ICONIP 2011

Relationship between phase and amplitude generalization errors in complex- and real-valued feedforward neural networks

Akira Hirose · Shotaro Yoshida

Abstract We compare the generalization characteristics of complex-valued and real-valued feedforward neural networks. We assume a task of function approximation with phase shift and/or amplitude change in signals having various coherence. Experiments demonstrate that complex-valued neural networks show smaller generalization error than real-valued networks having doubled input and output neurons in particular when the signals have high coherence, that is, high degree of wave nature. We also investigate the relationship between amplitude and phase errors. It is found in real-valued networks that abrupt change in amplitude is often accompanied by steep change in phase, which is a consequence of local minima in real-valued supervised learning.

Keywords Complex-valued neural network ·
Function approximation · Generalization

1 Introduction

Researches on complex-valued neural networks (CVNNs) have revealed various aspects of their dynamics. However, at the same time, it is true that a complex number is represented by a pair of real numbers, namely real and imaginary parts, or amplitude and phase. Actually, a variety of useful neural dynamics theories are obtained by paying attention to the real and imaginary parts [1, 10, 11] or amplitude and phase [2, 3]. This fact sometimes leads to an assumption that a CVNN is almost equivalent to a double-dimensional real-valued neural network.

In this paper, we compare complex- and real-valued neural networks by focusing on their generalization characteristics. We investigate the generalization ability of feedforward complex-valued and double-dimensional real-valued neural networks, in particular when they learn and process wave-related data for function approximation or filtering. We observe the characteristics by feeding signals that have various degrees of wave nature by mixing a sinusoidal wave and white noise. Computer experiments demonstrate that the generalization characteristics of CVNNs are different from those of double-dimensional real-valued neural networks (RVNNs) depending on the degree of wave nature of the signals, or the coherence.

This paper is an extension of a conference paper [8]. A statistical evaluation and discussion on this topic are also given in Ref. [9]. Contrarily, we concentrate on the relationship between the amplitude and phase response occurring with the time shift and amplitude changes in the input signals.

This paper is organized as follows. Section 2 reviews a property of complex numbers by representing them as 2×2 matrices and discuss its effect on the supervised learning in non-layered feedforward neural networks. Section 3 presents the construction of the computer experiments and learning dynamics. In Sect. 4, we show the difference in the generalization characteristics experimentally. We find that the generalization ability of CVNNs is higher than that of double-dimensional RVNNs especially when they process signals having high degree of wave nature. We discuss the characteristics specific to respective networks. Section 5 is the conclusion.

A. Hirose (✉) · S. Yoshida
Department of Electrical Engineering and Information Systems,
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku,
Tokyo 113-8656, Japan
e-mail: ahirose@ee.t.u-tokyo.ac.jp
URL: http://eis.t.u-tokyo.ac.jp

∂ Springer

2 Qualitative difference between complex- and real-valued neural networks

2.1 Complex number represented as real 2 × 2 matrix

First, we review the nature of a complex number [5]. As we focus on multiplication out of the four arithmetic operations of complex numbers, we can represent a complex number as a real 2 × 2 matrix. That is, with every complex number $c = a + \sqrt{-1}\,b$, where a and b are real numbers, we associate a $\boldsymbol{C}$-linear transformation

$$T_c : \boldsymbol{C} \to \boldsymbol{C}, \quad z \mapsto cz = ax - by + \sqrt{-1}(bx + ay) \tag{1}$$

If we identify $\boldsymbol{C}$ with $\boldsymbol{R}^2$ by

$$z = x + iy = \begin{pmatrix} x \\ y \end{pmatrix} \tag{2}$$

it follows that

$$
\begin{aligned}
T_c \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix} \\
&= \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}
\tag{3}
$$

In other words, the linear transformation T_c determined by $c = a + ib$ is described by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Generally, a mapping represented by a 2 × 2 matrix is non-commutative. However, in the present case, it becomes commutative.

The most important point of this representation lies in the fact that it clearly expresses the function specific to the complex numbers as the rotation and amplification or attenuation as

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{4}$$

where $r \equiv \sqrt{a^2 + b^2}$ and $\theta \equiv \arctan b/a$ denote amplification or attenuation of amplitude and rotation angle applied to the complex signals z, respectively.

2.2 Phase rotation and amplitude amplification/ attenuation in neural networks

Let us consider how the above feature of the complex number emerges in neural dynamics. Assume a task to realize a mapping that transforms an input $\boldsymbol{x}^{\mathbf{IN}}$ to an output $\boldsymbol{x}^{\mathbf{OUT}}$ shown in Fig. 1a through supervised learning that adjusts the synaptic weights w_{ji}. Simply, we have only a single teacher pair of input and output signals here. We consider a very simple feedforward neural network in the real number domain shown in Fig. 1b having a single-layer, 2-input, 2-output. For simplicity, we omit possible
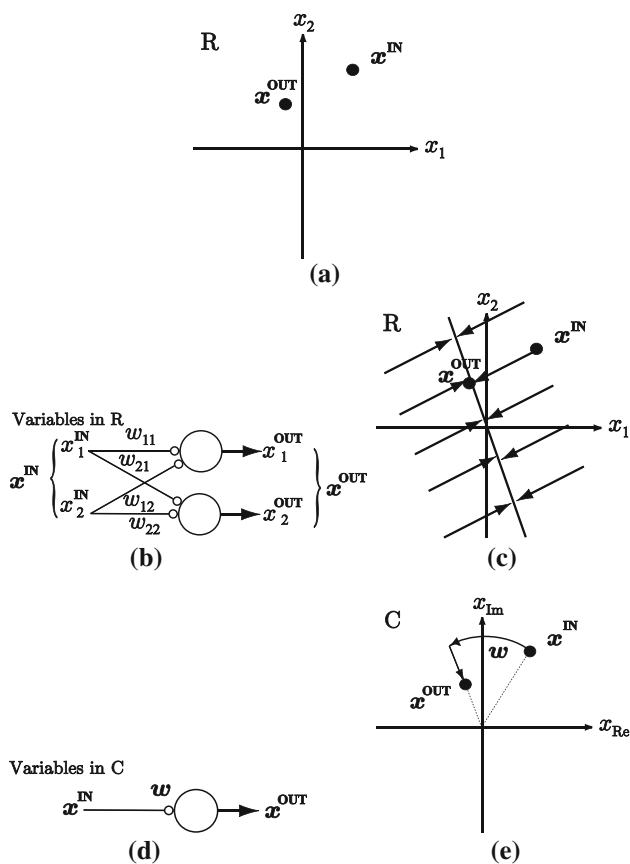
Fig. 1 **a** A task to map $\boldsymbol{x}^{\mathbf{IN}}$ to $\boldsymbol{x}^{\mathbf{OUT}}$, and a simple RVNN to learn the task having **b** 2-input 2-output single-layer structure and **c** a possible but degenerate solution that is often not useful, and a simple CVNN to learn the same task having **b** 1-input 1-output single-layer structure and **c** expected learning result [5]

nonlinearity at the neurons, that is, the activation function is the identity function. Then, the general input–output relationship is described by using four real numbers a, b, c, and d as

$$\begin{pmatrix} x_1^{\mathbf{OUT}} \\ x_2^{\mathbf{OUT}} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1^{\mathbf{IN}} \\ x_2^{\mathbf{IN}} \end{pmatrix} \tag{5}$$

In the present case, we have a variety of possible mappings realized by the learning because the number of parameters to be determined is larger than that of the condition, that is, the learning task is an ill-posed problem. The functional difference among the possible mapping functions emerges as the difference in the generalization characteristics. For example, learning can result in a degenerate mapping shown in Fig. 1c, which is often unuseful in practice.

Next let us consider the learning of the mapping in the complex domain, which transforms a complex value $\boldsymbol{x}^{\mathbf{IN}} = (x_1^{\mathbf{IN}}, x_2^{\mathbf{IN}})$ to another complex value $\boldsymbol{x}^{\mathbf{OUT}} = (x_1^{\mathbf{OUT}}, x_2^{\mathbf{OUT}})$. Figure 1d shows a CVNN, where the weight is a single complex value $w = |w| \exp(\sqrt{-1}\,\theta)$. The situation is expressed just like in (5) with the constraint (4) as

$$\begin{pmatrix} x_1^{\mathbf{OUT}} \\ x_2^{\mathbf{OUT}} \end{pmatrix} = \begin{pmatrix} |w|\cos\theta & -|w|\sin\theta \\ |w|\sin\theta & |w|\cos\theta \end{pmatrix} \begin{pmatrix} x_1^{\mathbf{IN}} \\ x_2^{\mathbf{IN}} \end{pmatrix} \quad (6)$$

The degree of freedom is reduced, and the arbitrariness of the solution is also reduced. That is, we have a unique solution in this case as follows. Figure 1d illustrates the result of the learning. The mapping is a combination of phase rotation and amplitude attenuation.

This property can be a great advantage when we deal with information related with waves such as electromagnetic wave, lightwave, and electron wave. This is an intuitive expectation, and investigated numerically in the following sections.

3 Construction of experiments and learning dynamics

We organize our experiment as follows.

- Preparation of input signals: Variously weighted summation of sinusoidal wave (coherent wave) and nonwave data, that is, white noise having random amplitude and phase (or real and imaginary parts).
- Definition of the task to learn: Identity mapping, which is expected to show the learning characteristics clearly, for the above signals with various degrees of wave nature.
- Evaluation of generalization: Observation of the generalization error when the input signals shift in time and/or when the amplitude is changed.

3.1 Forward processing and learning dynamics

3.1.1 Complex-valued neural network

We consider a layered feedforward network having input terminals, hidden neurons, and output neurons. In the case of a CVNN, we employ a phase-amplitude-type sigmoid activation function and the teacher-signal-backpropagation learning process [3, 7] with notations of

$$z^{\mathrm{I}} = [z_1, \ldots, z_i, \ldots, z_I, z_{I+1}]^T \quad \text{(Input signal vector)} \quad (7)$$

$$z^{\mathrm{H}} = [z_1, \ldots, z_h, \ldots, z_H, z_{H+1}]^T \quad (8)$$
$$\text{(Hidden-layer output signal vector)}$$

$$z^{\mathrm{O}} = [z_1, \ldots, z_o, \ldots, z_O]^T \quad \text{(Output-layer signal vector)} \quad (9)$$

$$\mathbf{W}^{\mathrm{H}} = [w_{hi}] \quad \text{(Hidden neuron weight matrix)} \quad (10)$$

$$\mathbf{W}^{\mathrm{O}} = [w_{oh}] \quad \text{(Output neuron weight matrix)} \quad (11)$$

where $[\cdot]^T$ means transpose. In (10) and (11), the weight matrices include additional weights w_{hI+1} and w_{oH+1}, equivalent to neural thresholds, where we add formal constant inputs $z_{I+1} = 1$ and $z_{H+1} = 1$ in (7) and (8), respectively. Respective signal vectors and synaptic weights are connected with one another through an activation function $f(z)$ as

$$z^{\mathrm{H}} = f(\mathbf{W}^{\mathrm{H}} z^{\mathrm{I}}), \quad z^{\mathrm{O}} = f(\mathbf{W}^{\mathrm{O}} z^{\mathrm{H}}) \quad (12)$$

where $f(z)$ is a function of each vector element $z\ (\in \boldsymbol{C})$ defined as

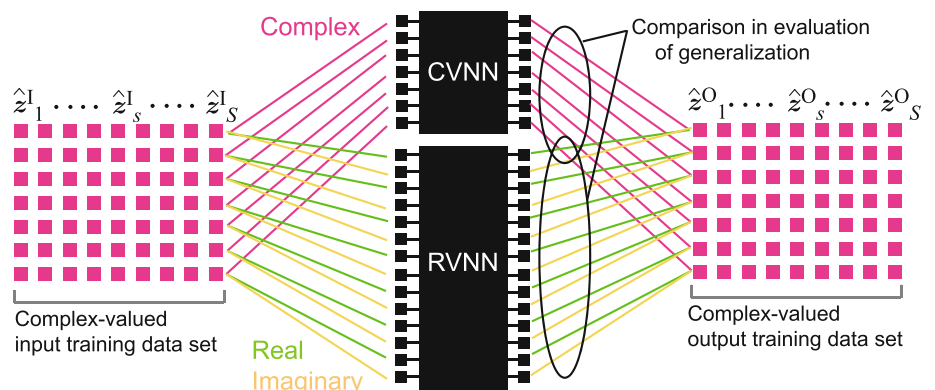$$f(z) = \tanh(|z|)\exp\left(\sqrt{-1}\,\arg z\right) \quad (13)$$

Figure 2 is a diagram to explain the supervised learning process. We prepare a set of teacher signals at the input $\hat{z}_s^{\mathrm{I}} = [\hat{z}_{1s}, \ldots, \hat{z}_{is}, \ldots, \hat{z}_{Is}, \hat{z}_{I+1\,s}]^T$ and the output $\hat{z}_s^{\mathrm{O}} = [\hat{z}_{1s}, \ldots, \hat{z}_{os}, \ldots, \hat{z}_{Os}]^T\ (s = 1, \ldots, s, \ldots S)$ for which we employ the teacher-signal backpropagation learning. We define an error function E to obtain the dynamics [3, 7] as

$$E \equiv \frac{1}{2} \sum_{s=1}^{S} \sum_{o=1}^{O} \left| z_o(\hat{z}_s^{\mathrm{I}}) - \hat{z}_{os} \right|^2 \quad (14)$$

$$|w_{oh}^{\mathrm{new}}| = |w_{oh}^{\mathrm{old}}| - K \frac{\partial E}{\partial |w_{oh}|} \quad (15)$$

$$\arg w_{oh}^{\mathrm{new}} = \arg w_{oh}^{\mathrm{old}} - K \frac{1}{|w_{oh}|} \frac{\partial E}{\partial(\arg w_{oh})} \quad (16)$$

Fig. 2 Schematic diagram of the learning process of complex- and double-dimensional real-valued feedforward neural networks for pairs of input–output teachers

$$\frac{\partial E}{\partial |w_{oh}|} = \left(1 - |z_o|^2\right)\left(|z_o| - |\hat{z}_o|\cos(\arg z_o - \arg \hat{z}_o)\right)$$
$$\times |z_h|\cos(\arg z_o - \arg \hat{z}_o - \arg w_{oh})$$
$$- |z_o||\hat{z}_o|\sin(\arg z_o - \arg \hat{z}_o)\frac{|z_h|}{\tanh^{-1}|z_o|}$$
$$\times \sin(\arg z_o - \arg \hat{z}_o - \arg w_{oh}) \quad (17)$$

$$\frac{1}{|w_{oh}|}\frac{\partial E}{\partial(\arg w_{oh})}$$
$$= \left(1 - |z_o|^2\right)\left(|z_o| - |\hat{z}_o|\cos(\arg z_o - \arg \hat{z}_o)\right)|z_h|$$
$$\times \sin(\arg z_o - \arg \hat{z}_o - \arg w_{oh})$$
$$+ |z_o||\hat{z}_o|\sin(\arg z_o - \arg \hat{z}_o)\frac{|z_h|}{\tanh^{-1}|z_o|}$$
$$\times \cos(\arg z_o - \arg \hat{z}_o - \arg w_{oh}) \quad (18)$$

where $(\cdot)^{\text{new}}$ and $(\cdot)^{\text{old}}$ indicate the update of the weights from $(\cdot)^{\text{old}}$ to $(\cdot)^{\text{new}}$ and K is a learning constant. The teacher signals at the hidden layer $\hat{z}^{\text{H}} = [\hat{z}_1, \ldots, \hat{z}_h, \ldots, \hat{z}_H]^T$ are obtained by making the output teacher vector itself $\hat{z}^{\text{O}}$ propagate backward as

$$\hat{z}^{\text{H}} = \left(f\left((\hat{z}^{\text{O}})^*\hat{\mathbf{W}}^{\text{O}}\right)\right)^* \quad (19)$$

where $(\cdot)^*$ denotes Hermite conjugate. Using $\hat{z}^{\text{H}}$, the hidden-layer neurons change their weights by following (15)–(18) with replacement of the suffixes o, h with h, i [4, 6].

3.1.2 Double-dimensional real-valued neural network

Similarly, the forward processing and learning of a double-dimensional RVNN is explained as follows. Figure 2 includes also this case. We represent a complex number as a pair of real numbers as $z_i = x_{2i-1} + \sqrt{-1}\,x_{2i}$. That is, we have a double-dimensional real input vector z_{R}^{I}, a double-dimensional hidden signal vector z_{R}^{H}, and a double-dimensional output signal vector z_{R}^{O}. A forward signal processing connects the signal vectors as well as hidden neuron weights $\mathbf{W}_{\text{R}}^{\text{H}}$ and output neuron weights $\mathbf{W}_{\text{R}}^{\text{O}}$ through a real-valued activation function f_{R} as

$$z_{\text{R}}^{\text{I}} = \left[\overbrace{x_1,\ x_2}^{\text{real\&imaginary}}, \ldots, x_{2i-1}, x_{2i}, \ldots, x_{2I-1}, x_{2I}, x_{2I+1}, x_{2I+2}\right]^T$$
$$\left(= z^{\text{I}}\right) \quad \text{(Input signal vector)} \quad (20)$$

$$z_{\text{R}}^{\text{H}} = [x_1, x_2, \ldots, x_{2h-1}, x_{2h}, \ldots, x_{2H-1}, x_{2H}, x_{2H+1}, x_{2H+2}]^T$$
$$\text{(Hidden-layer output signal vector)} \quad (21)$$

$$z_{\text{R}}^{\text{O}} = [x_1, x_2, \ldots, x_{2o-1}, x_{2o}, \ldots, x_{2O-1}, x_{2O}]^T \quad (22)$$
$$\text{(Output-layer signal vector)}$$

$$\mathbf{W}_{\text{R}}^{\text{H}} = [w_{\text{R}hi}] \quad \text{(Hidden neuron weight matrix)} \quad (23)$$

$$\mathbf{W}_{\text{R}}^{\text{O}} = [w_{\text{R}oh}] \quad \text{(Output neuron weight matrix)} \quad (24)$$

$$z_{\text{R}}^{\text{H}} = f_{\text{R}}\left(\mathbf{W}_{\text{R}}^{\text{H}}z_{\text{R}}^{\text{I}}\right), \quad z_{\text{R}}^{\text{O}} = f_{\text{R}}\left(\mathbf{W}_{\text{R}}^{\text{O}}z_{\text{R}}^{\text{H}}\right) \quad (25)$$

$$f_{\text{R}}(x) = \tanh(x) \quad (26)$$

where the thresholds are $w_{\text{R}h2I+1}$, $w_{\text{R}h2I+2}$, $w_{\text{R}h2H+1}$, and $w_{\text{R}h2H+2}$ with formal additional inputs $x_{2H+1} = 1$, $x_{2H+2} = 1$, $x_{2H+1} = 1$, and $x_{2H+2} = 1$. We employ the conventional error backpropagation learning.

4 Computer experiments

4.1 Experimental setup

We chose the identity mapping as the task to be learned to show the network characteristics most clearly. To generate input signals as a function of time $z(t)$ having several degrees of coherence, we added white Gaussian noise $n(t)$ to a sinusoidal wave $\sin \omega t$ (angular frequency ω) with various weighting as $z(t) = a_{\text{s}}\sin \omega t + a_{\text{n}}n(t)$ where a_{s} and a_{n} denote equivalent amplitude. Then, the degree of wave nature is expressed as the signal-to-noise ratio SNR $\equiv$ $a_{\text{s}}/a_{\text{n}}$ where SNR $= \infty$ means complete wave, while SNR $= 0$ corresponds to complete non-wave. The network parameters are as follows: Number of input neurons $I = 16$, hidden neurons $H = 25$, output neurons $O = 16$, learning constant $K = 0.01$, and the learning iteration = 3,000.

4.2 Results and discussion

Figures 3, 4, 5, 6 display typical examples of the learning curves and output signals when SNR $= \infty$, 20dB, 10dB, and 0dB, respectively. Figure 3a shows an example of the learning curve when SNR $= \infty$, that is, the signal is sinusoidal. We find that the learning is almost successfully completed for both the CVNN and RVNN. The learning errors converge roughly at zero, which means that there is only slight residual error at the learning teacher points. However, the curves are different from each other at the beginning of the learning. The curve of the CVNN shows quick decreases. Contrarily, that of the RVNN has a plateau just after the beginning and then steep decrease. This tendency is often observed in the RVNN to learn high-coherence signals, which implies that the RVNN is subject to local minima.
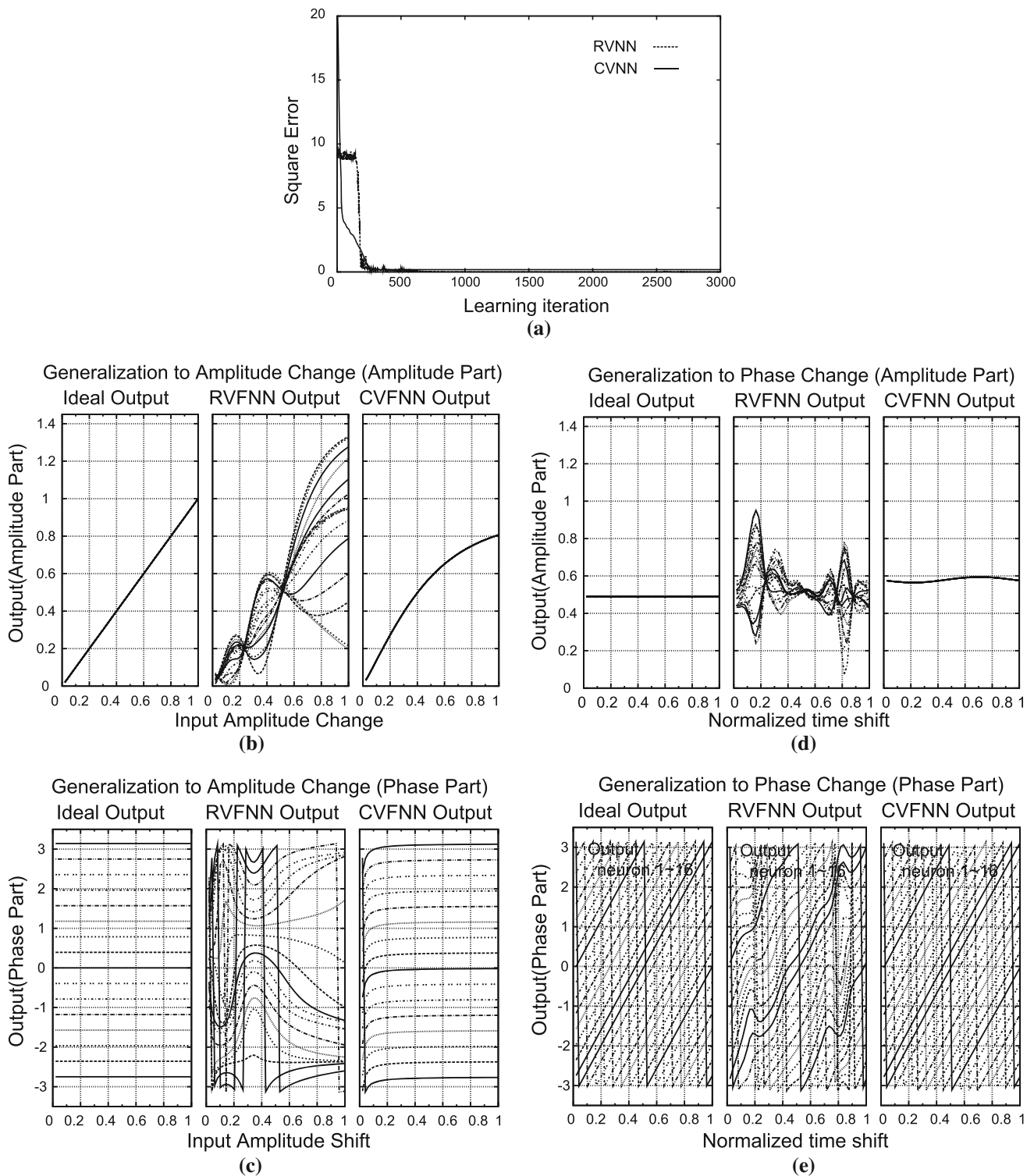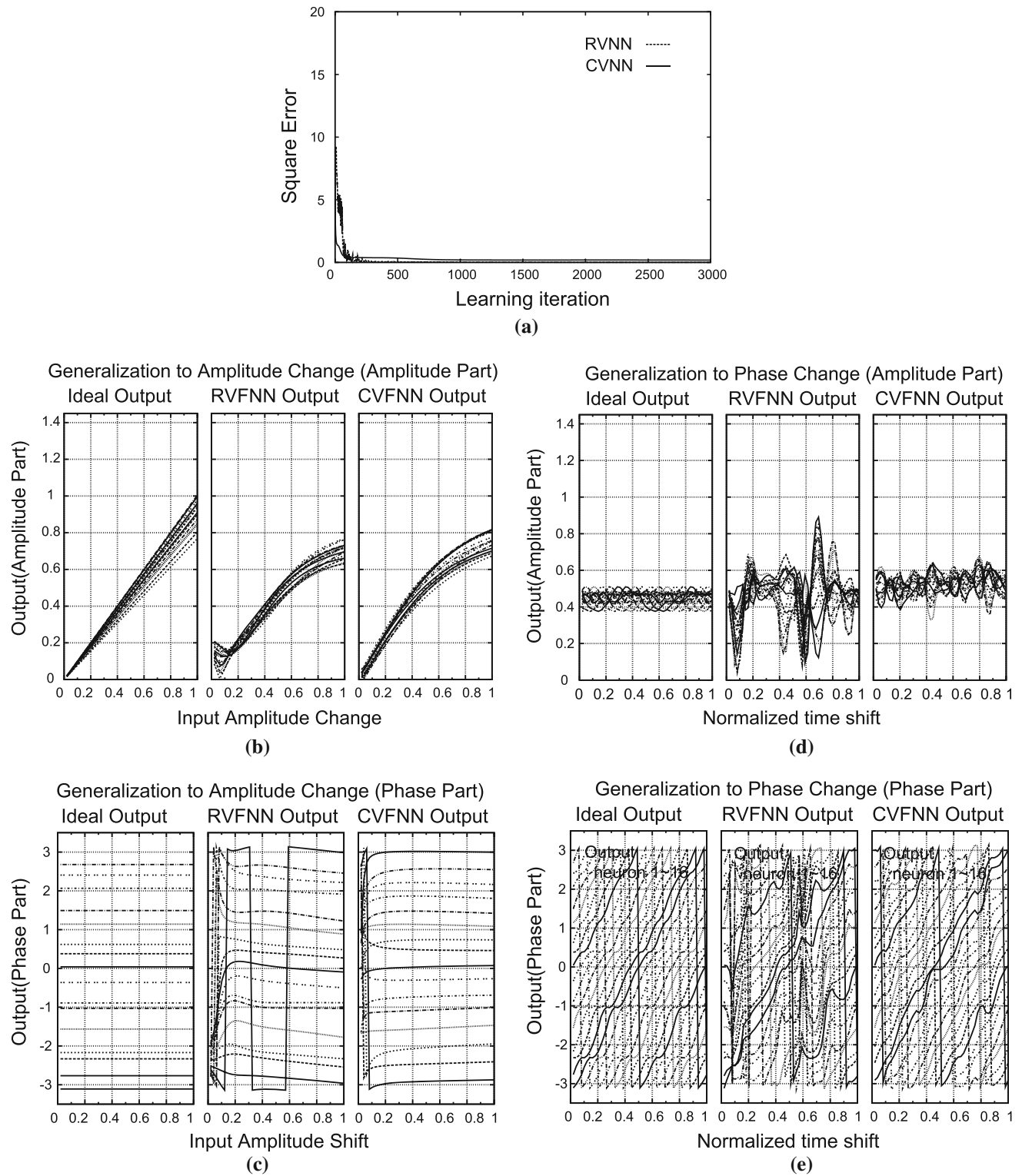
Fig. 3 Example of **a** learning curve, **b** amplitude, and **c** phase when the input signal amplitude gradually changed, and **d** amplitude and **e** phase when the input signal gradually shifts in time, in the real-valued and complex-valued neural networks (RVNN and CVNN) when no noise is added to sinusoidal signals ($S/N = \infty$)

After the learning, we feed other input signals to investigate the generalization. As mentioned above, the wavelength is adjusted to span over the 16 neural input terminals.

Figure 3b and c presents examples of the output amplitude and phase, respectively, showing from left-hand side to the right-hand side the ideal output of the identity mapping, the

(a)

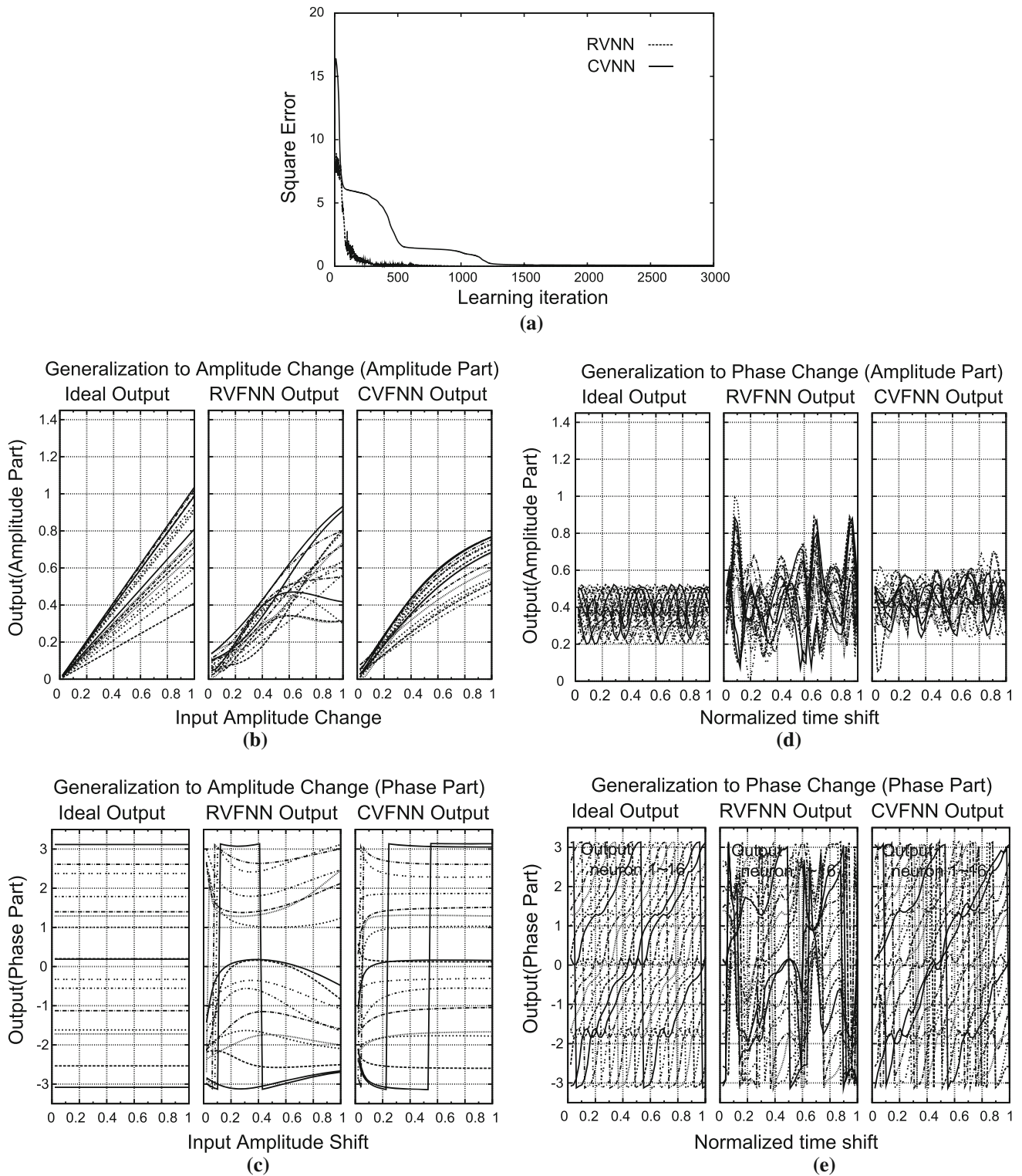
(b)

(c)

(d)

(e)

Fig. 4 Example of **a** learning curve, **b** amplitude, and **c** phase when the input signal amplitude gradually changed, and **d** amplitude and **e** phase when the input signal gradually shifts in time, in the real-valued and complex-valued neural networks (RVNN and CVNN) when no noise is added to sinusoidal signals (*S/N* = 20 dB)

RVNN outputs, and the CVNN outputs of the 16 output neurons. The horizontal axis shows the amplitude changing from 0 to 1. Figure 3d and e shows the output amplitude and phase when the input signal is shifted in time. The horizontal axes present the time shift *t* normalized by the unit-wave duration *T*.
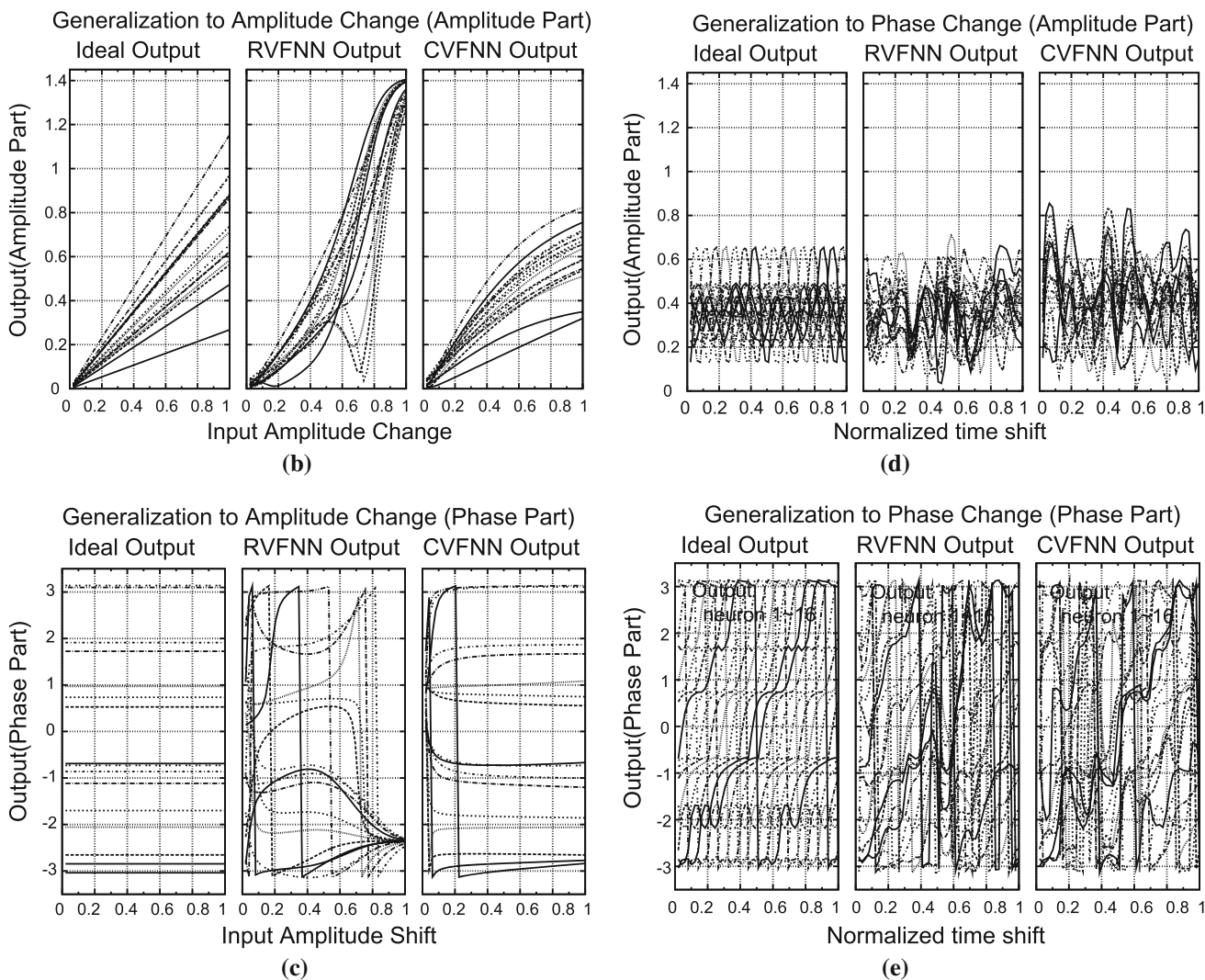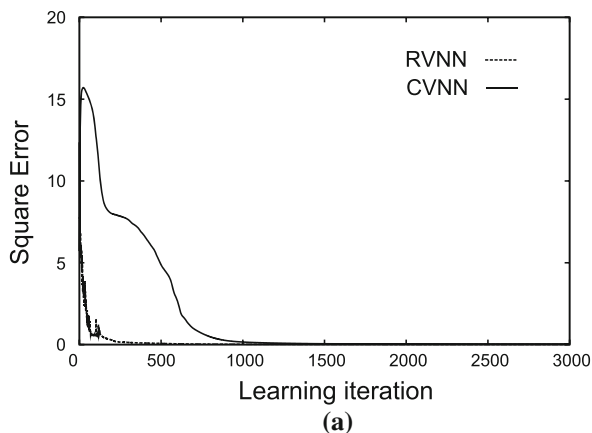
Fig. 5 Example of **a** learning curve, **b** amplitude, and **c** phase when the input signal amplitude gradually changed, and **d** amplitude and **e** phase when the input signal gradually shifts in time, in the real-valued and complex-valued neural networks (RVNN and CVNN) when no noise is added to sinusoidal signals ($S / N = 10$dB)

We find that the RVNN output amplitude and phase values are very different from the ideal ones. The learning has been conducted at 16 snapshots of the signal waveform, that is, four points in the amplitude ($a = 0, 0.25, 0.5, 0.75$) multiplied by four phase shifts, or time shifts t normalized by the signal wave period T ($t/T = 0$, 1/8, 2/8, 3/8) plus initial waveform phase at respective

Fig. 6 Example of **a** learning curve, **b** amplitude, and **c** phase when the input signal amplitude gradually changed, and **d** amplitude and **e** phase when the input signal gradually shifts in time, in the real-valued

and complex-valued neural networks (RVNN and CVNN) when no noise is added to sinusoidal signals ($S/N = 0$ dB)

neurons $2i\pi/16\,(i = 0, 1, 2, \ldots, 15)$. (For the details of the learning process, see Ref. [9].) In each charts in Fig. 3b, c, d, or e, we have 16 curves corresponding to the 16 neuron outputs.

Figure 3b shows the output amplitude response when the input signal amplitude is changed. The ideal output (left-hand side) is given as the proportional outputs where

16 curves are identical. However, the RVNN output amplitudes are largely different. At the learning points of 0, 0.25, and 0.5, the curves are forced to converge at almost ideal values. The phase values at 0.25 and 0.5 in Fig. 3c also show values near to the ideal ones. However, in total, it deviates very largely from the ideal line, though at the 0-amplitude point the phase value means nothing.

At the last learning point, amplitude of 0.75, the state is different. The amplitude values do not converge but are scattered instead. The phase values also differ from the ideal ones. The result implies that the network sought optimal solution in the scattered amplitude condition, which is a local minimum.

In contrast, the 16 amplitude outputs of the CVNN in Fig. 3b are identical with one another, which situation is the same as the ideal one, though the amplitude curves show saturation at the large amplitude region. This result reflects directly the saturation characteristic of the neuron activation function. As an optimal learning result, the network shows slightly larger amplitude in the small input-amplitude region, while a little smaller amplitude in the large input region. The phase values in Fig. 3c are similar to ideal ones, though at around 0-amplitude they are meaningless and deviating.

Next, we observe the responses to the time shift (or phase shift) of the input signal. The horizontal axes in Figs. 3d and e show the time shift normalized by the wave period T. Figure 3d presents the output amplitude when the input amplitude is fixed at 0.5. Ideally, it should be 0.5 constantly. However, the RVNN outputs deviate very largely again. The phase values in Fig. 3e also deviate from the ideal ones. The large phase error regions in (e) correspond to the regions of steep amplitude changes in (d). Contrarily, the CVNN output amplitude values in Fig. 3d are almost constant. The value is a little different from 0.5 because of the nonlinearity of the neuron activation function. The phase values of the CVNN in Fig. 3e are almost identical with the ideal lines.

As seen above, the CVNN presents better generalization ability in both amplitude and phase for coherent signals. Its feature is obvious in the response including the phase rotation observed clearly as the phase stability against input amplitude changes as well as the linear phase changes versus the input phase shift. These characteristics match the single neuron dynamics in Fig. 1e in Sect. 2, which shows that the elemental process consists of phase rotation as well as amplitude change if needed.

Figures 4, 5, 6 show the data for SNR = 20dB, 10dB, and 0dB, respectively. As the degree of wave nature decreases, the generalization error increases. However, in any SNR case, both the amplitude and phase of the CVNN exhibit better generalization than those of the RVNN.

Note that the time required for the learning can also be longer, in particular for the signals with lower degree of wave nature (smaller the SNR, lower the coherence). This fact is attributed to the smaller degree of freedom of the CVNNs described in Sect. 2.2.

The correspondence between steep changes in amplitude and phase is sometimes observed also in these low coherence cases. For example, in Fig. 6b where SNR = 0dB, at amplitude of about 0.75, the amplitude has a sharp dip for several neuron outputs. Correspondingly, in Fig. 6c, we can find phase change in the phase outputs of the same neurons. Such changes are observed only in the RVNN where there is no implicit limitation of phase-and-amplitude elemental dynamics.

5 Conclusion

This paper investigated numerically the generalization characteristics in the feedforward complex-valued and real-valued neural networks (CVNN and RVNN). We compared a CVNN and a double-dimensional RVNN in a simple case where the networks deal with the task of function approximation. Computer experiments demonstrated that the CVNN exhibits better generalization characteristics in particular for signals having high degree of wave nature, that is, coherence. This fact is attributed to the smaller degree of freedom of the CVNN than that of the RVNN, resulting in a learning tendency to assume phase rotation and amplitude amplification or attenuation. We also investigated the relationship between the amplitude and phase errors. We found in the RVNN that abrupt change in amplitude is often accompanied by a steep change in phase. This phenomenon is a consequence of local minima in the RVNN and not observed in the CVNN. These characteristics of the CVNN are expected to be used in many applications to deal with wave phenomena and wave-related information processing.

References

1. Benvenuto N, Piazza F (1992) On the complex backpropagation algorithm. IEEE Trans Sign Process 40:967–969
2. Georgiou GM, Koutsougeras C (1992) Complex domain back-propagation. IEEE Trans Circ Syst II 39(5):330–334
3. Hirose A (1992) Continuous complex-valued back-propagation learning. Electron Lett 28(20):1854–1855
4. Hirose A (1994) Applications of complex-valued neural networks to coherent optical computing using phase-sensitive detection scheme. Inform Sci Appl 2:103–117
5. Hirose A (2011) Nature of complex number and complex-valued neural networks. Frontiers of Electrical and Electronic Engineering in China 6(1):171–180

6. Hirose A (2012) Complex-valued neural networks, 2nd (ed) Springer, Heidelberg

7. Hirose A, Eckmiller R (1996) Behavior control of coherent-type neural networks by carrier-frequency modulation. IEEE Trans Neural Netw 7(4):1032–1034

8. Hirose A, Yoshida S (November 2011) Comparison of complex- and real-valued feedforward neural networks in their generalization ability. In: International Conference on Neural Information Processing (ICONIP) 2011 Shanghai. Springer, pp. 526–531

9. Hirose A, Yoshida S (2012) Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence. IEEE Trans Neural Netw Learn Syst 23:541–551

10. Leung H, Haykin S (1991) The complex backpropagation algorithm. IEEE Trans Sign Process 39:2101–2104

11. Nitta T (1997) An extension of the back-propagation algorithm to complex numbers. Neural Netw 10:1391–1415