimensional linear manifold

$$Z^+ = \{(x_1, x_2) : x_2 \geq 0, x_1 = W_{12} x_2\}$$

Continuous Attractors

$$\tau \frac{dx_i(t)}{dt} + x_i(t) = f\left(\sum_{j=1}^{n} w_{ij} x_j + b_i \right)$$

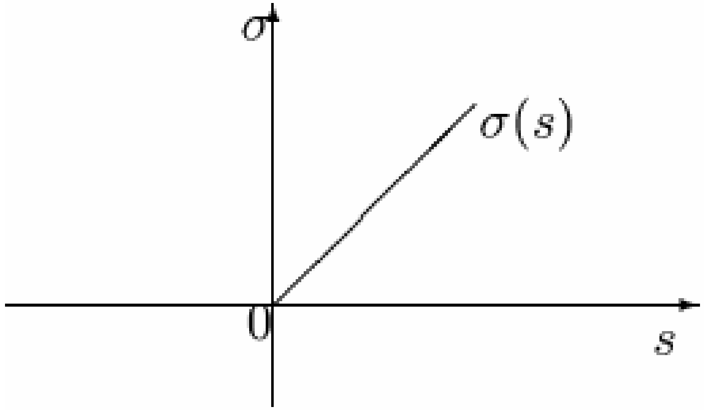


- Seung 1998 propose two important theoretical questions:
  - First, is it possible to implement continuous attractors in nonlinear networks?
  - Second, can nonlinearity make continuous attractors more robust?

Multistability in Nonlinear RNNs

- Zhang Yi, K. K. Tan and T. H. Lee, Multistability analysis for recurrent neural networks with unsaturating piecewise linear transfer functions, ***Neural Computation***, vol. 15, no. 3, pp. 639-662, 2003.
  - A global attractive compact set exist.
  - Conditions for calculating the attractive compact set are obtained.
  - Each trajectory converge – complete stable.

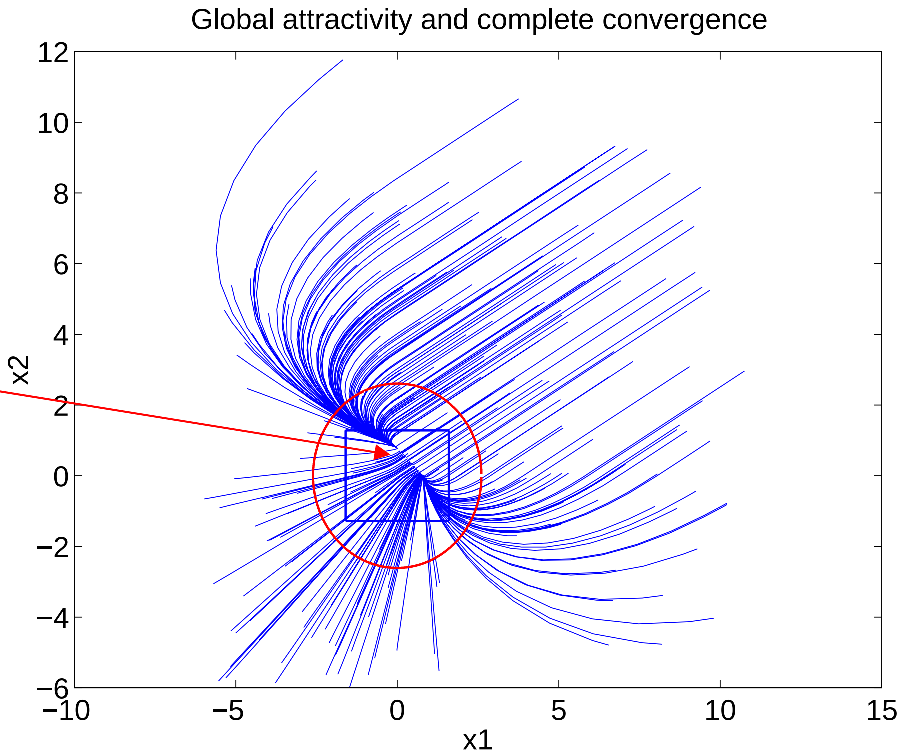
$$\dot{x}(t) = -x(t) + W\sigma(x(t) + h$$



Continuous Attractors

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = -\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -0.6 & 0.4 \end{bmatrix} \begin{bmatrix} \sigma(x_1(t)) \\ \sigma(x_2(t)) \end{bmatrix} + \begin{bmatrix} 0.8 \\ 0.48 \end{bmatrix}$$

Continuous Attractor

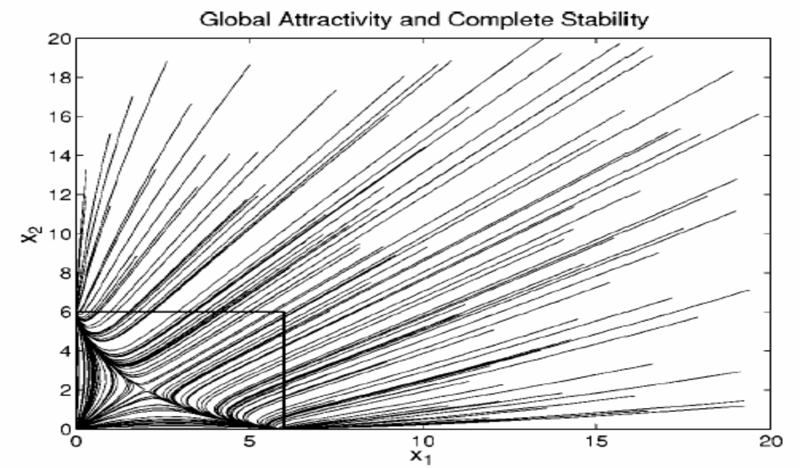


Multistability of Lotka-Volterra RNNs

- Derived from conventional membrane dynamics of competing neurons.
- Successful applications in many “winner-take-all” types of problems.

$$\dot{x}_i(t) = x_i(t)\left[h_i - x_i(t) + \sum_{j=1}^{n} \{a_{ij}x_j(t) + b_{ij}x_j[t-\tau_{ij}(t)]\}\right] \quad (i=1,\ldots,n)$$

Multistability of Lotka-Volterra RNNs

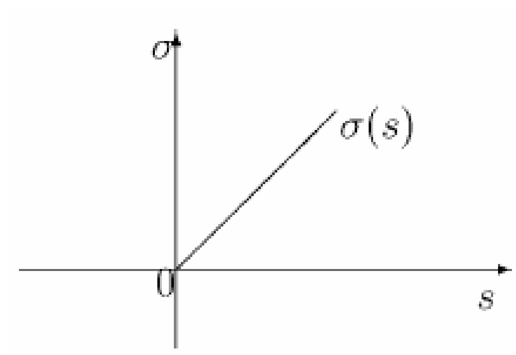


- A global attractive compact set exist.
- Conditions for calculating the attractive compact set are obtained.
- Each trajectory converge – complete stable.
- Z. Yi and K. K. Tan, Dynamic stability conditions for Lotka-Volterra recurrent neural networks with delays, ***Physical Review E***, 66, 011910 (2002).

Multistability of Discrete RNNs

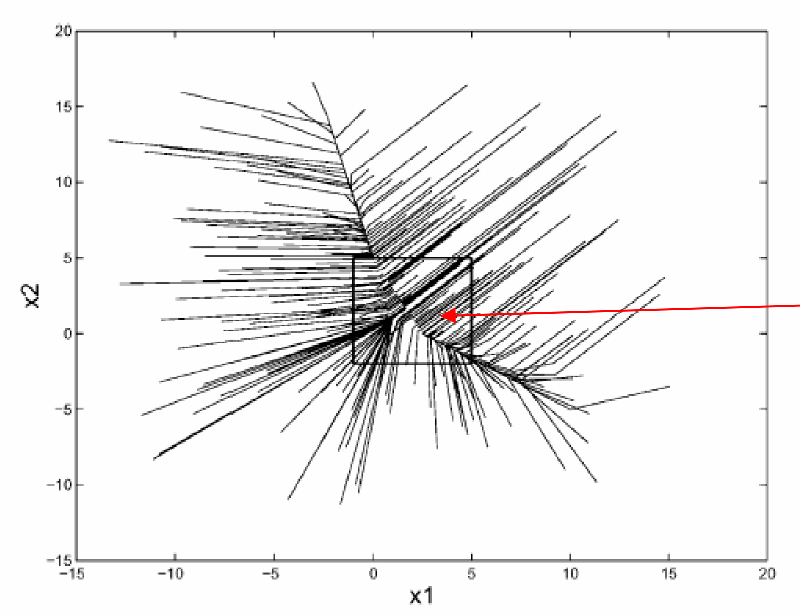
$$x_i(k+1) = \sum_{j=1}^{n} w_{ij}\sigma(x_j(k)) + h_i, \qquad (i = 1, \ldots, n)$$

$$x(k+1) = W\sigma(x(k) + h$$



Multistability of Discrete RNNs

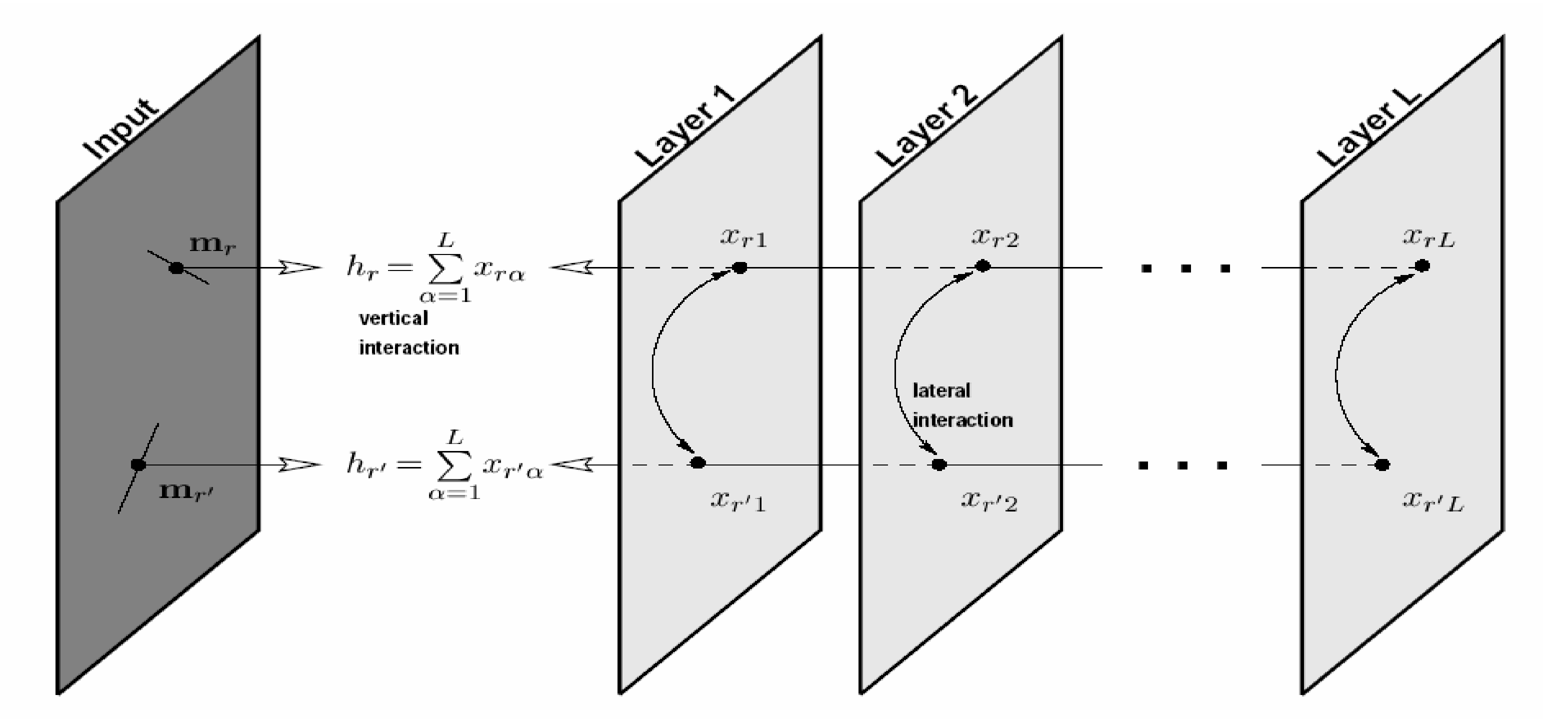
- A global attractive compact set exist.
- Conditions for calculating the attractive compact set are obtained.
- Each trajectory converge – complete stable.
- Z.Yi and K. K. Tan, Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions, ***IEEE Trans. Neural Networks***, vol. 15, no. 2, pp. 329-336, 2004.



Continuous Attractor

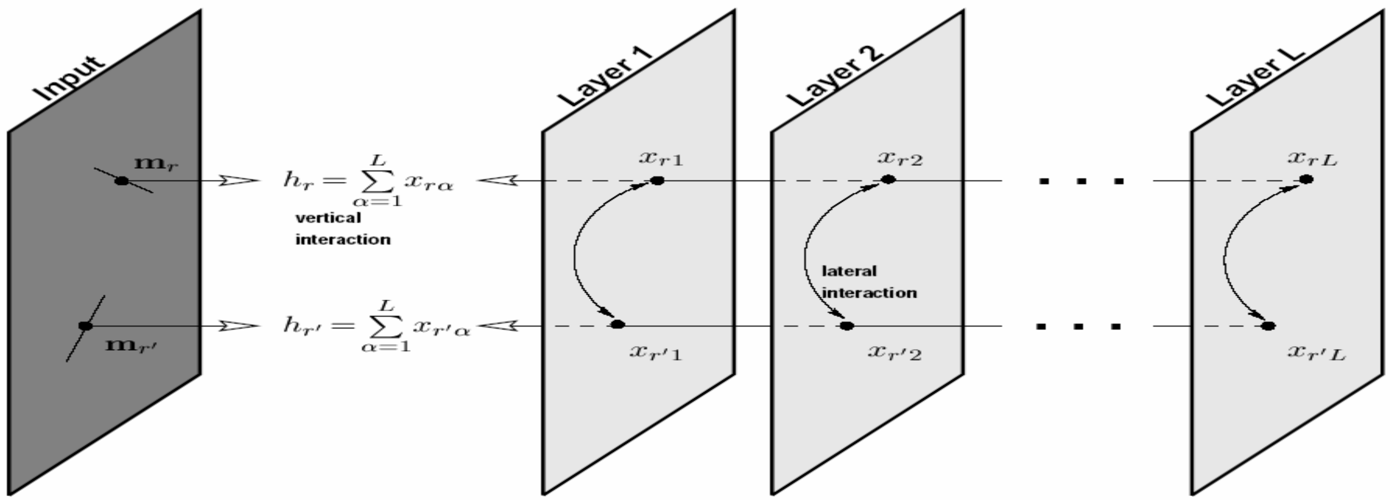
Multistability in CLM

Competitive Layer Model



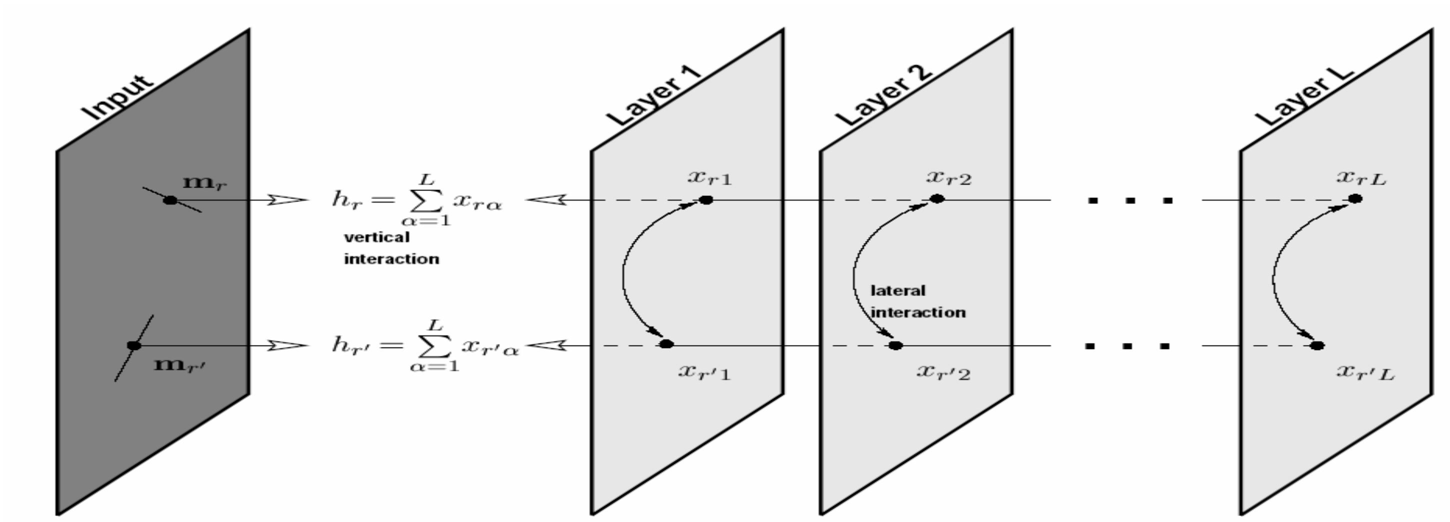
Competitive Layer Model

$$E = \frac{J_1}{2}\sum_{r}\left(\sum_{\alpha=1}^{L} x_{r\alpha} - h_r\right)^2 - \frac{1}{2}\sum_{\alpha=1}^{L}\sum_{rr'} f_{\mathbf{rr'}} x_{\mathbf{r}\alpha} x_{\mathbf{r'}\alpha}$$



Competitive Layer Model

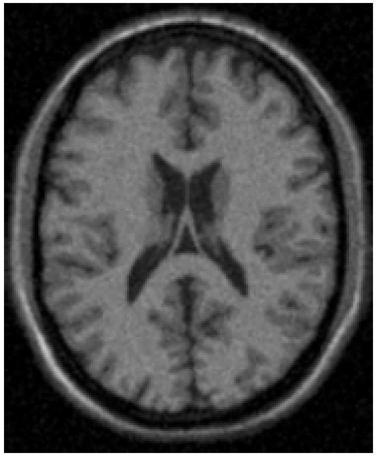
$$E = \frac{J_1}{2}\sum_{\mathbf{r}}\left(\sum_{\alpha=1}^{L} x_{\mathbf{r}\alpha} - h_{\mathbf{r}}\right)^2 - \frac{1}{2}\sum_{\alpha=1}^{L}\sum_{\mathbf{r}\mathbf{r}'} \mathrm{f}_{\mathbf{r}\mathbf{r}'} x_{\mathbf{r}\alpha} x_{\mathbf{r}'\alpha}$$

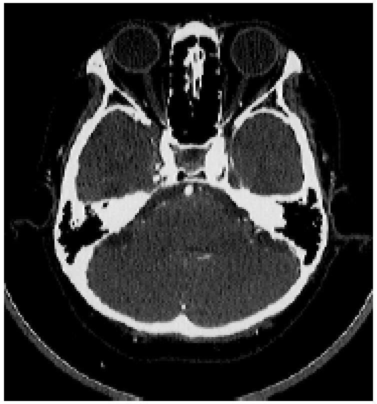


$$\begin{cases} x_p^l(k+1) = \alpha\left[\sum_{p'=1}^{P} a_{pp'}^l v_{p'}^l(k) - \sum_{l'=1}^{L} b_p^{ll'} v_p^{l'}(k) + h_p^l\right] \\ v_p^l(k) = \sigma\left(x_p^l(k)\right), \end{cases}$$

Medical Image Segmentation

- Medical Image Segmentation





Medical Image Segmentation

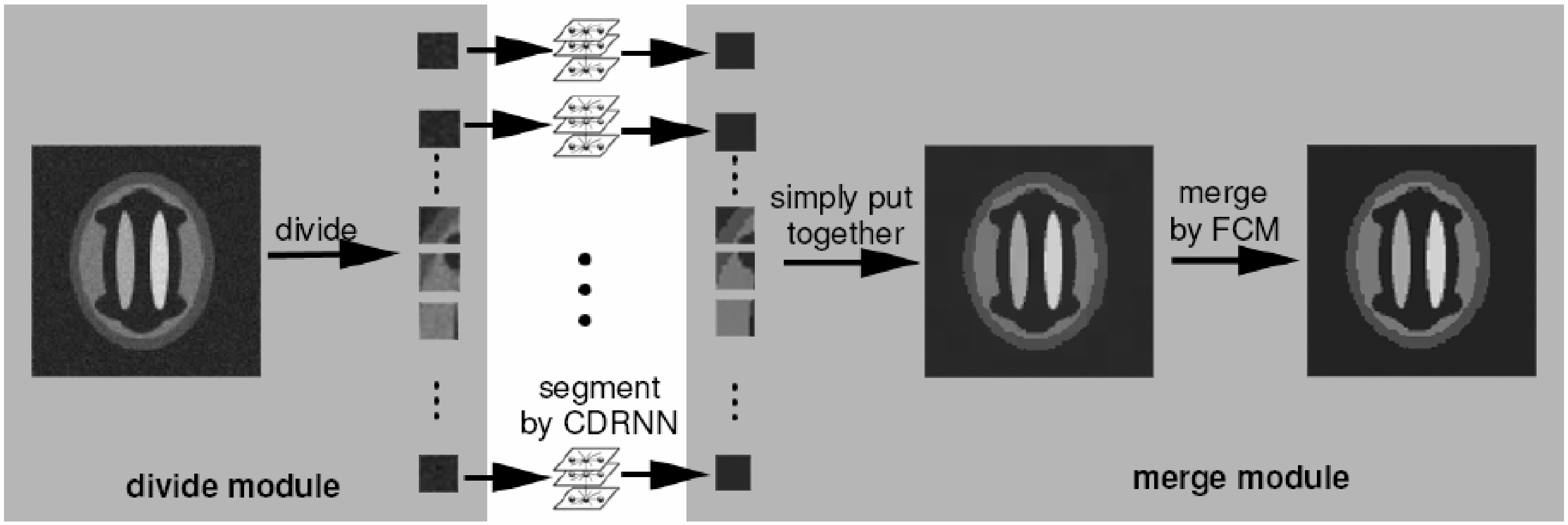
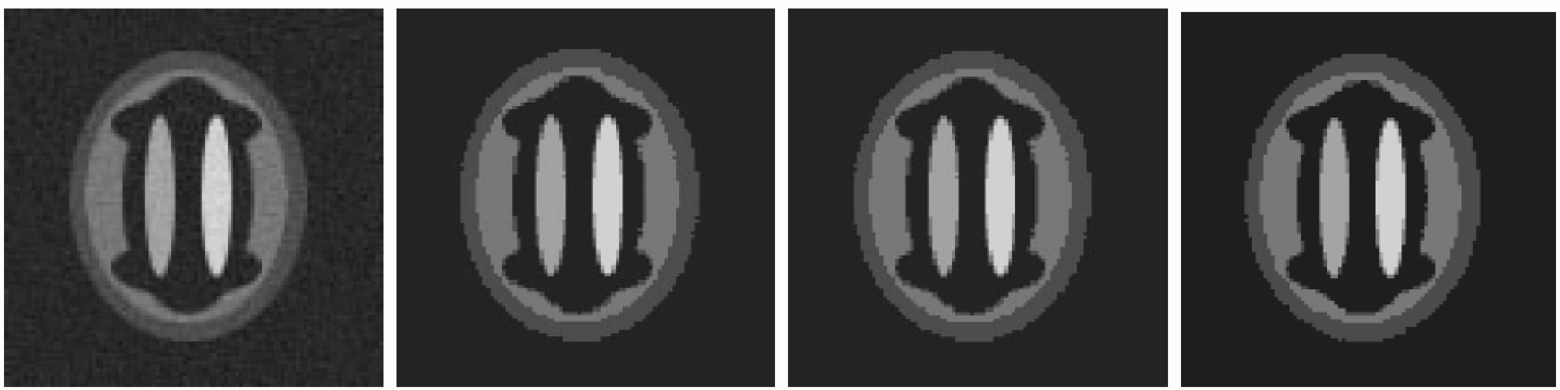


Fig. 3. Divide and merge system architecture.

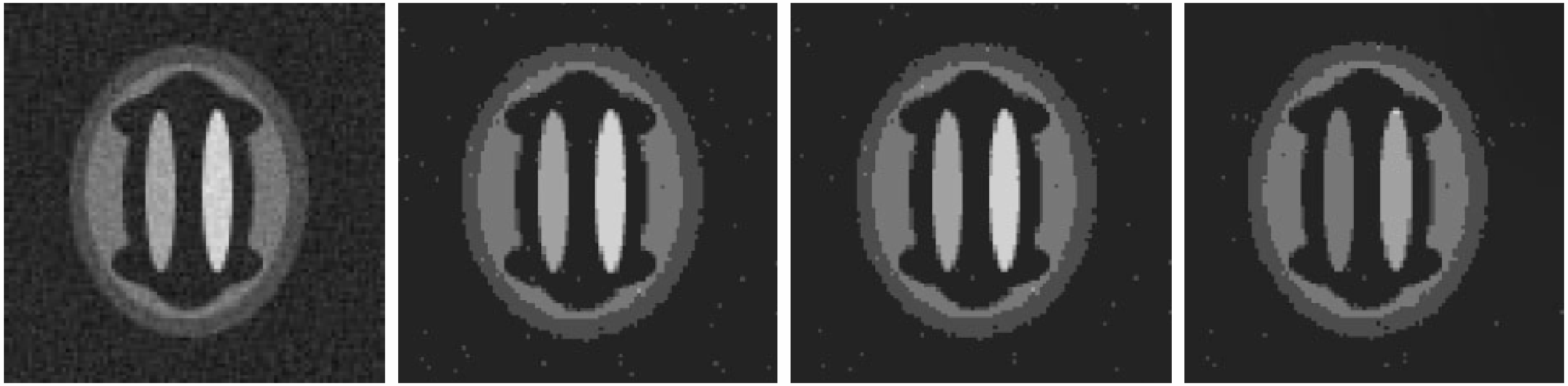
Medical Image Segmentation



(a) (b) (c) (d)

Fig. 7. (a) The original image with noise of 2%. (b), (c) and (d) are the segmentation results form FCM, CHNN and CDRNN, respectively. Due to the low noise level, all the three methods segmented the given image correctly.

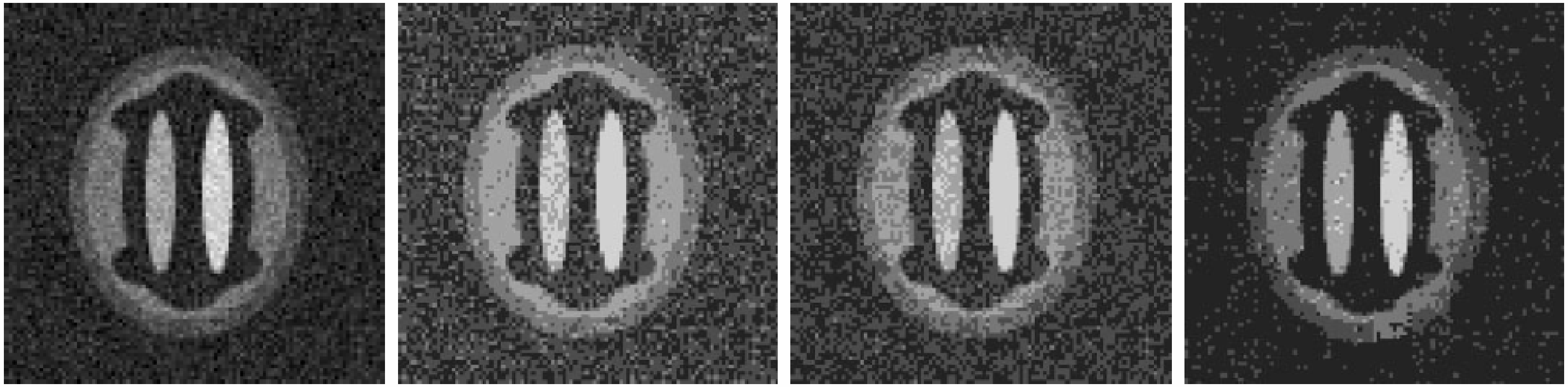
Medical Image Segmentation



(a) (b) (c) (d)

Fig. 8. (a) The original image with noise of 3%. (b), (c) and (d) are the segmentation results form FCM, CHNN and CDRNN, respectively. The three results are almost the same except for a little difference of the amount of isolated fragments.

Medical Image Segmentation



(a) (b) (c) (d)

Fig. 9. (a) The original image with heavy noise of 6%. (b), (c) and (d) are the segmentation results form FCM, CHNN and CDRNN, respectively. The proposed CDRNN method is obviously more accurate in image segmentation than the other two methods.

Medical Image Segmentation

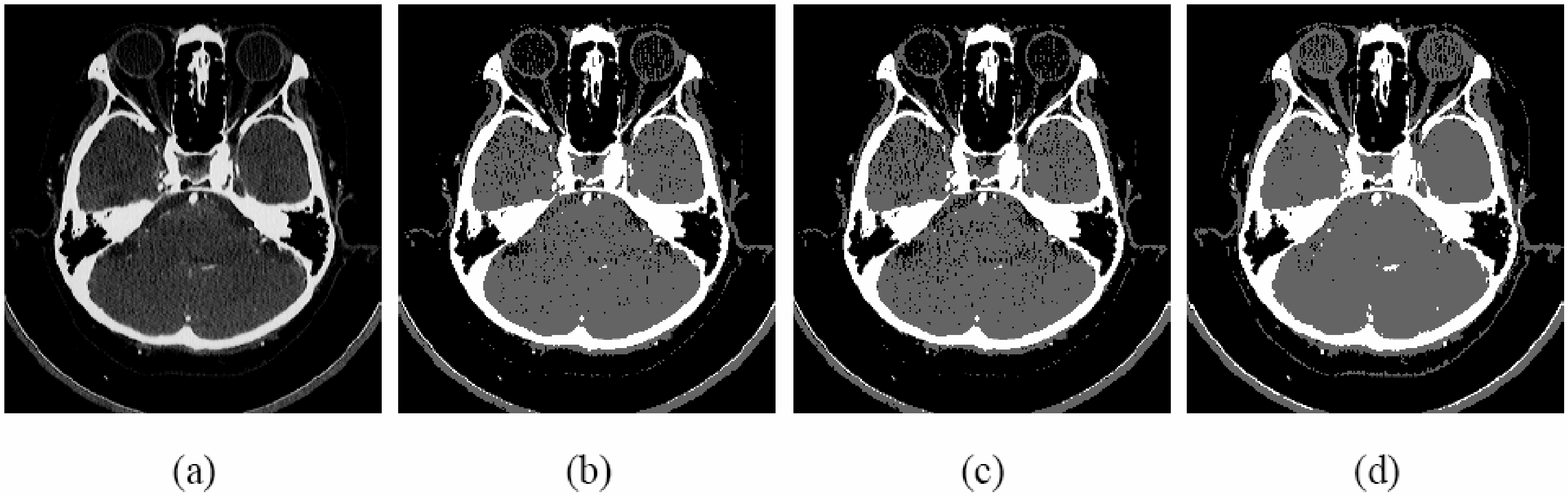
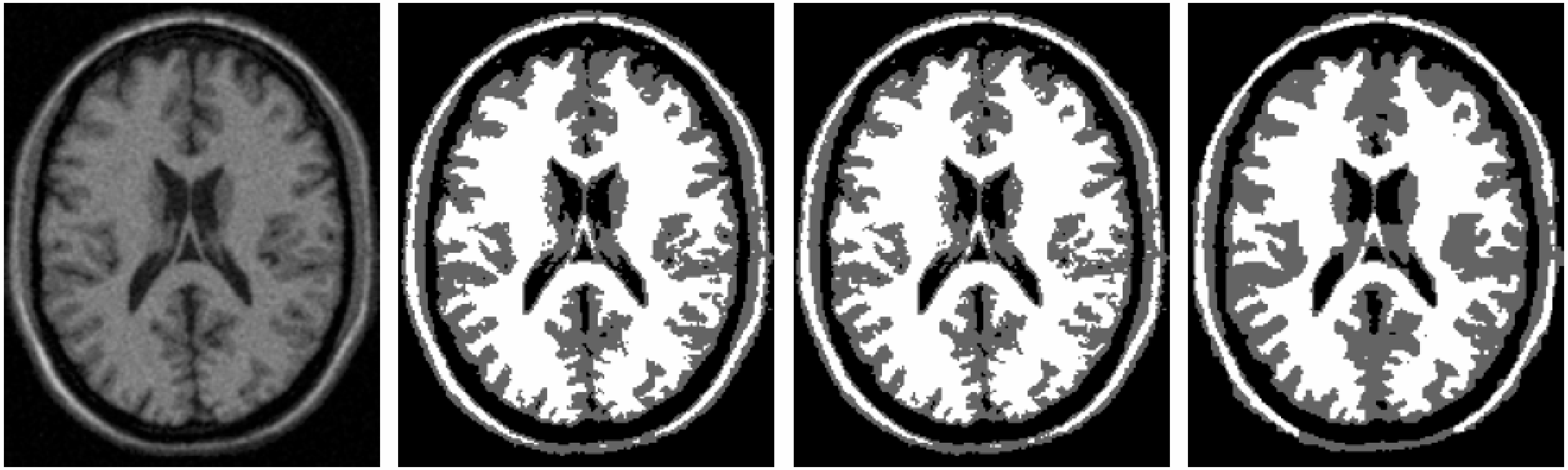


Fig. 11. (a) The original CT image. (b), (c) and (d) are the segmentation results by FCM, CHNN and CDRNN, respectively.

Medical Image Segmentation



(a) (b) (c) (d)

Fig. 12. (a) The original MRI brain image. (b), (c) and (d) are the segmentation results by FCM, CHNN and CDRNN, respectively.

Multistability in Learning Algorithms

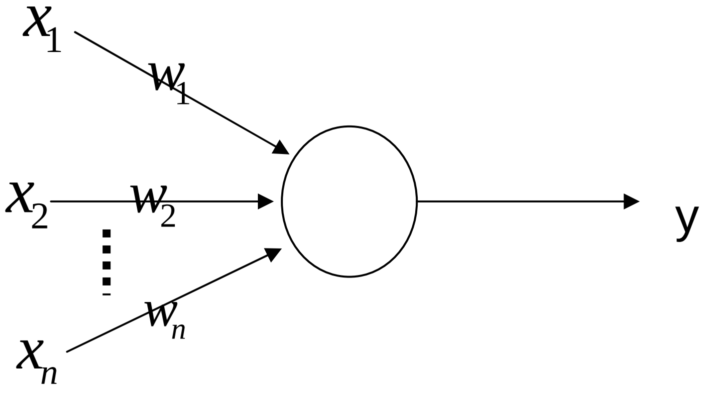
- A learning algorithm has many equilibrium points.
- Convergence study requires multistability analysis.
- Derive conditions for a learning algorithm to converge to a particular equilibrium point.

Oja's PCA Learning Algorithm

$$y(k) = w^T(k)x(k), \quad (k = 0, 1, 2, \cdots)$$

$$w(k+1) = w(k) + \eta y(k)\left[x(k) - y(k)w(k)\right]$$

- Convergence Analysis
  - DCT method
  - DDT method



Oja's PCA Learning Algorithm

Theorem 2: Given any constant l such that $1 \le l \le 2$, if

$$\eta\sigma \le \frac{3\sqrt[3]{2l} - 2}{2}$$

then, the set

$$S(l) = \left\{ w \,\middle|\, w \in R^n, w^T C w \le \lambda_p + \frac{l}{\eta} \right\}$$

is an invariant set of (2).

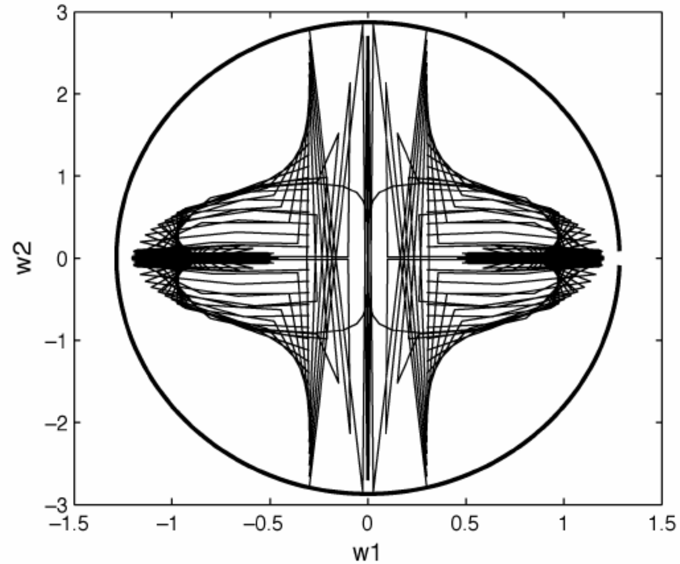


Fig. 1. Invariance of $S(2)$. 80 trajectories starting from points in $S(2)$ remain in $S(2)$.

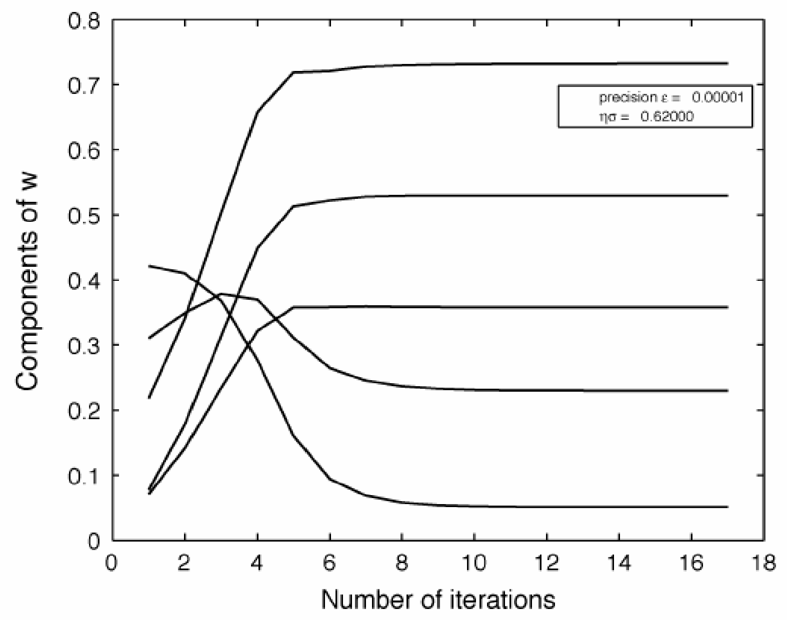
- Zhang Yi, etc., Convergence analysis of a deterministic discrete time system of Oja's PCA learning algorithm, ***IEEE Trans. Neural Networks***, vol. 16, no. 6, pp. 1318-1328.

Convergence of Oja's PCA Learning Algorithm

Theorem 3: Suppose that

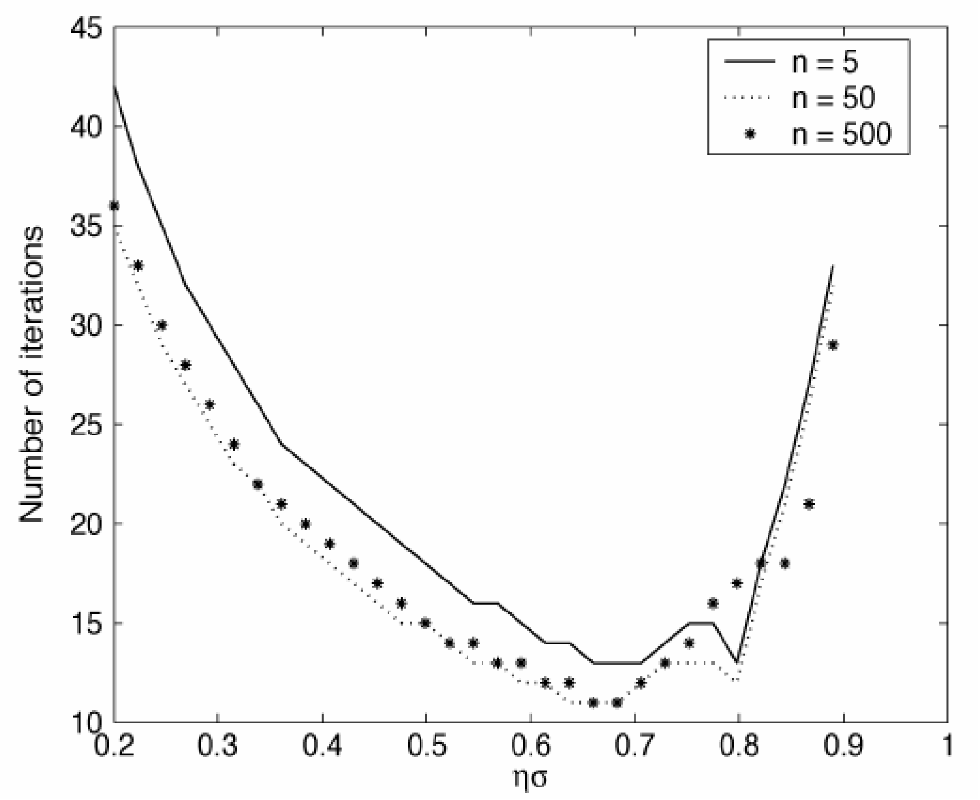
$$\eta\sigma \le \frac{3\sqrt[3]{2}-2}{2} \approx 0.8899$$

if $w(0)$ and $w(0) \notin V_\sigma^\perp$, then the trajectory of (2) starting from $w(0)$ will converge to a unit eigenvector associated with the largest eigenvalue of the correlation matrix C.



Convergence of Oja's PCA Learning Algorithm

$$\eta\sigma \approx 0.618.$$



Books on Multistability

- Zhang Yi and K. K. Tan, Convergence Analysis of Recurrent Neural Networks, Kluwer Academic Publisheres, Boston Hardbound, ISBN 1-4020-7694-0, 2004, 250pp.

- H. J. Tang, K. C. Tan and Zhang Yi, Neural Networks: Computational Models and Applications, Springer-Verlag, 2006.

Special Issue on Multistability

- Special issue: Multistability in Dynamical Systems
- *International Journal of Bifurcation and Chaos*
- Submission deadline: January 1, 2007.
- Guest Editors
  - Prof. Alexander N. Pisarchik
  - Prof. Celso Grebogi

Thanks !