

*Research Article*

# **Convergence of an Online Split-Complex Gradient Algorithm for Complex-Valued Neural Networks**

**Huisheng Zhang,<sup>1</sup> Dongpo Xu,<sup>2</sup> and Zhiping Wang<sup>1</sup>**

<sup>1</sup> *Department of Mathematics, Dalian Maritime University, Dalian 116026, China*

<sup>2</sup> *Department of Applied Mathematics, Harbin Engineering University, Harbin 150001, China*

Correspondence should be addressed to Huisheng Zhang, zhhuisheng@163.com

Received 1 September 2009; Accepted 19 January 2010

Academic Editor: Manuel De La Sen

Copyright © 2010 Huisheng Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The online gradient method has been widely used in training neural networks. We consider in this paper an online split-complex gradient algorithm for complex-valued neural networks. We choose an adaptive learning rate during the training procedure. Under certain conditions, by firstly showing the monotonicity of the error function, it is proved that the gradient of the error function tends to zero and the weight sequence tends to a fixed point. A numerical example is given to support the theoretical findings.

## **1. Introduction**

In recent years, neural networks have been widely used because of their outstanding capability of approximating nonlinear models. As an important search method in optimization theory, gradient algorithm has been applied in various engineering fields, such as adaptive control and recursive parametrical estimation [1–3]. Gradient algorithm is also a popular training method for neural networks (when used to train neural networks with hidden layers, gradient algorithm is also called BP algorithm) and can be done either in the online or in the batch mode [4]. In online training, weights are updated after the presentation of each training example, while in batch training, weights are not updated until all of the examples are inputted into the networks. As a result, batch gradient training algorithm is always used when the number of training samples is relatively small. However, in the case that a very large number of training samples are available, online gradient training algorithm is preferred.

Conventional neural networks' parameters are usually real numbers for dealing with real-valued signals [5, 6]. In many applications, however, the inputs and outputs of a system are best described as complex-valued signals and processing is done in complex space. In order to solve the problem in complex domain, complex-valued neural networks (CVNNs)

have been proposed in recent years [7–9], which are the extensions of the usual real-valued neural networks to complex numbers. Accordingly, there are two types of generalized gradient training algorithm for complex-valued neural networks: fully complex gradient algorithm [10–12] and split-complex gradient algorithm [13, 14]; both of which can be processed in online mode and batch mode. It has been pointed out that the split-complex gradient algorithm can avoid the problems resulting from the singular points [14].

Convergence is of primary importance for a training algorithm to be successfully used. There have been extensive research results concerning the convergence of gradient algorithm for real-valued neural networks (see, e.g., [15, 16] and the references cited therein), covering both of online mode and batch mode. In comparison, the convergence properties for complex gradient algorithm are seldom investigated. We refer the reader to [11, 12] for some convergence results of fully complex gradient algorithms and [17] for those of batch split-complex gradient (OSCG) algorithm for complex-valued neural networks has not yet been established in the literature, and this becomes our primary concern in this paper. Under certain conditions, by firstly showing the monotonicity of the error function, we prove that the gradient of the error function tends to zero and the weight sequence tends to a fixed point. A numerical example is also given to support the theoretical findings.

The remainder of this paper is organized as follows. The CVNN model and the OSCG algorithm are described in the next section. Section 3 presents the main results. The proofs of these results are postponed to Section 4. In Section 5 we give a numerical example to support our theoretical findings. The paper ends with some conclusions given in Section 6.

## 2. Network Structure and Learning Method

It has been shown that two-layered CVNN can solve many problems that cannot be solved by real-valued neural networks with less than three layers [13]. Thus, without loss of generalization, this paper considers a two-layered CVNN consisting of  $L$  input neurons and 1 output neuron. For any positive integer  $d$ , the set of all  $d$ -dimensional complex vectors is denoted by  $\mathbb{C}^d$  and the set of all  $d$ -dimensional real vectors is denoted by  $\mathbb{R}^d$ . Let us write  $\mathbf{w} = \mathbf{w}^R + i\mathbf{w}^I = (w_1, w_2, \dots, w_L)^T \in \mathbb{C}^L$  as the weight vector between the input neurons and output neuron, where  $w_l = w_l^R + iw_l^I$ ,  $w_l^R$  and  $w_l^I \in \mathbb{R}^1$ ,  $i = \sqrt{-1}$ , and  $l = 1, \dots, L$ . For input signals  $\mathbf{z} = (z_1, z_2, \dots, z_L)^T = \mathbf{x} + i\mathbf{y} \in \mathbb{C}^L$ , where  $\mathbf{x} = (x_1, x_2, \dots, x_L)^T \in \mathbb{R}^L$ , and  $\mathbf{y} = (y_1, y_2, \dots, y_L)^T \in \mathbb{R}^L$ , the input of the output neuron is

$$\begin{aligned} U &= U^R + iU^I = \sum_{l=1}^L (w_l^R x_l - w_l^I y_l) + i \sum_{l=1}^L (w_l^I x_l + w_l^R y_l) \\ &= \begin{pmatrix} \mathbf{w}^R \\ -\mathbf{w}^I \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + i \begin{pmatrix} \mathbf{w}^I \\ \mathbf{w}^R \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}. \end{aligned} \quad (2.1)$$

Here “ $\cdot$ ” denotes the inner product of two vectors.

For the convenience of using OSCG algorithm to train the network, we consider the following popular real-imaginary-type activation function [13]:

$$f_C(U) = f_R(U^R) + if_R(U^I) \quad (2.2)$$

for any  $U = U^R + iU^I \in \mathbb{C}^1$ , where  $f_R$  is a real function (e.g., sigmoid function). If simply denoting  $f_R$  as  $f$ , the network output  $O$  is given by

$$O = O^R + iO^I = f(U^R) + if(U^I). \quad (2.3)$$

Let the network be supplied with a given set of training examples  $\{\mathbf{z}^q, d^q\}_{q=1}^Q \subset \mathbb{C}^L \times \mathbb{C}^1$ . For each input  $\mathbf{z}^q = \mathbf{x}^q + i\mathbf{y}^q$  ( $1 \leq q \leq Q$ ) from the training set, we write  $U^q = U^{q,R} + iU^{q,I}$  as the input for the output neuron and  $O^q = O^{q,R} + iO^{q,I}$  as the actual output. The square error function can be represented as follows:

$$\begin{aligned} E(\mathbf{w}) &= \frac{1}{2} \sum_{q=1}^Q (O^q - d^q)(O^q - d^q)^* = \frac{1}{2} \sum_{q=1}^Q \left[ (O^{q,R} - d^{q,R})^2 + (O^{q,I} - d^{q,I})^2 \right] \\ &= \sum_{q=1}^Q \left[ \mu_{qR}(U^{q,R}) + \mu_{qI}(U^{q,I}) \right], \end{aligned} \quad (2.4)$$

where “\*” signifies complex conjugate, and

$$\mu_{qR}(t) = \frac{1}{2} (f(t) - d^{q,R})^2, \quad \mu_{qI}(t) = \frac{1}{2} (f(t) - d^{q,I})^2, \quad t \in \mathbb{R}^1, \quad 1 \leq q \leq Q. \quad (2.5)$$

The neural network training problem is to look for the optimal choice  $\mathbf{w}^*$  of the weights so as to minimize approximation error. The gradient method is often used to solve the minimization problem. Differentiating  $E(\mathbf{w})$  with respect to the real parts and imaginary parts of the weight vectors, respectively, gives

$$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}^R} = \sum_{q=1}^Q \left[ \mu'_{qR}(U^{q,R}) \mathbf{x}^q + \mu'_{qI}(U^{q,I}) \mathbf{y}^q \right], \quad (2.6)$$

$$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}^I} = \sum_{q=1}^Q \left[ -\mu'_{qR}(U^{q,R}) \mathbf{y}^q + \mu'_{qI}(U^{q,I}) \mathbf{x}^q \right]. \quad (2.7)$$

Now we describe the OSCG algorithm. Given initial weights  $\mathbf{w}^0 = \mathbf{w}^{0,R} + i\mathbf{w}^{0,I}$  at time 0, OSCG algorithm updates the weight vector  $\mathbf{w}$  by dealing with the real part  $\mathbf{w}^R$  and  $\mathbf{w}^I$  separately:

$$\begin{aligned} \mathbf{w}^{mQ+q,R} &= \mathbf{w}^{mQ+q-1,R} - \eta_m \left[ \mu'_{qR}(\mathbf{w}^{mQ+q-1,R} \cdot \mathbf{x}^q - \mathbf{w}^{mQ+q-1,I} \cdot \mathbf{y}^q) \mathbf{x}^q \right. \\ &\quad \left. + \mu'_{qI}(\mathbf{w}^{mQ+q-1,I} \cdot \mathbf{x}^q + \mathbf{w}^{mQ+q-1,R} \cdot \mathbf{y}^q) \mathbf{y}^q \right], \\ \mathbf{w}^{mQ+q,I} &= \mathbf{w}^{mQ+q-1,I} - \eta_m \left[ -\mu'_{qR}(\mathbf{w}^{mQ+q-1,R} \cdot \mathbf{x}^q - \mathbf{w}^{mQ+q-1,I} \cdot \mathbf{y}^q) \mathbf{y}^q \right. \\ &\quad \left. + \mu'_{qI}(\mathbf{w}^{mQ+q-1,I} \cdot \mathbf{x}^q + \mathbf{w}^{mQ+q-1,R} \cdot \mathbf{y}^q) \mathbf{x}^q \right], \\ &\quad m = 0, 1, \dots, \quad q = 1, 2, \dots, Q. \end{aligned} \quad (2.8)$$

For  $k, q = 1, 2, \dots, Q$ , and  $m = 0, 1, \dots$ , denote that

$$\begin{aligned}
U^{mQ+k,q,R} &= \mathbf{w}^{mQ+k,R} \cdot \mathbf{x}^q - \mathbf{w}^{mQ+k,I} \cdot \mathbf{y}^q, \\
U^{mQ+k,q,I} &= \mathbf{w}^{mQ+k,I} \cdot \mathbf{x}^q + \mathbf{w}^{mQ+k,R} \cdot \mathbf{y}^q, \\
\Delta \mathbf{w}^{mQ+k,R} &= \mathbf{w}^{mQ+k,R} - \mathbf{w}^{mQ+k-1,R}, \\
\Delta \mathbf{w}^{mQ+k,I} &= \mathbf{w}^{mQ+k,I} - \mathbf{w}^{mQ+k-1,I}, \\
\mathbf{P}_R^{m,q,k} &= \mu'_{qR} \left( U^{mQ+k-1,q,R} \right) \mathbf{x}^q + \mu'_{qI} \left( U^{mQ+k-1,q,I} \right) \mathbf{y}^q, \\
\mathbf{P}_I^{m,q,k} &= -\mu'_{qR} \left( U^{mQ+k-1,q,R} \right) \mathbf{y}^q + \mu'_{qI} \left( U^{mQ+k-1,q,I} \right) \mathbf{x}^q.
\end{aligned} \tag{2.9}$$

Then (2.8) can be rewritten as

$$\begin{aligned}
\Delta \mathbf{w}^{mQ+q,R} &= -\eta_m \mathbf{P}_R^{m,q,q}, \\
\Delta \mathbf{w}^{mQ+q,I} &= -\eta_m \mathbf{P}_I^{m,q,q}.
\end{aligned} \tag{2.10}$$

Given  $0 < \eta_0 \leq 1$  and a positive constant  $N$ , we choose learning rate  $\eta_m$  as

$$\frac{1}{\eta_{m+1}} = \frac{1}{\eta_m} + N, \quad m = 0, 1, \dots \tag{2.11}$$

Equation (2.11) can be rewritten as

$$\eta_m = \frac{\eta_0}{1 + N\eta_0 m} = O\left(\frac{1}{m}\right), \tag{2.12}$$

and this implies that

$$\eta_m \leq \frac{1}{Nm}. \tag{2.13}$$

This type of learning rate is often used in the neural network training [16].

For the convergence analysis of OSCG algorithm, similar to the batch version of split-complex gradient algorithm [17], we shall need the following assumptions.

(A1) There exists a constant  $c_1 > 0$  such that

$$\max_{t \in \mathbb{R}^1} \{|f(t)|, |f'(t)|, |f''(t)|\} \leq c_1. \tag{2.14}$$

(A2) The set  $\Phi_0 = \{\mathbf{w} \mid \partial E(\mathbf{w})/\partial \mathbf{w}^R = 0, \partial E(\mathbf{w})/\partial \mathbf{w}^I = 0\}$  contains only finite points.

### 3. Main Results

In this section, we will give several lemmas and the main convergence theorems. The proofs of those results are postponed to the next section.

In order to derive the convergence theorem, we need to estimate the values of the error function (2.4) at two successive cycles of the training iteration. Denote that

$$\mathbf{r}_R^{m,k} = \mathbf{p}_R^{m,k,k} - \mathbf{p}_R^{m,k}, \quad \mathbf{r}_I^{m,k} = \mathbf{p}_I^{m,k,k} - \mathbf{p}_I^{m,k}, \quad m = 0, 1, \dots, k = 1, \dots, Q, \quad (3.1)$$

where  $\mathbf{p}_R^{m,k} \triangleq \mathbf{p}_R^{m,k,1}$ , and  $\mathbf{p}_I^{m,k} \triangleq \mathbf{p}_I^{m,k,1}$ . The first lemma breaks the changes of error function (2.4) at two successive cycles of the training iteration into several terms.

**Lemma 3.1.** *Suppose Assumption (A1) is valid. Then one has*

$$\begin{aligned} E(\mathbf{w}^{(m+1)Q}) - E(\mathbf{w}^{mQ}) &= -\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\ &\quad - \eta_m \left[ \left( \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_R^{m,k} \right) + \left( \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_I^{m,k} \right) \right] \\ &\quad + \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I}), \end{aligned} \quad (3.2)$$

where  $\rho_{m,q,R} = (1/2)\mu_{qR}''(t_1^{m,q})(U^{(m+1)Q,q,R} - U^{mQ,q,R})^2$ ,  $\rho_{m,q,I} = (1/2)\mu_{qI}''(t_2^{m,q})(U^{(m+1)Q,q,I} - U^{mQ,q,I})^2$ , each  $t_1^{m,q} \in \mathbb{R}^1$  lies on the segment between  $U^{(m+1)Q,q,R}$  and  $U^{mQ,q,R}$ , and each  $t_2^{m,q} \in \mathbb{R}^1$  lies on the segment between  $U^{(m+1)Q,q,I}$  and  $U^{mQ,q,I}$ .

The second lemma gives the estimations on some terms of (3.2).

**Lemma 3.2.** *Suppose Assumptions (A1) and (A2) hold, for  $0 < \eta_0 \leq 1$ , then one has*

$$\left\| \mathbf{r}_R^{m,k} \right\| \leq c_2 \eta_m \sum_{q=1}^{k-1} \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right), \quad k = 2, \dots, Q, \quad (3.3)$$

$$\left\| \mathbf{r}_I^{m,k} \right\| \leq c_2 \eta_m \sum_{q=1}^{k-1} \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right),$$

$$\begin{aligned} &\left| \eta_m \left( \left( \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_R^{m,k} \right) + \left( \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_I^{m,k} \right) \right) \right| \\ &\leq c_3 \eta_m^2 \left( \left( \sum_{q=1}^Q \left\| \mathbf{p}_R^{m,q} \right\| \right)^2 + \left( \sum_{q=1}^Q \left\| \mathbf{p}_I^{m,q} \right\| \right)^2 \right), \end{aligned} \quad (3.4)$$

$$\left| \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I}) \right| \leq c_4 \eta_m^2 \left( \left( \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\| \right)^2 + \left( \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\| \right)^2 \right), \quad (3.5)$$

where  $c_i$  ( $i = 2, 3, 4$ ) are constants and  $m = 0, 1, \dots$

From Lemmas 3.1 and 3.2, we can derive the following lemma.

**Lemma 3.3.** *Suppose Assumptions (A1) and (A2) hold, for  $0 < \eta_0 \leq 1$ , then one has*

$$E(\mathbf{w}^{(m+1)Q}) - E(\mathbf{w}^{mQ}) \leq -\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + c_5 \eta_m^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \quad (3.6)$$

where  $c_5$  is a constant.

With the above Lemmas 3.1–3.3, we can prove the following monotonicity result of OSCG algorithm.

**Theorem 3.4.** *Let  $\{\eta_m\}$  be given by (2.11) and let the weight sequence  $\{\mathbf{w}^{mQ}\}$  be generated by (2.8). Then under Assumption (A1), there are positive numbers  $\tilde{N}$  and  $\tilde{\eta}$  such that for any  $N > \tilde{N}$  and  $0 < \eta_0 < \min\{1, \tilde{\eta}\}$  one has*

$$E(\mathbf{w}^{(m+1)Q}) \leq E(\mathbf{w}^{mQ}), \quad \forall m \geq 0. \quad (3.7)$$

To give the convergence theorem, we also need the following estimation.

**Lemma 3.5.** *Let  $\{\eta_m\}$  be given by (2.11). Then under Assumption (A1), there are the same positive numbers  $\tilde{N}$  and  $\tilde{\eta}$  chosen as Theorem 3.4 such that for any  $N > \tilde{N}$  and  $0 < \eta_0 < \min\{1, \tilde{\eta}\}$  one has*

$$\sum_{m=1}^{\infty} \frac{1}{m} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) < \infty. \quad (3.8)$$

The following lemma gives an estimate of a series, which is essential for the proof of the convergence theorem.

**Lemma 3.6** (see [16]). *Suppose that a series*

$$\sum_{n=1}^{\infty} \frac{a_n^2}{n} \quad (3.9)$$

is convergent and  $a_n \geq 0$ . If there exists a constant  $c_6 > 0$  such that

$$|a_n - a_{n+1}| < \frac{c_6}{n}, \quad (3.10)$$

then

$$\lim_{n \rightarrow \infty} a_n = 0. \quad (3.11)$$

The following lemma will be used to prove the convergence of the weight sequence.

**Lemma 3.7.** *Suppose that the function  $\mathcal{E} : \mathbb{R}^{2L} \rightarrow \mathbb{R}^1$  is continuous and differentiable on a compact set  $\Phi \subset \mathbb{R}^{2L}$  and that  $\Phi_1 = \{\theta \mid \partial \mathcal{E}(\theta) / \partial(\theta) = 0\}$  contains only finite points. If a sequence  $\{\theta^n\}_{n=1}^{\infty} \subset \Phi$  satisfies*

$$\lim_{n \rightarrow \infty} \|\theta^{n+1} - \theta^n\| = 0, \quad \lim_{n \rightarrow \infty} \left\| \frac{\partial \mathcal{E}(\theta^n)}{\partial \theta} \right\| = 0, \quad (3.12)$$

then there exists a point  $\theta^* \in \Phi_1$  such that  $\lim_{n \rightarrow \infty} \theta^n = \theta^*$ .

Now we are ready to give the main convergence theorem.

**Theorem 3.8.** *Let  $\{\eta_m\}$  be given by (2.11) and let the weight sequence  $\{\mathbf{w}^n\}$  be generated by (2.8). Then under Assumption (A1), there are positive numbers  $\tilde{N}$  and  $\tilde{\eta}$  such that for any  $N > \tilde{N}$  and  $0 < \eta_0 < \min\{1, \tilde{\eta}\}$  one has*

$$\lim_{n \rightarrow \infty} \left\| \frac{\partial E(\mathbf{w}^n)}{\partial \mathbf{w}^R} \right\| = 0, \quad \lim_{n \rightarrow \infty} \left\| \frac{\partial E(\mathbf{w}^n)}{\partial \mathbf{w}^I} \right\| = 0. \quad (3.13)$$

Furthermore, if Assumption (A2) also holds, then there exists a point  $\mathbf{w}^* \in \Phi_0$  such that

$$\lim_{n \rightarrow \infty} \mathbf{w}^n = \mathbf{w}^*. \quad (3.14)$$

## 4. Proofs

*Proof of Lemma 3.1.* Using Taylor's formula, we have

$$\begin{aligned} & \mu_{qR}(U^{(m+1)Q,q,R}) - \mu_{qR}(U^{mQ,q,R}) \\ &= \mu'_{qR}(U^{mQ,q,R}) (U^{(m+1)Q,q,R} - U^{mQ,q,R}) + \frac{1}{2} \mu''_{qR}(t_1^{m,q}) (U^{(m+1)Q,q,R} - U^{mQ,q,R})^2 \\ &= \mu'_{qR}(U^{mQ,q,R}) \begin{pmatrix} \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \\ -\mathbf{w}^{(m+1)Q,I} + \mathbf{w}^{mQ,I} \end{pmatrix} \cdot \begin{pmatrix} x^q \\ y^q \end{pmatrix} \\ & \quad + \frac{1}{2} \mu''_{qR}(t_1^{m,q}) \left( \begin{pmatrix} \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \\ -\mathbf{w}^{(m+1)Q,I} + \mathbf{w}^{mQ,I} \end{pmatrix} \cdot \begin{pmatrix} x^q \\ y^q \end{pmatrix} \right)^2, \end{aligned} \quad (4.1)$$

where  $t_1^{m,q}$  lies on the segment between  $U^{(m+1)Q,q,R}$  and  $U^{mQ,q,R}$ . Similarly we also have a point  $t_2^{m,q}$  between  $U^{(m+1)Q,q,I}$  and  $U^{mQ,q,I}$  such that

$$\begin{aligned}
& \mu_{qI}(U^{(m+1)Q,q,I}) - \mu_{qI}(U^{mQ,q,I}) \\
&= \mu'_{qI}(U^{mQ,q,I})(U^{(m+1)Q,q,I} - U^{mQ,q,I}) + \frac{1}{2}\mu''_{qI}(t_2^{m,q})(U^{(m+1)Q,q,I} - U^{mQ,q,I})^2 \\
&= \mu'_{qI}(U^{mQ,q,I}) \left( \frac{\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}}{\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}} \right) \cdot \begin{pmatrix} \mathbf{x}^q \\ \mathbf{y}^q \end{pmatrix} \\
& \quad + \frac{1}{2}\mu''_{qI}(t_2^{m,q}) \left( \left( \frac{\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}}{\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}} \right) \cdot \begin{pmatrix} \mathbf{x}^q \\ \mathbf{y}^q \end{pmatrix} \right)^2.
\end{aligned} \tag{4.2}$$

From (2.8) and (2.10) we have

$$\begin{aligned}
\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} &= \sum_{q=1}^Q \Delta \mathbf{w}^{mQ+q,R} = -\eta_m \sum_{q=1}^Q \mathbf{p}_R^{m,q,q}, \\
\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} &= \sum_{q=1}^Q \Delta \mathbf{w}^{mQ+q,I} = -\eta_m \sum_{q=1}^Q \mathbf{p}_I^{m,q,q}.
\end{aligned} \tag{4.3}$$

Combining (2.4), (2.9), (3.1), and (4.1)–(4.3), then we have

$$\begin{aligned}
& E(\mathbf{w}^{(m+1)Q}) - E(\mathbf{w}^{mQ}) \\
&= \sum_{q=1}^Q \left( \left[ \mu_{qR}(U^{(m+1)Q,q,R}) - \mu_{qR}(U^{mQ,q,R}) \right] + \left[ \mu_{qI}(U^{(m+1)Q,q,I}) - \mu_{qI}(U^{mQ,q,I}) \right] \right) \\
&= \sum_{q=1}^Q \left( \mu'_{qR}(U^{mQ,q,R}) \left[ -\eta_m \sum_{k=1}^Q \mathbf{p}_R^{m,k,k} \cdot \mathbf{x}^q + \eta_m \sum_{k=1}^Q \mathbf{p}_I^{m,k,k} \cdot \mathbf{y}^q \right] \right. \\
& \quad \left. + \mu'_{qI}(U^{mQ,q,I}) \left[ -\eta_m \sum_{k=1}^Q \mathbf{p}_I^{m,k,k} \cdot \mathbf{x}^q - \eta_m \sum_{k=1}^Q \mathbf{p}_R^{m,k,k} \cdot \mathbf{y}^q \right] \right) + \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I}) \\
&= \sum_{q=1}^Q \left( \left[ \mu'_{qR}(U^{mQ,q,R}) \mathbf{x}^q + \mu'_{qI}(U^{mQ,q,I}) \mathbf{y}^q \right] \sum_{k=1}^Q (-\eta_m \mathbf{r}_R^{m,k} - \eta_m \mathbf{p}_R^{m,k}) \right. \\
& \quad \left. + \left[ -\mu'_{qR}(U^{mQ,q,R}) \mathbf{y}^q + \mu'_{qI}(U^{mQ,q,I}) \mathbf{x}^q \right] \sum_{k=1}^Q (-\eta_m \mathbf{r}_I^{m,k} - \eta_m \mathbf{p}_I^{m,k}) \right) \\
& \quad + \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I})
\end{aligned}$$



$$\begin{aligned}
&= -\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
&\quad - \eta_m \left[ \left( \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_R^{m,k} \right) + \left( \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_I^{m,k} \right) \right] \\
&\quad + \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I}),
\end{aligned} \tag{4.4}$$

where

$$\rho_{m,q,R} = \frac{1}{2} \mu''_{qR} (t_1^{m,q}) \left( U^{(m+1)Q,q,R} - U^{mQ,q,R} \right)^2, \quad \rho_{m,q,I} = \frac{1}{2} \mu''_{qI} (t_2^{m,q}) \left( U^{(m+1)Q,q,I} - U^{mQ,q,I} \right)^2. \tag{4.5}$$

□

*Proof of Lemma 3.2.* From (2.5) and Assumption (A1) we know that functions  $\mu_{qR}(t)$ ,  $\mu_{qI}(t)$ ,  $\mu'_{qR}(t)$ ,  $\mu'_{qI}(t)$ ,  $\mu''_{qR}(t)$ , and  $\mu''_{qI}(t)$  ( $1 \leq q \leq Q$ ) are all bounded. Thus there is a constant  $c_7$  such that

$$\max \left\{ \|\mu_{qR}(t)\|, \|\mu_{qI}(t)\|, \|\mu'_{qR}(t)\|, \|\mu'_{qI}(t)\|, \|\mu''_{qR}(t)\|, \|\mu''_{qI}(t)\| \right\} \leq c_7, \quad t \in \mathbb{R}^1, \quad 1 \leq q \leq Q. \tag{4.6}$$

By (2.9), (2.10), (3.1), and the Mean-Value Theorem, for  $2 \leq k \leq Q$  and  $m = 0, 1, \dots$ , we have

$$\begin{aligned}
\|\mathbf{r}_R^{m,k}\| &= \|\mathbf{p}_R^{m,k,k} - \mathbf{p}_R^{m,k}\| \\
&= \|\mu'_{kR} (U^{mQ+k-1,k,R}) \mathbf{x}^k + \mu'_{kI} (U^{mQ+k-1,k,I}) \mathbf{y}^k - \mu'_{kR} (U^{mQ,k,R}) \mathbf{x}^k - \mu'_{kI} (U^{mQ,k,I}) \mathbf{y}^k\| \\
&= \|\mu''_{kR} (t_3^{m,k}) (U^{mQ+k-1,k,R} - U^{mQ,k,R}) \mathbf{x}^k + \mu''_{kI} (t_4^{m,k}) (U^{mQ+k-1,k,I} - U^{mQ,k,I}) \mathbf{y}^k\| \\
&= \|\mu''_{kR} (t_3^{m,k}) [(\mathbf{w}^{mQ+k-1,R} - \mathbf{w}^{mQ,R}) \cdot \mathbf{x}^k - (\mathbf{w}^{mQ+k-1,I} - \mathbf{w}^{mQ,I}) \cdot \mathbf{y}^k] \mathbf{x}^k \\
&\quad + \mu''_{kI} (t_4^{m,k}) [(\mathbf{w}^{mQ+k-1,I} - \mathbf{w}^{mQ,I}) \cdot \mathbf{x}^k + (\mathbf{w}^{mQ+k-1,R} - \mathbf{w}^{mQ,R}) \cdot \mathbf{y}^k] \mathbf{y}^k\| \\
&\leq \tilde{c}_1 \left[ \left\| \sum_{q=1}^{k-1} (-\eta_m \mathbf{r}_R^{m,q} - \eta_m \mathbf{p}_R^{m,q}) \right\| + \left\| \sum_{q=1}^{k-1} (-\eta_m \mathbf{r}_I^{m,q} - \eta_m \mathbf{p}_I^{m,q}) \right\| \right] \\
&\leq \tilde{c}_1 \eta_m \sum_{q=1}^{k-1} \left[ (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + (\|\mathbf{r}_R^{m,q}\| + \|\mathbf{r}_I^{m,q}\|) \right],
\end{aligned} \tag{4.7}$$

where  $\tilde{c}_1 = c_7 \max_{1 \leq q \leq Q} (\|\mathbf{x}^q\|^2 + \|\mathbf{y}^q\|^2)$ . Similarly we have

$$\|\mathbf{r}_I^{m,k}\| \leq \tilde{c}_1 \eta_m \sum_{q=1}^{k-1} \left[ (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + (\|\mathbf{r}_R^{m,q}\| + \|\mathbf{r}_I^{m,q}\|) \right], \quad 2 \leq k \leq Q, \quad m = 0, 1, \dots \quad (4.8)$$

In particular, as  $\|\mathbf{r}_R^{m,1}\| = \|\mathbf{r}_I^{m,1}\| = 0$ , for  $k = 2$ , we can get

$$\max \left\{ \|\mathbf{r}_R^{m,2}\|, \|\mathbf{r}_I^{m,2}\| \right\} \leq \tilde{c}_2 \eta_m \sum_{q=1}^{k-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|), \quad (4.9)$$

where  $\tilde{c}_2 = \tilde{c}_1$ . For  $3 \leq s \leq Q$ ,  $2 \leq k \leq s-1$ , suppose that

$$\|\mathbf{r}_R^{m,k}\| \leq \tilde{c}_s \eta_m \sum_{q=1}^{k-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|), \quad \|\mathbf{r}_I^{m,k}\| \leq \tilde{c}_s \eta_m \sum_{q=1}^{k-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|), \quad (4.10)$$

where  $\tilde{c}_s$  are nonnegative constants. Recalling  $\eta_m \leq 1$ , then we have

$$\begin{aligned} \|\mathbf{r}_R^{m,s}\| &\leq \tilde{c}_1 \eta_m \sum_{q=1}^{s-1} \left[ (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + (\|\mathbf{r}_R^{m,q}\| + \|\mathbf{r}_I^{m,q}\|) \right] \\ &= \tilde{c}_1 \eta_m \left[ \sum_{q=1}^{s-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + \sum_{q=2}^{s-1} \|\mathbf{r}_R^{m,q}\| + \sum_{q=2}^{s-1} \|\mathbf{r}_I^{m,q}\| \right] \\ &\leq \tilde{c}_1 \eta_m \left[ \sum_{q=1}^{s-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + 2 \sum_{q=2}^{s-1} \left( \tilde{c}_q \eta_m \sum_{l=1}^{q-1} (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) \right) \right] \\ &\leq \tilde{c}_1 \eta_m \left[ \sum_{q=1}^{s-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) + 2Q \eta_m \tilde{c}_L \sum_{l=1}^{s-1} (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) \right] \\ &\leq \tilde{c}_s \eta_m \sum_{q=1}^{s-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \end{aligned} \quad (4.11)$$

where  $\tilde{c}_s = 1 + 2Q\tilde{c}_L$  and  $\tilde{c}_L = \max\{\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_{s-1}\}$ . Similarly, we also have

$$\|\mathbf{r}_I^{m,s}\| \leq \tilde{c}_s \eta_m \sum_{q=1}^{s-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|). \quad (4.12)$$

Thus, by setting  $c_2 = \max\{\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_Q\}$ , we have (3.3). Now we begin to prove (3.4). Using (3.3) and Cauchy-Schwartz inequality, we have

$$\begin{aligned}
& \left| \eta_m \left( \left( \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_R^{m,k} \right) + \left( \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right) \cdot \left( \sum_{k=1}^Q \mathbf{r}_I^{m,k} \right) \right) \right| \\
& \leq \eta_m \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\| \times \sum_{k=1}^Q \left( c_2 \eta_m \sum_{l=1}^{k-1} (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) \right) \\
& \quad + \eta_m \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\| \times \sum_{k=1}^Q \left( c_2 \eta_m \sum_{l=1}^{k-1} (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) \right) \\
& \leq c_2 Q \eta_m^2 \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\| \times \sum_{l=1}^Q (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) + c_2 Q \eta_m^2 \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\| \\
& \quad \times \sum_{l=1}^Q (\|\mathbf{p}_R^{m,l}\| + \|\mathbf{p}_I^{m,l}\|) \\
& \leq c_8 \eta_m^2 \left( \left( \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\| \right)^2 + \left( \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\| \right)^2 \right),
\end{aligned} \tag{4.13}$$

where  $c_8 = 2c_2Q$ . This validates (3.4). Finally, we show (3.5). Using (2.10), (3.1), (3.3), and (4.3), we have

$$\begin{aligned}
\|\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}\| &= \eta_m \left\| \sum_{k=1}^Q (\mathbf{r}_R^{m,k} + \mathbf{p}_R^{m,k}) \right\| \\
&\leq \eta_m \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| + c_2 \eta_m^2 \sum_{k=1}^Q \sum_{q=1}^{k-1} (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \\
&\leq \eta_m \sum_{k=1}^Q (\|\mathbf{p}_R^{m,k}\| + \|\mathbf{p}_I^{m,k}\|) + c_2 Q \eta_m^2 \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \\
&\leq c_9 \eta_m \sum_{k=1}^Q (\|\mathbf{p}_R^{m,k}\| + \|\mathbf{p}_I^{m,k}\|),
\end{aligned} \tag{4.14}$$

where  $c_9 = c_2Q + 1$ . Similarly we also have

$$\|\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}\| \leq c_9 \eta_m \sum_{k=1}^Q (\|\mathbf{p}_R^{m,k}\| + \|\mathbf{p}_I^{m,k}\|). \tag{4.15}$$

This together with (2.9) and (4.6) leads to

$$\begin{aligned}
& \left| \sum_{q=1}^Q (\rho_{m,q,R} + \rho_{m,q,I}) \right| \\
&= \frac{1}{2} \left| \sum_{q=1}^Q \left( \mu''_{qR}(t_1^{m,q}) (U^{(m+1)Q,q,R} - U^{mQ,q,R})^2 + \mu''_{qI}(t_2^{m,q}) (U^{(m+1)Q,q,I} - U^{mQ,q,I})^2 \right) \right| \\
&= \frac{1}{2} \left| \sum_{q=1}^Q \left[ \mu''_{qR}(t_1^{m,q}) \left( \begin{pmatrix} \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \\ -\mathbf{w}^{(m+1)Q,I} + \mathbf{w}^{mQ,I} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}^q \\ \mathbf{y}^q \end{pmatrix} \right)^2 \right. \right. \\
&\quad \left. \left. + \mu''_{qI}(t_2^{m,q}) \left( \begin{pmatrix} \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \\ \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}^q \\ \mathbf{y}^q \end{pmatrix} \right)^2 \right] \right| \\
&\leq c_{10} \left( \left\| \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right\|^2 + \left\| \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right\|^2 \right) \\
&\leq 2c_9 c_{10} \eta_m^2 \left( \sum_{k=1}^Q (\left\| \mathbf{p}_R^{m,k} \right\| + \left\| \mathbf{p}_I^{m,k} \right\|) \right)^2 \\
&\leq c_4 \eta_m^2 \left( \left( \sum_{q=1}^Q \left\| \mathbf{p}_R^{m,q} \right\| \right)^2 + \left( \sum_{q=1}^Q \left\| \mathbf{p}_I^{m,q} \right\| \right)^2 \right), \tag{4.16}
\end{aligned}$$

where  $c_{10} = Qc_7 \max_{1 \leq q \leq Q} \{ \|\mathbf{x}^q\|^2 + \|\mathbf{y}^q\|^2 \}$  and  $c_4 = 4c_9 c_{10}$ . This completes the proof.  $\square$

*Proof of Lemma 3.3.* Recalling Lemmas 3.1 and 3.2, we conclude that

$$\begin{aligned}
& E(\mathbf{w}^{(m+1)Q}) - E(\mathbf{w}^{mQ}) \\
&\leq -\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + (c_3 + c_4) \eta_m^2 \left( \left( \sum_{q=1}^Q \left\| \mathbf{p}_R^{m,q} \right\| \right)^2 + \left( \sum_{q=1}^Q \left\| \mathbf{p}_I^{m,q} \right\| \right)^2 \right) \\
&\leq -\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + Q(c_3 + c_4) \eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\|^2 + \left\| \mathbf{p}_I^{m,q} \right\|^2 \right). \tag{4.17}
\end{aligned}$$

Then (3.6) is obtained by letting  $c_5 = Q(c_3 + c_4)$ .  $\square$

*Proof of Theorem 3.4.* In virtue of (3.6), the core to prove this lemma is to verify that

$$\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \geq c_5 \eta_m^2 \sum_{k=1}^Q \left( \left\| \mathbf{p}_R^{m,k} \right\|^2 + \left\| \mathbf{p}_I^{m,k} \right\|^2 \right). \quad (4.18)$$

In the following we will prove (4.18) by induction. First we take  $\eta_0$  such that

$$\eta_0 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{0,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{0,k} \right\|^2 \right) \geq c_5 \eta_0^2 \sum_{k=1}^Q \left( \left\| \mathbf{p}_R^{0,k} \right\|^2 + \left\| \mathbf{p}_I^{0,k} \right\|^2 \right). \quad (4.19)$$

For  $m \geq 0$  suppose that

$$\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \geq c_5 \eta_m^2 \sum_{k=1}^Q \left( \left\| \mathbf{p}_R^{m,k} \right\|^2 + \left\| \mathbf{p}_I^{m,k} \right\|^2 \right). \quad (4.20)$$

Next we will prove that

$$\eta_{m+1} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 \right) \geq c_5 \eta_{m+1}^2 \sum_{k=1}^Q \left( \left\| \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \mathbf{p}_I^{m+1,k} \right\|^2 \right). \quad (4.21)$$

Notice that

$$\begin{aligned} \mathbf{p}_R^{m+1,k} - \mathbf{p}_R^{m,k} &= \left( \mu'_{kR} \left( U^{(m+1)Q,k,R} \right) - \mu'_{kR} \left( U^{mQ,k,R} \right) \right) \mathbf{x}^k + \left( \mu'_{kI} \left( U^{(m+1)Q,k,I} \right) - \mu'_{kI} \left( U^{mQ,k,I} \right) \right) \mathbf{y}^k \\ &= \left( \mu''_{kR} \left( t_5^{m,k} \right) \left( U^{(m+1)Q,k,R} - U^{mQ,k,R} \right) \right) \mathbf{x}^k + \left( \mu''_{kI} \left( t_6^{m,k} \right) \left( U^{(m+1)Q,k,I} - U^{mQ,k,I} \right) \right) \mathbf{y}^k \\ &= \mu''_{kR} \left( t_5^{m,k} \right) \left( \left( \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right) \cdot \mathbf{x}^k - \left( \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right) \cdot \mathbf{y}^k \right) \mathbf{x}^k \\ &\quad + \mu''_{kI} \left( t_6^{m,k} \right) \left( \left( \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right) \cdot \mathbf{x}^k + \left( \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right) \cdot \mathbf{y}^k \right) \mathbf{y}^k, \end{aligned} \quad (4.22)$$

where  $t_5^{m,k}$  lies on the segment between  $U^{(m+1)Q,k,R}$  and  $U^{mQ,k,R}$ , and  $t_6^{m,k}$  lies on the segment between  $U^{(m+1)Q,k,I}$  and  $U^{mQ,k,I}$ . Similar to (4.14), we also have the following estimation:

$$\begin{aligned}
\left\| \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right\| &= \eta_m \left\| \sum_{k=1}^Q \left( \mathbf{r}_R^{m,k} + \mathbf{p}_R^{m,k} \right) \right\| \\
&\leq \eta_m \sum_{k=1}^Q \left\| \mathbf{r}_R^{m,k} \right\| + \eta_m \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| \\
&\leq c_2 \eta_m^2 \sum_{k=1}^Q \sum_{q=1}^{k-1} \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) + \eta_m \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| \\
&\leq c_2 Q \eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) + \eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\| \right) \\
&= \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) + c_{11} \eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right), \\
\left\| \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right\| &\leq \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) + c_{11} \eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right),
\end{aligned} \tag{4.23}$$

where  $c_{11} = c_2 Q$ . By (4.6) and (4.22)-(4.23) we know that there are positive constants  $c_{12}$  and  $c_{13}$  such that

$$\begin{aligned}
\left\| \mathbf{p}_R^{m+1,k} \right\| &\leq \left\| \mathbf{p}_R^{m,k} \right\| + c_{12} \left\| \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right\| + c_{13} \left\| \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right\| \\
&\leq \left\| \mathbf{p}_R^{m,k} \right\| + c_{14} \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) + c_{15} \eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right),
\end{aligned} \tag{4.24}$$

where  $c_{14} = c_{12} + c_{13}$ ,  $c_{15} = c_{11}(c_{12} + c_{13})$ . Taking squares of the two sides of the above inequality gives

$$\begin{aligned}
\left\| \mathbf{p}_R^{m+1,k} \right\|^2 &\leq \left\| \mathbf{p}_R^{m,k} \right\|^2 + c_{14}^2 \eta_m^2 \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 + c_{15}^2 \eta_m^4 \left( \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) \right)^2 \\
&\quad + 2 \left( c_{14} \eta_m \left\| \mathbf{p}_R^{m,k} \right\| \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) \right. \\
&\quad \left. + c_{15} \eta_m^2 \left\| \mathbf{p}_R^{m,k} \right\| \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) \right. \\
&\quad \left. + c_{14} c_{15} \eta_m^3 \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \|\mathbf{p}_R^{m,k}\|^2 + c_{14}^2 \eta_m^2 \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 + c_{15}^2 \eta_m^4 \left( \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \right)^2 \\
&\quad + c_{14} \eta_m \left( \|\mathbf{p}_R^{m,k}\|^2 + \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 \right) \\
&\quad + c_{15} \eta_m^2 \left( \|\mathbf{p}_R^{m,k}\|^2 + \left( \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \right)^2 \right) \\
&\quad + c_{14} c_{15} \eta_m^3 \left( \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 + \left( \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \right)^2 \right) \\
&= \|\mathbf{p}_R^{m,k}\|^2 + (c_{14}^2 \eta_m + c_{14} + c_{14} c_{15} \eta_m^2) \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 \\
&\quad + (c_{14} + c_{15} \eta_m) \eta_m \|\mathbf{p}_R^{m,k}\|^2 + (c_{15}^2 \eta_m^3 + c_{15} \eta_m + c_{14} c_{15} \eta_m^2) \\
&\quad \times \eta_m \left( \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \right)^2.
\end{aligned} \tag{4.25}$$

Now we sum up the above inequality over  $k = 1, \dots, Q$  and obtain

$$\begin{aligned}
\sum_{k=1}^Q \|\mathbf{p}_R^{m+1,k}\|^2 &\leq \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 + Q(c_{14}^2 \eta_m + c_{14} + c_{14} c_{15} \eta_m^2) \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 \\
&\quad + (c_{14} + c_{15} \eta_m) \eta_m \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 \\
&\quad + Q(c_{15}^2 \eta_m^3 + c_{15} \eta_m + c_{14} c_{15} \eta_m^2) \eta_m \left( \sum_{q=1}^Q (\|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\|) \right)^2 \\
&\leq \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 + 2Q(c_{14}^2 \eta_m + c_{14} + c_{14} c_{15} \eta_m^2) \eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right)^2 \\
&\quad + (c_{14} + c_{15} \eta_m) \eta_m \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 \\
&\quad + 2Q(c_{15}^2 \eta_m^3 + c_{15} \eta_m + c_{14} c_{15} \eta_m^2) \eta_m \left( \left( \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\| \right)^2 + \left( \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\| \right)^2 \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 + 2Q(c_{14}^2\eta_m + c_{14} + c_{14}c_{15}\eta_m^2)\eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\|^2 + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\|^2 \right) \\
&\quad + (c_{14} + c_{15}\eta_m)\eta_m \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \\
&\quad + 2Q^2(c_{15}^2\eta_m^3 + c_{15}\eta_m + c_{14}c_{15}\eta_m^2)\eta_m \left( \sum_{q=1}^Q \|\mathbf{p}_R^{m,q}\|^2 + \sum_{q=1}^Q \|\mathbf{p}_I^{m,q}\|^2 \right).
\end{aligned} \tag{4.26}$$

Let

$$c_{16} = \max \left\{ 2Q(c_{14}^2\eta_m + c_{14} + c_{14}c_{15}\eta_m^2), c_{14} + c_{15}\eta_m + 2Q^2(c_{15}^2\eta_m^3 + c_{15}\eta_m + c_{14}c_{15}\eta_m^2) \right\}, \tag{4.27}$$

then

$$\begin{aligned}
\sum_{k=1}^Q \|\mathbf{p}_R^{m+1,k}\|^2 &\leq \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 + c_{16}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
&\quad + c_{16}\eta_m \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right).
\end{aligned} \tag{4.28}$$

On the other hand, from (4.22) we have

$$\begin{aligned}
\sum_{k=1}^Q \mathbf{p}_R^{m+1,k} &= \sum_{k=1}^Q \mathbf{p}_R^{m,k} + \sum_{k=1}^Q \mu_{kR}''(t_5^{m,k}) \left( (\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}) \cdot \mathbf{x}^k - (\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}) \cdot \mathbf{y}^k \right) \mathbf{x}^k \\
&\quad + \sum_{k=1}^Q \mu_{kI}''(t_6^{m,k}) \left( (\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}) \cdot \mathbf{x}^k + (\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}) \cdot \mathbf{y}^k \right) \mathbf{y}^k.
\end{aligned} \tag{4.29}$$

Similar to the deduction of (4.24), from (4.29) we have

$$\begin{aligned}
\left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\| &\geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| - Qc_{12} \|\mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R}\| - Qc_{13} \|\mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I}\| \\
&\geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| - Qc_{14}\eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) - Qc_{15}\eta_m^2 \sum_{q=1}^Q \left( \|\mathbf{p}_R^{m,q}\| + \|\mathbf{p}_I^{m,q}\| \right).
\end{aligned} \tag{4.30}$$



It can be easily verified that, for any positive numbers  $a, b, c$ , if  $a \geq b - c$ ,

$$a^2 \geq b^2 - 2bc. \quad (4.31)$$

Applying (4.31) to (4.30) implies that

$$\begin{aligned} \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 &\geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 - 2 \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| \\ &\quad \times \left( Qc_{14}\eta_m \left( \left\| \sum_{q=1}^Q \mathbf{p}_R^{m,q} \right\| + \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \right) + Qc_{15}\eta_m^2 \sum_{q=1}^Q \left( \left\| \mathbf{p}_R^{m,q} \right\| + \left\| \mathbf{p}_I^{m,q} \right\| \right) \right) \\ &\geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 - 2Qc_{14}\eta_m \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 - 2Qc_{14}\eta_m \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\| \left\| \sum_{q=1}^Q \mathbf{p}_I^{m,q} \right\| \\ &\quad - 2Qc_{15}\eta_m^2 \left( \sum_{k=1}^Q \left\| \mathbf{p}_R^{m,k} \right\| \right)^2 - 2Qc_{15}\eta_m^2 \left( \sum_{k=1}^Q \left\| \mathbf{p}_R^{m,k} \right\| \right) \left( \sum_{q=1}^Q \left\| \mathbf{p}_I^{m,q} \right\| \right) \\ &\geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 - 2Qc_{14}\eta_m \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 - Qc_{14}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\ &\quad - 2Qc_{15}\eta_m^2 \left( \sum_{k=1}^Q \left\| \mathbf{p}_R^{m,k} \right\| \right)^2 - Qc_{15}\eta_m^2 \left( \left( \sum_{k=1}^Q \left\| \mathbf{p}_R^{m,k} \right\| \right)^2 + \left( \sum_{k=1}^Q \left\| \mathbf{p}_I^{m,k} \right\| \right)^2 \right). \end{aligned} \quad (4.32)$$

Similarly, we can obtain the counterpart of (4.28) as

$$\begin{aligned} \sum_{k=1}^Q \left\| \mathbf{p}_I^{m+1,k} \right\|^2 &\leq \sum_{k=1}^Q \left\| \mathbf{p}_I^{m,k} \right\|^2 + c_{16}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\ &\quad + c_{16}\eta_m \sum_{k=1}^Q \left( \left\| \mathbf{p}_R^{m,k} \right\|^2 + \left\| \mathbf{p}_I^{m,k} \right\|^2 \right), \end{aligned} \quad (4.33)$$

and the counterpart of (4.32) as

$$\begin{aligned} \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 &\geq \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 - 2Qc_{14}\eta_m \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 - Qc_{14}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\ &\quad - 2Qc_{15}\eta_m^2 \left( \sum_{k=1}^Q \left\| \mathbf{p}_I^{m,k} \right\| \right)^2 - Qc_{15}\eta_m^2 \left( \left( \sum_{k=1}^Q \left\| \mathbf{p}_R^{m,k} \right\| \right)^2 + \left( \sum_{k=1}^Q \left\| \mathbf{p}_I^{m,k} \right\| \right)^2 \right). \end{aligned} \quad (4.34)$$

From (4.28) and (4.33) we have

$$\begin{aligned}
& \sum_{k=1}^Q \|\mathbf{p}_R^{m+1,k}\|^2 + \sum_{k=1}^Q \|\mathbf{p}_I^{m+1,k}\|^2 \\
& \leq \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\|^2 + \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\|^2 \\
& \quad + 2c_{16}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + 2c_{16}\eta_m \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right).
\end{aligned} \tag{4.35}$$

From (4.32) and (4.34) we have

$$\begin{aligned}
& \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 \\
& \geq \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 - 4Qc_{14}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4Qc_{15}\eta_m^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right).
\end{aligned} \tag{4.36}$$

Using (2.11) and (4.36), we can get

$$\begin{aligned}
& \frac{1}{\eta_{m+1}} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 \right) \\
& \geq \frac{1}{\eta_m} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + N \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4Qc_{14} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) - 4NQc_{14}\eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4Qc_{15}\eta_m \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& \quad - 4NQc_{15}\eta_m^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right).
\end{aligned} \tag{4.37}$$

Multiplying (4.37) with  $\eta_{m+1}^2$  gives

$$\begin{aligned}
& \eta_{m+1} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 \right) \\
& \geq \frac{\eta_{m+1}^2}{\eta_m} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + N\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4Qc_{14}\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4NQc_{14}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad - 4Qc_{15}\eta_m\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& \quad - 4NQc_{15}\eta_m^2\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right).
\end{aligned} \tag{4.38}$$

Using (4.20) and (4.35), we obtain

$$\begin{aligned}
& c_5\eta_{m+1}^2 \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m+1,k}\|^2 + \sum_{k=1}^Q \|\mathbf{p}_I^{m+1,k}\|^2 \right) \\
& \leq \frac{\eta_{m+1}^2}{\eta_m} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) + 2c_5c_{16}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& \quad + 2c_5c_{16}\eta_m\eta_{m+1}^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right).
\end{aligned} \tag{4.39}$$

Combining (4.38) and (4.39) we have

$$\begin{aligned}
& \eta_{m+1} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m+1,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m+1,k} \right\|^2 \right) \\
& \geq c_5\eta_{m+1}^2 \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m+1,k}\|^2 + \sum_{k=1}^Q \|\mathbf{p}_I^{m+1,k}\|^2 \right) \\
& \quad + N\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right)
\end{aligned}$$

$$\begin{aligned}
& -4Qc_{14}\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& -4NQc_{14}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& -4Qc_{15}\eta_m\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& -4NQc_{15}\eta_m^2\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& -2c_5c_{16}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& -2c_5c_{16}\eta_m\eta_{m+1}^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right).
\end{aligned} \tag{4.40}$$

Thus to validate (4.21) we only need to prove the following inequality:

$$\begin{aligned}
N\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) & \geq 4Qc_{14}\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& + 4NQc_{14}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& + 4Qc_{15}\eta_m\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& + 4NQc_{15}\eta_m^2\eta_{m+1}^2 \left( \left( \sum_{k=1}^Q \|\mathbf{p}_R^{m,k}\| \right)^2 + \left( \sum_{k=1}^Q \|\mathbf{p}_I^{m,k}\| \right)^2 \right) \\
& + 2c_5c_{16}\eta_m\eta_{m+1}^2 \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) \\
& + 2c_5c_{16}\eta_m\eta_{m+1}^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right).
\end{aligned} \tag{4.41}$$

Recalling (4.20) and  $\eta_m \leq \eta_0$  ( $m \geq 0$ ), it is easy to see that  $(N/6)\eta_{m+1}^2 (\|\sum_{k=1}^Q \mathbf{p}_R^{m+1,k}\|^2 + \|\sum_{k=1}^Q \mathbf{p}_I^{m+1,k}\|^2) \geq$  each term of (4.41) can be assured for  $N \geq \tilde{N}$  and  $0 \leq \eta_0 \leq \min\{1, \tilde{\eta}\}$  by setting

$$\begin{aligned} \tilde{N} &= \max \left\{ 24Qc_{14}, 24Q \frac{c_{15}}{c_5}, 12c_{16} \right\}, \\ \tilde{\eta} &= \min \left\{ \frac{2}{3}Qc_{14}, \frac{2c_{16}}{3c_{15}}, \frac{c_5 c_{16} \tilde{N}}{12}, \frac{\|\sum_{k=1}^Q \mathbf{p}_R^{0,k}\|^2 + \|\sum_{k=1}^Q \mathbf{p}_I^{0,k}\|^2}{c_5 \sum_{k=1}^Q \|\mathbf{p}_R^{0,k}\|^2 + \|\mathbf{p}_I^{0,k}\|^2} \right\}. \end{aligned} \quad (4.42)$$

This, thus, validates (4.41). As a result, (4.18) and (3.7) are proved.  $\square$

*Proof of Lemma 3.5.* From Lemma 3.3 we have

$$E(\mathbf{w}^{mQ}) - E(\mathbf{w}^{(m+1)Q}) \geq \eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) - c_5 \eta_m^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right). \quad (4.43)$$

Sum the above inequalities up over  $m = 1, \dots, M$ , then

$$\begin{aligned} E(\mathbf{w}^Q) - E(\mathbf{w}^{(M+1)Q}) \\ \geq \sum_{m=1}^M \left( \eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) - c_5 \eta_m^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \right). \end{aligned} \quad (4.44)$$

Note that  $E(\mathbf{w}^{(M+1)Q}) \geq 0$  for  $M \geq 0$ . Setting  $M \rightarrow \infty$ , we have

$$\sum_{m=1}^{\infty} \left( \eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) - c_5 \eta_m^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \right) \leq E(\mathbf{w}^Q) < \infty. \quad (4.45)$$

Using (2.9) and (4.6), we can find a constant  $c_{17}$  such that

$$\sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \leq c_{17}. \quad (4.46)$$

This together with (2.13) leads to

$$\sum_{m=1}^{\infty} \left( c_5 \eta_m^2 \sum_{k=1}^Q \left( \|\mathbf{p}_R^{m,k}\|^2 + \|\mathbf{p}_I^{m,k}\|^2 \right) \right) \leq \sum_{m=1}^{\infty} \frac{c_5 c_{17}}{N^2} \frac{1}{m^2} < \infty. \quad (4.47)$$

Thus from (4.45) and (4.47) it holds that

$$\sum_{m=1}^{\infty} \eta_m \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) < \infty. \quad (4.48)$$

Recalling  $\eta_m = O(1/m)$  from (2.12) gives

$$\sum_{m=1}^{\infty} \frac{1}{m} \left( \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 + \left\| \sum_{k=1}^Q \mathbf{p}_I^{m,k} \right\|^2 \right) < \infty. \quad (4.49)$$

□

*Proof of Lemma 3.6.* This lemma is the same as Lemma 2.1 of [16]. □

*Proof of Lemma 3.7.* This result is almost the same as Theorem 14.1.5 in [18], and the details of the proof are omitted. □

*Proof of Theorem 3.8.* Using (2.9), (4.6), (4.14), and (4.15), we can find a constant  $c_{18}$  such that

$$\left\| \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right\| \leq c_{18} \eta_m, \quad \left\| \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right\| \leq c_{18} \eta_m. \quad (4.50)$$

From (2.6) and (4.22) we have

$$\begin{aligned} & \frac{\partial E(\mathbf{w}^{(m+1)Q})}{\partial \mathbf{w}^R} - \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} \\ &= \sum_{k=1}^Q \left( \left[ \mu'_{kR} \left( U^{(m+1)Q,k,R} \right) - \mu'_{kR} \left( U^{mQ,k,R} \right) \right] \mathbf{x}^k + \left[ \mu'_{kI} \left( U^{(m+1)Q,k,I} \right) - \mu'_{kI} \left( U^{mQ,k,I} \right) \right] \mathbf{y}^k \right) \\ &= \sum_{k=1}^Q \left( \mu''_{kR} \left( t_5^{m,k} \right) \left( \left( \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right) \cdot \mathbf{x}^k - \left( \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right) \cdot \mathbf{y}^k \right) \mathbf{x}^k \right. \\ & \quad \left. + \mu''_{kI} \left( t_6^{m,k} \right) \left( \left( \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right) \cdot \mathbf{x}^k + \left( \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right) \cdot \mathbf{y}^k \right) \mathbf{y}^k \right), \end{aligned} \quad (4.51)$$

where  $t_5^{m,k}$  and  $t_6^{m,k}$  are defined in (4.22). Thus, from (4.6), (4.50), and Cauchy-Schwartz inequality, there exists a constant  $c_{19}$  such that for any vector  $\mathbf{e} \in \mathbb{R}^L$

$$\begin{aligned}
& \left\| \frac{\partial E(\mathbf{w}^{(m+1)Q})}{\partial \mathbf{w}^R} \cdot \mathbf{e} \right\| - \left\| \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} \cdot \mathbf{e} \right\| \\
& \leq \left| \frac{\partial E(\mathbf{w}^{(m+1)Q})}{\partial \mathbf{w}^R} \cdot \mathbf{e} - \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} \cdot \mathbf{e} \right| \\
& \leq c_{19} \|\mathbf{e}\| \left( \left\| \mathbf{w}^{(m+1)Q,R} - \mathbf{w}^{mQ,R} \right\| + \left\| \mathbf{w}^{(m+1)Q,I} - \mathbf{w}^{mQ,I} \right\| \right) \\
& \leq 2c_{18}c_{19} \|\mathbf{e}\| \eta_m \\
& \leq \frac{2c_{18}c_{19} \|\mathbf{e}\|}{N} \frac{1}{m}.
\end{aligned} \tag{4.52}$$

Using (2.6), (2.9), and Lemma 3.5, we have

$$\sum_{m=1}^{\infty} \left( \frac{1}{m} \left| \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} \cdot \mathbf{e} \right|^2 \right) = \sum_{m=1}^{\infty} \left( \frac{1}{m} \left| \left( \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right) \cdot \mathbf{e} \right|^2 \right) \leq \|\mathbf{e}\|^2 \sum_{m=1}^{\infty} \left( \frac{1}{m} \left\| \sum_{k=1}^Q \mathbf{p}_R^{m,k} \right\|^2 \right) < \infty. \tag{4.53}$$

From (4.52), (4.53), and Lemma 3.6 it holds that

$$\lim_{m \rightarrow \infty} \left| \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} \cdot \mathbf{e} \right| = 0. \tag{4.54}$$

Since  $\mathbf{e}$  is arbitrary in  $\mathbb{R}^L$ , we have

$$\lim_{m \rightarrow \infty} \frac{\partial E(\mathbf{w}^{mQ})}{\partial \mathbf{w}^R} = 0. \tag{4.55}$$

Therefore, when  $q = 0$ , we complete the proof of  $\lim_{m \rightarrow \infty} (\partial E(\mathbf{w}^{mQ+q}) / \partial \mathbf{w}^R) = 0$ , and we can similarly show that  $\lim_{m \rightarrow \infty} (\partial E(\mathbf{w}^{mQ+q}) / \partial \mathbf{w}^R) = 0$  for  $q = 1, \dots, Q$ . Thus, we have shown that

$$\lim_{n \rightarrow \infty} \frac{\partial E(\mathbf{w}^n)}{\partial \mathbf{w}^R} = 0. \tag{4.56}$$

In a similar way, we can also prove that

$$\lim_{n \rightarrow \infty} \frac{\partial E(\mathbf{w}^n)}{\partial \mathbf{w}^I} = 0. \tag{4.57}$$

Thus, (3.13) is obtained from (4.56) and (4.57).

Next we begin to prove (3.14). Using (2.10), we have

$$\left\| \Delta \mathbf{w}^{mQ+q,R} \right\| = \eta_m \left\| \mathbf{p}_R^{m,q,q} \right\|. \quad (4.58)$$

Similar to (4.46), we know that  $\left\| \mathbf{p}_R^{m,q,q} \right\|$  is bounded. Recalling (2.13) makes us conclude that

$$\lim_{m \rightarrow \infty} \left\| \mathbf{w}^{mQ+q,R} - \mathbf{w}^{mQ+q-1,R} \right\| = \lim_{m \rightarrow \infty} \left\| \Delta \mathbf{w}^{mQ+q,R} \right\| = \lim_{m \rightarrow \infty} \eta_m \left\| \mathbf{p}_R^{m,q,q} \right\| = 0, \quad q = 1, \dots, Q, \quad (4.59)$$

which implies that

$$\lim_{n \rightarrow \infty} \left\| \mathbf{w}^{n,R} - \mathbf{w}^{n-1,R} \right\| = 0. \quad (4.60)$$

Similarly, we have

$$\lim_{n \rightarrow \infty} \left\| \mathbf{w}^{n,I} - \mathbf{w}^{n-1,I} \right\| = 0. \quad (4.61)$$

Write

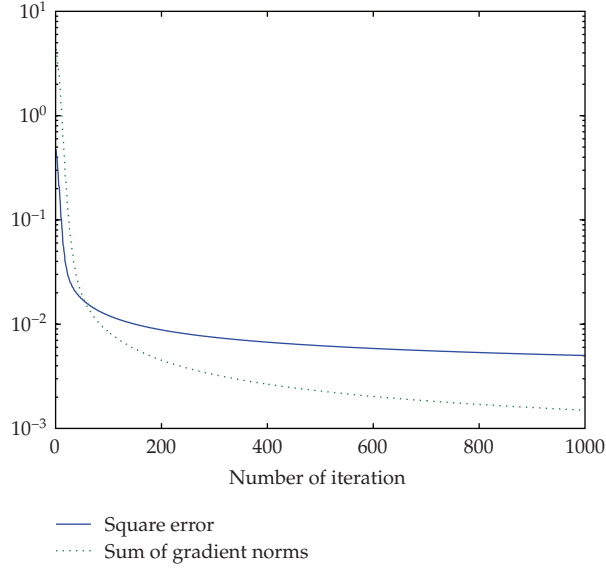
$$\boldsymbol{\theta} = \left( \left( \mathbf{w}^R \right)^T, \left( \mathbf{w}^I \right)^T \right)^T, \quad \boldsymbol{\theta}^n = \left( \left( \mathbf{w}^{n,R} \right)^T, \left( \mathbf{w}^{n,I} \right)^T \right)^T, \quad (4.62)$$

then the square error function  $E(\mathbf{w})$  can be looked as a real-valued function  $\mathcal{E}(\boldsymbol{\theta})$ . Thus from (4.56), (4.57), (4.60), and (4.61) we have

$$\lim_{n \rightarrow \infty} \frac{\partial \mathcal{E}(\boldsymbol{\theta}^n)}{\partial \boldsymbol{\theta}} = 0, \quad \lim_{n \rightarrow \infty} \left\| \boldsymbol{\theta}^n - \boldsymbol{\theta}^{n-1} \right\| = 0. \quad (4.63)$$

Furthermore, from Assumption (A2) we know that the set  $\{\boldsymbol{\theta} \mid \partial \mathcal{E}(\boldsymbol{\theta}) / \partial(\boldsymbol{\theta}) = 0\}$  contains only finite points. Thus, the sequence  $\{\boldsymbol{\theta}^n\}_{n=1}^{\infty}$  here satisfies all the conditions needed in Lemma 3.7. As a result, there is a  $\boldsymbol{\theta}^*$  which satisfies that  $\lim_{n \rightarrow \infty} \boldsymbol{\theta}^n = \boldsymbol{\theta}^*$ . Since  $\boldsymbol{\theta}^n$  consists of the real and imaginary parts of  $\mathbf{w}^n$ , we know that there is a  $\mathbf{w}^*$  such that  $\lim_{n \rightarrow \infty} \mathbf{w}^n = \mathbf{w}^*$ . We, thus, complete the proof of (3.14).  $\square$





**Figure 1:** Convergence behavior of OSCG algorithm for solving XOR problem (sum of gradient norms =  $\|\partial E(\mathbf{w}^n)/\partial \mathbf{w}^R\| + \|\partial E(\mathbf{w}^n)/\partial \mathbf{w}^I\|$ ).

## 5. Numerical Example

In this section we illustrate the convergence behavior of the OSCG algorithm by using a simple numerical example. The well-known XOR problem is a benchmark in literature of neural networks. As in [13], the training samples of the encoded XOR problem for CVNN are presented as follows:

$$\begin{aligned} \left\{ \mathbf{z}^1 = -1 - i, d^1 = 1 \right\}, & \quad \left\{ \mathbf{z}^2 = -1 + i, d^2 = 0 \right\}, \\ \left\{ \mathbf{z}^3 = 1 - i, d^3 = 1 + i \right\}, & \quad \left\{ \mathbf{z}^4 = 1 + i, d^4 = i \right\}. \end{aligned} \quad (5.1)$$

This example uses a network with two input nodes (including a bias node) and one output node. The transfer function is  $\text{tansig}(\cdot)$  in MATLAB, which is a commonly used sigmoid function. The parameter  $\eta_0$  is set to be 0.1 and  $N$  is set to be 1. We carry out the test with the initial components of the weights stochastically chosen in  $[-0.5, 0.5]$ . Figure 1 shows that the gradients tend to zero and the square error decreases monotonically as the number of iteration increases and at last tends to a constant. This supports our theoretical analysis.

## 6. Conclusion

In this paper we investigate some convergence properties of an OSCG training algorithm for two-layered CVNN. We choose an adaptive learning rate in the algorithm. Under the condition that the activation function and its up to the second-order derivative are bounded, it is proved that the error function is monotonely decreasing during the training process.

With this result, we further prove that the gradient of the error function tends to zero and the weight sequence tends to a fixed point. A numerical example is given to support our theoretical analysis. We mention that those results are interestingly similar to the convergence results of batch split-complex gradient training algorithm for CVNN given in [17]. Thus our results can also be a theoretical explanation for the relationship between the OSCG algorithm and the batch split-complex algorithm. The convergence results in this paper can be generalized to a more general case, that is, multilayer CVNN.

## Acknowledgments

The authors wish to thank the associate editor and the anonymous reviewers for their helpful comments and valuable suggestions regarding this paper. This work is supported by the National Natural Science Foundation of China (70971014).

## References

- [1] S. Alonso-Quesada and M. De La Sen, "Robust adaptive control with multiple estimation models for stabilization of a class of non-inversely stable time-varying plants," *Asian Journal of Control*, vol. 6, no. 1, pp. 59–73, 2004.
- [2] P. Zhao and O. P. Malik, "Design of an adaptive PSS based on recurrent adaptive control theory," *IEEE Transactions on Energy Conversion*, vol. 24, no. 4, pp. 884–892, 2009.
- [3] J. Cortès, "Distributed Kriged Kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.
- [4] D. R. Wilson and T. R. Martinez, "The general inefficiency of batch training for gradient descent learning," *Neural Networks*, vol. 16, no. 10, pp. 1429–1451, 2003.
- [5] J. Li, Y. Diao, M. Li, and X. Yin, "Stability analysis of discrete Hopfield neural networks with the nonnegative definite monotone increasing weight function matrix," *Discrete Dynamics in Nature and Society*, vol. 2009, Article ID 673548, 10 pages, 2009.
- [6] Q. Zhu and J. Cao, "Stochastic stability of neural networks with both Markovian jump parameters and continuously distributed delays," *Discrete Dynamics in Nature and Society*, vol. 2009, Article ID 490515, 20 pages, 2009.
- [7] G. M. George and C. Koutsougeras, "Complex domain backpropagation," *IEEE Transactions on Circuits and Systems II*, vol. 39, no. 5, pp. 330–334, 1992.
- [8] T. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 967–969, 1992.
- [9] A. Hirose, *Complex-Valued Neural Networks*, Springer, New York, NY, USA, 2006.
- [10] T. Kim and T. Adali, "Fully complex backpropagation for constant envelop signal processing," in *Proceedings of the IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing X*, pp. 231–240, Sydney, Australia, December 2000.
- [11] A. I. Hanna and D. P. Mandic, "A data-reusing nonlinear gradient descent algorithm for a class of complex-valued neural adaptive filters," *Neural Processing Letters*, vol. 17, no. 1, pp. 85–91, 2003.
- [12] S. L. Goh and D. P. Mandic, "Stochastic gradient-adaptive complex-valued nonlinear neural adaptive filters with a gradient-adaptive step size," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1511–1516, 2007.
- [13] T. Nitta, "Orthogonality of decision boundaries in complex-valued neural networks," *Neural Computation*, vol. 16, no. 1, pp. 73–97, 2004.
- [14] S. S. Yang, S. Siu, and C. L. Ho, "Analysis of the initial values in split-complex backpropagation algorithm," *IEEE Transactions on Neural Networks*, vol. 19, no. 9, pp. 564–573, 2008.
- [15] H. Zhang, W. Wu, F. Liu, and M. Yao, "Boundedness and convergence of online gradient method with penalty for feedforward neural networks," *IEEE Transactions on Neural Networks*, vol. 20, no. 6, pp. 1050–1054, 2009.
- [16] W. Wu, G. Feng, Z. Li, and Y. Xu, "Deterministic convergence of an online gradient method for BP neural networks," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 533–540, 2005.

- [17] H. Zhang, C. Zhang, and W. Wu, "Convergence of batch split-complex backpropagation algorithm for complex-valued neural networks," *Discrete Dynamics in Nature and Society*, vol. 2009, Article ID 329173, 16 pages, 2009.
- [18] J. Ortega and W. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY, USA, 1970.