Complex Valued Neural Network With Möbius Activation Function

Necati Özdemir * **Beyza Billur İskender** **
Nihal Yılmaz Özgür ***

* *Department of Mathematics, Faculty of Science and Arts, Balikesir*
University, Balikesir, TURKEY (e-mail:nozdemir@balikesir.edu.tr)
** *e-mail:biskender@balikesir.edu.tr*
*** *e-mail:nihal@balikesir.edu.tr*

Abstract: In this work, we propose a new type of activation function for a complex valued neural network (CVNN). This activation function is a special Möbius transformation classified as reflection. It is bounded outside of the unit disk and has partial continuous derivatives but not differentiable since it does not satisfy the Cauchy-Riemann equalities. However, the fixed points set of this function is a circle. Therefore, we employ this activation to a specific complex valued Hopfield neural network (CVHNN) and increase the number of fixed points. Using of this activation function leads us also to guarantee the existence of fixed points of the CVHNN. It is shown that the fixed points are all attractive which indicates that storage capacity of the CVHNN is enlarged.

Keywords: Complex valued neural networks, Hopfield neural networks, Möbius transformation, Reflection in a circle, Lyapunov stability.

1. INTRODUCTION

A complex valued neural network (CVNN) is a neural network that processes information in complex plane $\mathbb{C}$, Hirose (2006). It becomes very attractive field at the end of the 1980s and applicable to optoelectronics, imaging, remote sensing, quantum neural devices and systems, spatiotemporal analysis of physiological neural systems, and artificial neural information processing, see Gangal et al. (2007).

For CVNNs, the main task is to find a suitable activation function in a variety of complex functions. Despite the activation function of real valued neural networks (RVNNs) is chosen to be smooth and bounded generally as a sigmoid function, in complex plane these properties are not convenience for the nature of neural networks. Because of Liouville's theorem; the analytic and bounded functions on entire complex plane are constant.

There are several complex activation functions proposed in the literature. The basic ones of them are given below.

The sigmoid function was also used for CVNNs by Leung and Haykin (1991)

$$f(z) = \frac{1}{1 + e^{-z}},$$

but this function has singular points at every $z = (2n + 1)i\pi$, $n \in \mathbb{Z}$. They avoided this problem by scaling the input data to some region of complex plane. Later on, the sigmoid function was adapted to CVNNs as

$$f(z) = \frac{1}{1 + e^{-Rez}} + i\frac{1}{1 + e^{-Imz}}$$

★ Corresponding author

by Birx and Pipenberg (1992); Benvenuto and Piazza (1992). Also, *tanh* function which has singular points at every $z = (n + \frac{1}{2})i\pi$, $n \in \mathbb{Z}$ was adapted to CVNNs as real-imaginary type activation function

$$f(z) = \tanh(Rez) + i\tanh(Imz)$$

by Kechriotis and Monalakos (1994); Kinouchi and Hagiwara (1995) and as amplitude-phase type activation function

$$f(z) = \tanh(|z|)\exp(i\arg(z))$$

by Hirose (1994).

The other activation functions are given below:

$$f(z) = \frac{z}{|z|}$$

Noest (1988),

$$f(z) = \frac{z}{c + \frac{1}{r}|z|},$$

Georgiou and Koutsougeras (1992),

$$f(z) = \frac{Rez}{c + \frac{1}{r}|Rez|} + i\frac{Imz}{c + \frac{1}{r}|Imz|},$$

or

$$f(z) = \frac{|z|}{c + \frac{1}{r}|z|}\exp\left[i\left\{\arg z - \frac{1}{2^n}\sin(2^n\arg z)\right\}\right]$$

Kuroe and Taniguchi (2005), in which c and $\frac{1}{r}$ are positive constants and $-\pi < \arg z < \pi$. Detailed comparison for these types of activation functions can be found in Kuroe and Taniguchi (2009). In addition, Kim and Adali (2002) presented a set of elementary transcendental functions whose components are bounded almost everywhere and analytic functions to employ backpropagation. *tanh* function is one of them and the singularities of the function was avoided by restricting the domain of interest to a circle of radius $\frac{\pi}{2}$.

Another approach to chose activation functions of CVNNs using conformal mappings was proposed by Clarke (1990). He emphasized that the elegant theory of conformal mappings can be applied to find other activation functions. He gave the following function

$$f(z) = \frac{(\cos\theta + i\sin\theta)(z - \alpha)}{1 - \bar{\alpha}z},$$

where θ is a rotation angle, α is a complex constant with $|\alpha| < 1$ and $\bar{\alpha}$ denotes complex conjugate of α. This function is the general conformal mapping that transform unit disk in the complex plane onto itself and also a Möbius transformation. Furthermore, Möbius transformations were used in RVNNs by Mandic (2000). He showed that sigmoidal or *tanh* types of activation functions for a RVNN satisfy the conditions of a Möbius transformation. To base on the observation of "fixed points of a neural network are determined by fixed points of the employed activation function" he deduced "the existence for fixed points of the activation function are guaranteed by the Möbius transformation".

In a neural network, information is stored as asymptotically stable states, see Abu-Mostafa and St.Jacques (1985). In this work, we aim to increase the number of fixed points by considering this fact. Therefore, we choose the activation function as a reflection type Möbius transformation whose fixed points set is the unit circle.

The paper is organised as follows. In Section 2, the new type of activation function is presented. The basic properties of this function are given briefly. In Section 3, we apply the proposed activation function to a CVNN known as Hopfield Neural Network and analyse the number of fixed points. In Section 4, we investigate the stability characteristic of the fixed points using Lyapunov approach. In Section 5, we give some conclusions. Finally, we deal with the future works in Section 6.

2. MÖBIUS TRANSFORMATIONS AS ACTIVATION FUNCTION

A Möbius transformation is defined as

$$f(z) = \frac{az + b}{cz + d} \tag{1}$$

where $a, b, c, d \in \mathbb{C}$ and $ad - bc = 1$. It is a conformal mapping of the complex plane and also known as linear fractional or bilinear transformation. Such a Möbius transformation has at most two fixed points if it is not identity transformation $f(z) = z$. Detailed information could be found in Jones and Singerman (1994); Needham (2000).

Möbius transformations with real coefficients can be classified into

$$G_1 = \left\{ f : f(z) = \frac{az + b}{cz + d}, \ a, b, c, d \in \mathbb{R}, \ ad - bc = 1 \right\} \tag{2}$$

and

$$G_2 = \left\{ g : g(z) = \frac{a\bar{z} + b}{c\bar{z} + d}, \ a, b, c, d \in \mathbb{R}, \ ad - bc = -1 \right\}. \tag{3}$$

The transformations belong to the components of $G = G_1 \cup G_2$ are bijective transformations of extended complex plane. The transformations in G_1 are conformal mappings

and have at most two fixed points. Any transformation belongs to G_2 is anti-conformal mapping and can be classified according to the value of $a + d$. If $a + d \neq 0$, the transformation is called as a glide-reflection and has two fixed points on the real axis. If $a + d = 0$, the transformation is called as a reflection and has infinite number of fixed points on a circle centered at $\frac{a}{c}$ and of radius $\frac{1}{|c|}$. To utilize the infinite number of fixed points, we use reflection type Möbius transformation as an activation function. We begin to analyse these types of activation functions by choosing a simple reflection transformation

$$f(z) = \frac{1}{\bar{z}}. \tag{4}$$

This transformation maps unit circle onto itself, outside of the unit circle to its inside and inside of the unit circle to its outside. It is not differentiable since it does not satisfy the Cauchy-Riemann equalities and has a singularity at $z = 0$. This transformation is bounded only for the points at the outside of the unit circle, see Figure 1. Therefore, we consider the signals from the set of

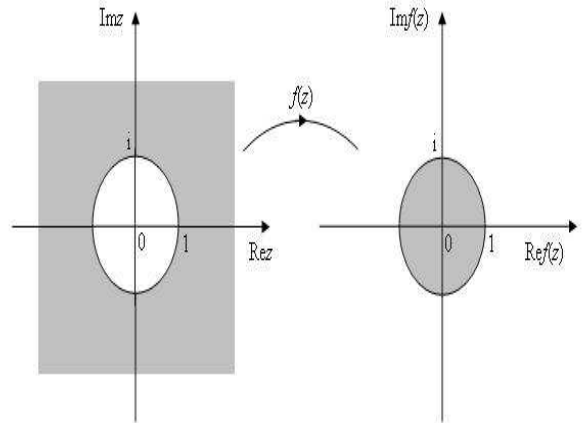$$B = \{ z : |z| \geq 1 \}. \tag{5}$$

Fig. 1. Geometric interpretation of the reflection transformation given by Eq.(4)

Remark: Let γ be the circle centered at p and of radius r. Then it is known that the reflection transformation in the circle γ is denoted by $I_\gamma(z)$ and defined as follows:

$$I_\gamma(z) = \frac{r^2}{\bar{z} - \bar{p}} + p. \tag{6}$$

This indicates that any circle in complex plane can be represented by a unique Möbius transformation, see Brickman (1993). When the circle is centered on real axis, the transformation is to be a reflection. Indeed, we have

$$I_\gamma(z) = \frac{p\bar{z} + r^2 - |p|^2}{\bar{z} - \bar{p}}$$

Now, we divide the numerator and the denominator of this transformation with r, then we have $ad - bc = -1$ and $a + d = \frac{p}{r} - \frac{\bar{p}}{r} = 0$ since $p \in \mathbb{R}$. It means that we have the advantage of to determine the fixed point circle

in the complex plane. In this paper, we analyse the CVNN whose fixed points set is chosen as the circle centered at the origin with radius $r = 1$. The analysis is also valid for the circles that is centered at the origin and with any values of radius. From Eq.(6) the reflection transformation of a circle centered at the origin and of radius r is

$$f(z) = \frac{r^2}{\bar{z}}. \qquad (7)$$

This transformation is bounded in the following set

$$B = \{z : |z| \geq r\}$$

and maps outside of the circle with radius r to inside of the circle with radius r. Thus, we can deduce that the domain of the input signals and the fixed points circle can be adjusted.

In the following section, we give a CVNN model to analyse the advantages of the new type activation function.

3. COMPLEX VALUED HOPFIELD NEURAL NETWORK

Hopfield neural network can be considered as a class of nonlinear and autonomous system, see Li et al. (1985); Leblebicioğlu et al. (2001). We consider this class of system in complex plane in order to interest complex valued Hopfield neural network (CVHNN) given by

$$\dot{z}(t) = -H(z(t))(-Tz(t) + F(z(t)) - U),$$

where $T \in \mathbb{C}^{n \times n}$, $U \in \mathbb{C}^n$ are matrices, $z(t) \in \mathbb{C}^n$ is state vector, $F(z) = (f_1(z_1), f_2(z_2), \cdots, f_n(z_n))^T : \mathbb{C}^n \to \mathbb{C}^n$ is an activation function and $H(z) : \mathbb{C}^n \to \mathbb{C}^{n \times n}$ is a nonlinear function.

The activation function is chosen as in Eq.(4):

$$f_j(z_j) = \frac{1}{\bar{z}_j}, \quad j = 1, 2, ..., n. \qquad (8)$$

To obtain correspondence between fixed points of the activation function and fixed points of the network, we select $T \in \mathbb{R}^{n \times n}$ and $U = 0$. Hence, we interest the CVHNN with the form of

$$\dot{z}(t) = -H(z(t))(-Tz(t) + F(z(t))). \qquad (9)$$

Fixed points of the Eq.(9) are calculated by the following equation:

$$-H(z)(-Tz + F(z)) = 0.$$

Assume that $H(z)$ is a nonsingular matrix then the fixed points are

$$F(z) = Tz$$

which correspond to the fixed points of the activation function.

4. STABILITY OF FIXED POINTS

As mentioned above, information are stored in a neural network as asymptotically stable states. A stable state is a fixed point of the neural network and also known as equilibrium point. Therefore, it is important to increase the number of stable states in a neural network. By using the activation function in Eq.(8) we increase the number of fixed points, but now we must know whether the fixed points are stable or not. We investigate stability of the fixed points by using Lyapunov stability.

Definition 1. $E(z)$ is a Lyapunov function of the CVHNN if $E(z)$ is a mapping $E : \mathbb{C}^n \to \mathbb{R}$ and the derivative of E along the trajectory of CVHNN satisfies $\dot{E}(z) \leq 0$. Furthermore, $\dot{E}(z) = 0$ if and only if $\dot{z} = 0$.

If all equilibrium points of the network are isolated and CVHNN given by Eq.(9) has a Lyapunov function, then no nontrivial periodic solution exists and each solution of the network converges to an equilibrium point as $t \to \infty$, see Kuroe et al. (2003).

An equilibrium point is isolated if it has no other equilibrium points in its vicinity, or there could be a continuum (compact and connected set) of equilibrium points, Khalil (1996). The fixed points of the CVHNN are isolated since they are on a circle. Therefore, the following theorem gives the stability of the fixed points. Here, we use the inner product defined on $\mathbb{C}^n$ as

$$\langle z_1, z_2 \rangle = z_2^* z_1,$$

where $z_1, z_2 \in \mathbb{C}^n$ and $(\cdot)^*$ denotes the conjugate transpose.

Theorem 2. If the matrix $T \in \mathbb{R}^{n \times n}$ is symmetric and the matrix $\text{Re}[H(z)]$ is positive definite, then the function

$$E(z) = -\frac{1}{2}z^* T z + Re\left[\sum_{j=1}^{n} \int_0^{\bar{z}_j} \bar{f}_j(s)\, ds\right] \qquad (10)$$

is a Lyapunov function of the CVHNN given by Eq.(9).

Proof. We can write Eq.(10) in the component wise form as

$$E(z) = -\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \bar{z}_j T_{jk} z_k + \text{Re}\left[\sum_{j=1}^{n} \int_0^{\bar{z}_j} \bar{f}_j(s)\, ds\right].$$

To show the monotonic decreasing of E with time t we compute $\dot{E}(z)$. Differentiating the first term of E gives

$$-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{d\bar{z}_j}{dt}T_{jk}z_k + \frac{dz_j}{dt}T_{jk}\bar{z}_k + \frac{d\bar{z}_k}{dt}T_{jk}z_j + \frac{dz_k}{dt}T_{jk}\bar{z}_j\right).$$

By using the symmetry property of the T matrix, this term can be arranged as follow

$$-\text{Re}\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(T_{jk}z_k\frac{d\bar{z}_j}{dt}\right)\right].$$

Therefore,

$$\dot{E}(z) = -\text{Re}\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(T_{jk}z_k\frac{d\bar{z}_j}{dt}\right)\right] + \text{Re}\left[\sum_{j=1}^{n}\bar{f}_j(\bar{z}_j)\frac{d\bar{z}_j}{dt}\right].$$

Using the property of $\bar{f}_j(\bar{z}_j) = f_j(z_j)$, this equation can be written in the matrix form as

$$\dot{E}(z) = -\text{Re}\left[(Tz - F(z))\dot{z}^*\right].$$

Substituting $\dot{z}^*$ into the above equation gives

$$\dot{E}(z) = -\text{Re}\left[(Tz - F(z))(Tz - F(z))^* H(z)^*\right]$$

$$= -\text{Re}\left[(Tz - F(z))\right]^2 \text{Re}\left[H(z)^*\right]$$

which is negative for positive definite $\text{Re}[H(z)]$ matrix and also equal to zero if and only if $\dot{z}(t) = 0$ ∎

Theorem 2 shows that the proposed activation function leads to infinite number of stable equilibrium points. Consequently, number of the stored information is increased.

5. CONCLUSIONS

This paper is constructed on the idea of interesting geometric properties of Möbius transformations which are conformal mappings of the complex plane. Because of a special class of Möbius transformation defined in Eq.(3)

$$g(z) = \frac{a\bar{z} + b}{c\bar{z} + d}, \ a, b, c, d \in \mathbb{R}, \ ad - bc = -1$$

has infinite number of fixed points if $a + d = 0$, we think to combine this property with complex valued neural network (CVNN) and aim to increase the number of stored information in a CVNN. Thus, we have used a simple Möbius transformation in the type of reflection $f(z) = \frac{1}{\bar{z}}$ that maps the unit circle onto itself, its inside to its outside and its outside to its inside. Since this function has singularity at the origin and unbounded in the unit circle, we have restricted the input signals of the CVNN to the outside of the unit circle. Therefore, we have guaranteed the boundedness of the function which is an important feature for neural networks. We have employed this activation function to a specific complex valued Hopfield Neural Network (CVHNN) and showed that the fixed points of the activation function are the fixed points of the CVHNN. Finally, we have proved that the fixed points are stable for positive real valued function $\mathrm{Re}\,[H(z)] > 0$.

6. FUTURE WORKS

First off all, we purpose to determine the domain of attraction of the fixed points. Also, we want to try more general Möbius transformation given in Eq.(3). Hence, it is possible to choose place of the fixed points circle in the complex plane because of the property of "every circle in the complex plane can be determined uniquely by a Möbius transformation". In addition, it maybe possible to define different activation functions for each neurons and get different fixed points circles from each of them.

REFERENCES

A. Hirose. *Complex Valued Neural Networks*. Studies in Computational Intelligence, volume 32, Springer-Verlag, 2006.

A.S. Gangal, P.K. Kalra and D.S. Chauhan. Performance evaluation of complex valued neural networks using various error functions. World Academy of Science, volume Engineering and Technology 29, pages 27–32, 2007.

H. Leung and S. Haykin. The complex backpropagation algorithm. *IEEE Transactions on Signal Processing*, volume 39, pages 2101-2104, 1991.

D.L. Birx and S.J. Pipenberg. Chaotic oscillators and complex mapping feed forward networks (CMFFNS) for signal detection in noisy environments. *Proceedings of IEEE International Joint Conference on Neural Networks*, volume II, pages 881–888, 1992.

N. Benvenuto and F. Piazza. On the complex back-propagation algorithm. *IEEE Transactions on Signal Processing*, volume 40, pages 967–969, 1992.

G. Kechriotis and E.S. Monalakos. Training fully recurrent neural networks with complex weights. *IEEE Transactions on Circuits and Systems-II*, volume 41, pages 235-238, 1994.

M. Kinouchi and M. Hagiwara. Learning temporal sequences by complex neurons with local feedback. *Proceedings of IEEE International Conference on Neural Networks*, volume IV, pages 3165-3169, 1995.

A. Hirose. Applications of complex-valued neural networks to coherent optical computing using phase-sensitive detection scheme. *Information Sciences-Applications*, volume 2, pages 103–117, 1994.

A.J. Noest. Associative memory in sparse phasor neural networks. *Europhysics Letters*, volume 6, pages 469-474, 1988.

G.M. Georgiou and C. Koutsougeras. Complex domain backpropagation. *IEEE Transactions on Circuits and Systems-II*, volume 39, pages 330-334, 1992.

Y. Kuroe and T. Taniguchi. Models of self-correlation type complex-valued associative memories and their dynamics. In W. Duch et. al., editor, *Artificial Neural Networks: Biological Inspirations-ICANN 2005, Lecture Notes in Computer Science*, volume 3696, pages 185–192. Springer-Verlag, 2005.

Y. Kuroe and T. Taniguchi. Models of complex-valued Hopfield-type neural networks and their dynamics. In T. Nitta, editor, *Complex-Valued Neural Networks*, pages 123–141. Information Science Reference, Hersey New York, 2009.

T. Kim and T. Adali. Fully complex multi-layer perceptron network for nonlinear signal processing. *Journal of VLSI Signal Processing*, volume 32, pages 29–43, 2002.

T. Clarke. Generalization of Neural Network to the Complex Plane. *Proceedings of International Joint Conference on Neural Networks*, volume 2, pages 435–440, 1990.

D.P. Mandic. The use of Möbius transformations in neural networks and signal processing. *Neural Networks for Signal Processing X*, volume 1, pages 185–194, 2000.

Y.S. Abu-Mostafa and J. St.Jacques. Information capacity of the Hopfield model. *IEEE Transactions on Information Theory*, volume IT-31, pages 461–464, 1985.

G.A. Jones and D. Singerman. Complex Functions an Algebraic and Geometric Viewpoint. *Cambridge University Press*, Great Britain, 1994.

T. Needham. Visual Complex Analysis. *Oxford University Press*, Great Britain, 2000.

L. Brickman. The symmetry principle for Möbius transformations. *American Mathematical Monthly*, volume 100, pages 781–782, 1993.

J.H. Li, A.N. Michel and W. Porod. Qualitative analysis and synthesis of a class of neural networks. *IEEE Transactions on Circuits and Systems*, volume 35, pages 976–986, 1988.

K. Leblebicioğlu, U. Halıcı and O. Çelebi. Infinite dimensional Hopfield neural networks. *Nonlinear Analysis*, volume 47, pages 5807–5813, 2001.

Y. Kuroe, M. Yoshida and T. Mori. On activation functions for complex-valued neural networks - existence of energy functions. In O. Kaynak et al, editor, *ICANN/ICONIP 2003, LNCS 2714*, pages 985–992. Springer-Verlag, Berlin Heidelberg, 2003.

H.K. Khalil. Nonlinear Systems. *Prentice Hall*, United States of America, 1996.