9

Adaptive Optical-Phase Equalizer

In the framework of the complex-valued neural networks dealing with phase values adaptively, we can realize various adaptive subsystems required in optical communications such as a learning phase equalizer. Modern optical communications attains a high degree of development mainly in trunk lines. Moreover, near-future networks provide subscribers with high-speed and multichannel information transmission over all-optical routers and switches. Thereby, we have to compensate the fiber dispersion varying with successively switched optical routes. The dispersion variation is very large since the high-speed multichannel optical communications occupies a wide frequency bandwidth. The optical-phase equalizer to be presented in this chapter can be one of the principles useful in such applications. As an example, we consider a system with supervised learning here.

9.1 System Construction

Figure 9.1 shows the neuron, the basic element, constructing the system. Its main characteristic is the multiple connections between m-th single input and n-th single neurons. Each connection possesses its own time delay and transparency. The connections altogether generate interference dependent on the optical carrier frequency.

We can directly relate the amplitude and phase of lightwave to those of signals in complex-valued neural networks. Then, to modulate phase in parallel two-dimensionally, we use a parallel-aligned liquid-crystal spatial light modulator (PAL-SLM), as we did in Chapter 8. To modulate amplitude, on the other hand, we use a PAL-SLM or ordinary (polarization-type) spatial light modulator (SLM) in combination with a polarizer. We express the input signal as $x_m \equiv |x_m| \exp(i\alpha_m)$, output signal as $y_n \equiv |y_n| \exp(i\beta_n)$, and connection weight as $w_{nm,h} \equiv |w_{nm,h}| \exp(i2\pi f \tau_{nm,h})$, respectively, where m and n are indices for input and output vectors, while h is that for multiple connections mentioned above.

A. Hirose: Complex-Valued Neural Networks, SCI 400, pp. 143–149.
springerlink.com

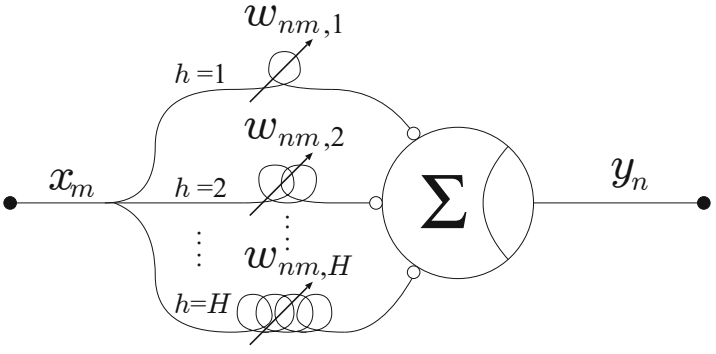


Fig. 9.1 Element constructing the adaptive phase equalizer. (Reprinted from Fig.1 in [209]: Sotaro Kawata and Akira Hirose: A coherent optical neural network that learns desirable phase values in frequency domain by using multiple optical-path differences, *Opt. Lett.,* 28(24):2524–2526, 2003, with permission.)

The purpose of the network is to generate a desirable phase value against input phase by adaptive processing. The task is classified into the function approximation described in Section 4.4. We realize function approximation in a frequency-dependent manner by the method explained in Section 4.4.5. The processing conducted by the neuron is expressed as

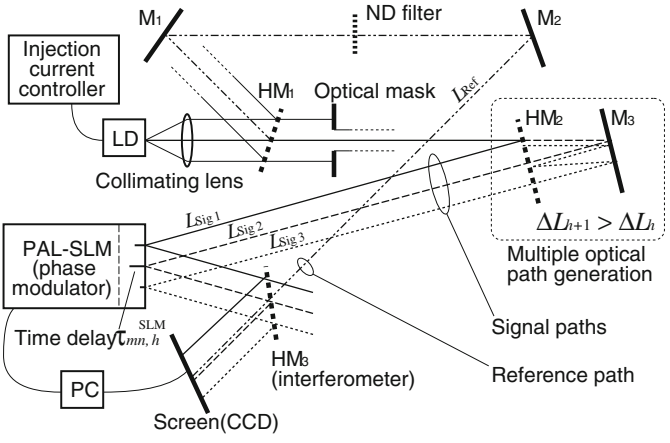
$$y_n = g\left(\sum_m \sum_h \left(|w_{nm,h}| \exp(i2\pi f \tau_{nm,h})\ x_m\right)\right) \tag{9.1}$$

$$g(u) \equiv A\ \tanh(B\ |u|)\ \exp(i \arg(u)) \tag{9.2}$$

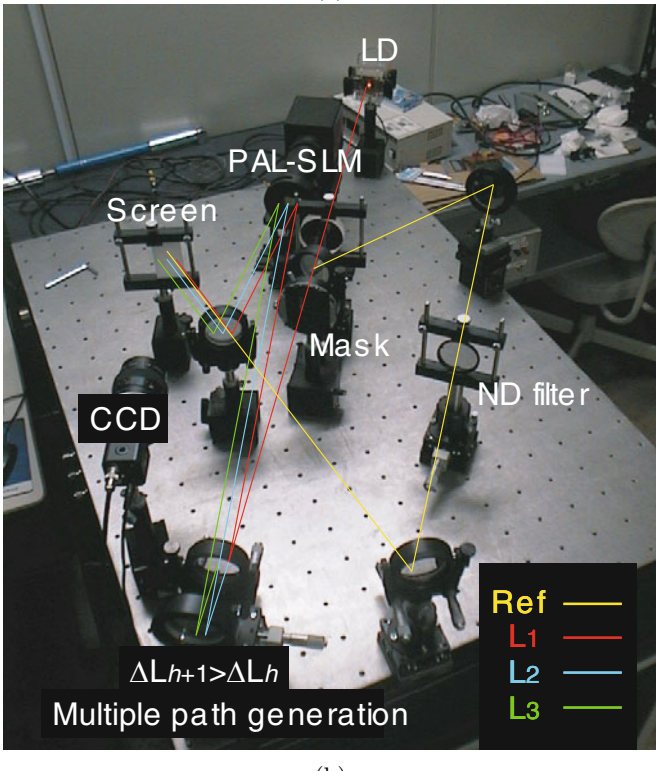
All the weighted inputs are summed in the complex domain to yield the internal state u. The activation function $g(u)$ is an amplitude-phase-type function defined in (9.2), A is saturation amplitude in output signals, and B is amplitude gain.

9.2 Optical Setup

We consider one of the simplest examples, i.e., an optical circuit consisting of a single neuron [209],[192]. Though the construction is simple, the function is fulfilling. We implement the neuron shown in Fig.9.1 as an optical circuit illustrated in Fig.9.2(a). Figure 9.2(b) is a photograph of the optical setup, where a three-armed optical interferometer is constructed as a set of self-homodyne circuits. We examine the phase of the circuit output by self-homodyning. We realize a frequency modulation by employing a semiconductor laser diode. The basis of the optical circuit is the same as that in the associative memory presented in Chapter 8.



(a)



(b)

Fig. 9.2 (a)Construction [209] and (b)photograph of the optical setup. (Reprinted from Fig.2 in [209] in figure caption of Fig.9.1 with permission.)

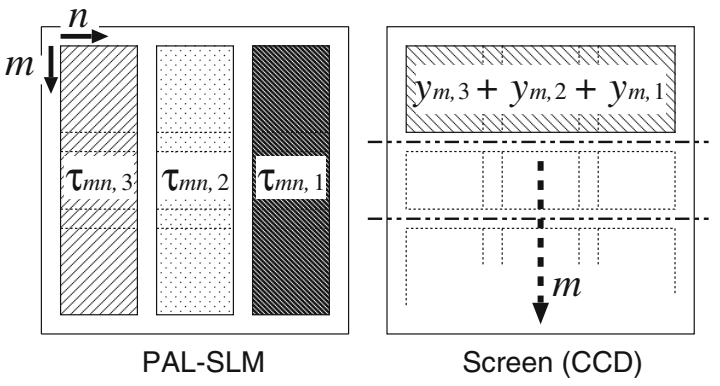


Fig. 9.3 Assignments of signals on the PAL-SLM and the CCD. (Reprinted from Fig.3 in [209] in figure caption of Fig.9.1 with permission.)

The delay length for input signals of the neuron is physically represented by the difference between the signal optical-path length $L_{\rm Sigh}$ and the reference length $L_{\rm Ref}$, i.e., $\Delta L_h \equiv L_{\rm Sigh} - L_{\rm Ref}$, where we assume $\Delta L_1 < \Delta L_2 < \Delta L_3$. The delay time $\tau_{nm,h}$ of the connection weight $w_{nm,h}$ in total is the sum of the delay time of the SLM modulation, shown in Fig.9.3, and the delay in the optical-path difference, i.e., $\tau^{\rm SLM}_{nm,h} + \Delta L_h/c$. Here we assume $|w_{nm,h}|$=1 for simplicity. The optical sum and the nonlinearity in $g(\cdot)$ is realized by the optical detector (CCD: Charge Coupled Device) that detects interference result as shown in Fig.9.3.

9.3 Dynamics of Output Phase-Value Learning

We synthesize a desired output signal (phase) by adjusting the connection weights $w_{nm,h}$ in learning. As we discussed in Chapter 4, we have two frameworks in supervised learning in single-layered complex-valued neural networks, i.e., the complex-valued steepest-descent method and the complex-valued correlation learning superficially identical to the complex-valued Hebbian rule. Here we employ the delay-time learning, (4.97), which is the frequency-dependent version of the complex-valued correlation learning mentioned in Section 4.3.8. That is, we have

$$\tau \frac{d\tau_{nm,h}}{dt} = \frac{1}{2\pi f} \frac{|y_n||x_m|}{|w_{nm,h}|} \sin(\beta_n - \alpha_m - 2\pi f \tau_{nm,h}) \tag{9.3}$$

where τ without index is the time constant determining the learning speed.

The learning procedure is described as follows. First, we determine an arbitrary frequency $\hat{f}$. Then we present signal sets $(\hat{f}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ to the system, where $\hat{\boldsymbol{y}}$ is the output to be learned for an input $\hat{\boldsymbol{x}}$. We update the connection

weights $w_{nm,h}$ according to (9.3) for all the connections. We repeat the above process for all the frequency–signal sets $\hat{f}$, $\hat{\boldsymbol{x}}$, and $\hat{\boldsymbol{y}}$ to be learned.

We choose initial weight values $w_{nm,h}$ such that the transparency is $|w_{nm,h}|=1$ and that the delay time is $\tau_{nm,h}=\theta_0/(2\pi f_0) + \Delta L_h/c$ where θ_0 is random value in the range of $0 \sim 2\pi$, where f_0 is the optical center frequency (the frequency without modulation). As the learning proceeds, the system learns to generate the desired output $\hat{\boldsymbol{y}}$ for input $\hat{\boldsymbol{x}}$ when the carrier frequency is $f = \hat{f}$. In the case that $f \neq \hat{f}$, the output can be different.

9.4 Performance of Phase Equalization

The setup for the optical experiment is shown in Fig.9.2. First, we confirm the learning dynamics in simulation using parameters in the optical experiments to be conducted. Then we proceed to an actual experiment. We prepare various desirable output phase values at four carrier-frequency points from 472.002 THz to 472.008 THz with an interval of 2GHz, and make the system learn the values. The upper limit of the learning iteration has been chosen at 200. We investigate the generalization characteristics in the frequency domain. We also define an error function as (9.4), and examine its evolution, i.e.,

$$E \equiv \frac{1}{2} \sum_{\mu} |\boldsymbol{y}(\hat{\boldsymbol{x}}_\mu) - \hat{\boldsymbol{y}}_\mu|^2 \tag{9.4}$$

where μ is the index for memorized vectors. In the following simulation and experiment, the learning is effective not only for phase but also for amplitude. However, with the present purpose of phase equalization, we investigate the output phase values in particular.

Figure 9.4 shows typical results of simulation (broken curves) and optical experiment (crosses). Closed circles in Fig.9.4 show the teacher signals that the system should learn. Figure 9.4(a) shows random output phase values before learning. However, for the output values at 80 iterations in Fig.9.4(b) and those at 200 iterations in Fig.9.4(c), we find that the system learns gradually the desirable outputs. Figure 9.4(d) presents the reduction of the error function. The almost monotonic decrease reveals the practical effectiveness of the learning dynamics.

By combining the processing elements shown here, we will obtain more complex phase curves in the frequency domain. In combination with the scaling in the frequency sensitivity, we can realize various types of dispersion compensation, besides wavelength-selecting routers. The scaling is actualized by preparing appropriate rough delays ΔL. The frequency-dependent learning will ultimately lead to novel parallelism based on frequency-domain multiplexing in neural networks utilizing the vast frequency bandwidth in lightwave.

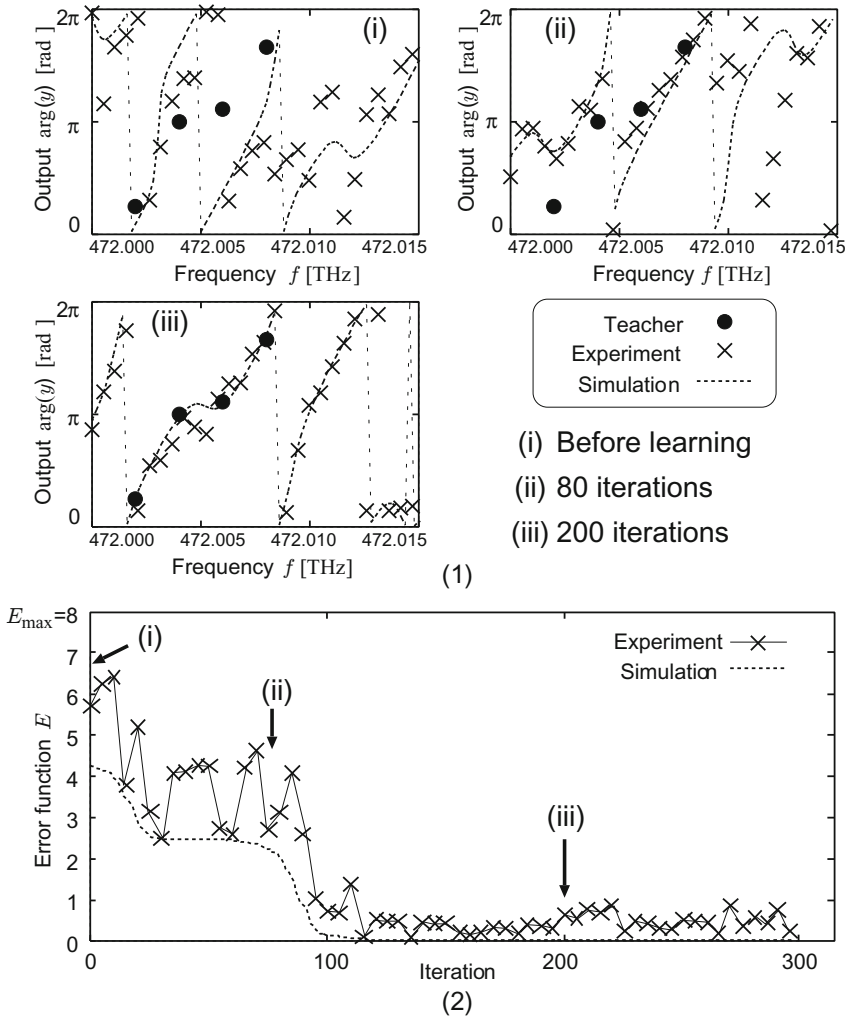


Fig. 9.4 Results of the optical experiment. Output phase values (a)before learning, (b)at 80 learning iterations, and (c)at 200 learning iterations. (d)Change in the error function. (Reprinted from Fig.5 in [209] in figure caption of Fig.9.1 with permission.)

Note that, in the present experiment mentioned here, we used a two-dimensional spatial light modulator (SLM) and other bulky optics. Therefore, the system is not so small. However, the principle of the learning and self-organization is widely applicable so that this type of system can be constructed based on, for instance, optical waveguides and photonic crystals. With such devices, we can design micro-optical circuits with high tolerance to mechanical turbulence.

9.5 Summary

In this chapter, we described a lightwave phase equalizer as a communications-network application of a coherent neural network. The learning dynamics and characteristics were presented. There are many other application areas in optical communications. One example is an adaptive recognition and classification of optical binary phase shift keying (BPSK) labels in photonic label routing for high-speed optical networks [210]. Such learning networks that treat the phase of waves adaptively are directly applicable also to sonic, ultrasonic, and other wave-related systems. They are also expected to pioneer novel future quantum devices that are highly adaptive.