Backpropagating neurons from bichromatic interaction with a three-level system

David Kagan and Harry Friedmann

Optical implementation of a backpropagating neuron by means of a nonlinear Fabry-Perot etalon requires thresholding a forward signal beam while the transmittance of a backpropagating beam is multiplied by the differential of the forward signal. This is achievable by inputting a bichromatic field to a three-level system in an optical cavity. The response characteristics of this device have the added possibility of adaptability of the threshold by the backward probe input intensity.

I. Introduction

Considerable interest has been shown recently in optical implementation of neural network models for associative memory applications.[1] One of the most interesting properties of a neural network is the ability to learn dynamically the interconnections that correspond to a desired behavior through an iterative adaptation of the weight matrices. One of the methods for learning or training the neural system involves error driven adaptation. Error driven behavioral modification has the property that it can sense the system performance and adapt the weights in a manner which will compensate for device imperfections that cause the unwanted behavior. In Ref. 2 a multilayer optical learning network is described that uses self-aligning volume holograms to interconnect nonlinear elements acting as individual neurons. The basic learning algorithm is now reviewed; more details can be found in the above-mentioned reference.

A set of training patterns is presented repetitively in a random sequence to the network. Initially the output patterns differ from the desired response. An error pattern is formed, optically, by taking the difference between the output and the desired response. This error pattern is sent backward through the output neurons into the system to correct the weights of the interconnection matrices. For most problems of interest it is necessary to adapt complex nonlinear decision surfaces. This is done by stacking single layer networks to form a multilayer system which can learn to classify more complex patterns. The layers are connected holographically and adaptive training is performed by the backpropagating error signal

through the system's various layers. The backpropagating algorithm requires that the transmission function of the error signal be proportional to the derivative of the forward signal sigmoid transfer function. The derivative will be peaked where the nonlinear sigmoid forward transfer function has large differential gain. If the neuron is operating at this level the connecting weights will be strongly modified by the error signal and the neuron is forced to a decision. This last condition requires a specific nonlinear optical gate.

Nonlinear optical gates are usually Fabry-Perot etalons containing some optically sensitive material. The nonlinear transfer characteristics are easily obtained but a backward signal which is the derivative of the forward mode presents a problem. In Ref. 2 this was solved by using two closely spaced cavities both addressed by the same forward and backward beams. In this case one cavity is optimized to produce a sigmoid response to the forward beam while blocking the backward signal, the other being resonant to the error signal. The response of the cavity which is resonant to the backpropagating beam is dependent on the forward propagating beam, which tunes the cavity right across the resonance. If the foward signal is of low intensity, the cavity is detuned from resonance. As the foward signal intensity increases the cavity is tuned into resonance. When the forward signal saturates, the cavity is detuned once more so that its transmission decreases. This method has some obvious difficulties. The foward signal is not allowed to resonate in the cavity resonant to the backpropagating error beam. This being the case, a very intense signal will be required to tune the cavity response right through the resonance (linearly scan the Airy pattern). A further disadvantage of this system is the requirement of two cavities per neuron, which add volume and expense to the system. In this paper we suggest a nonlinear cavity which will give a forward nonlinear signal and an error probe which approximates the desired derivative extremely well. The gate will also have an adaptive threshold which is a distinct advantage for learning procedures. The nonlinear Fabry-
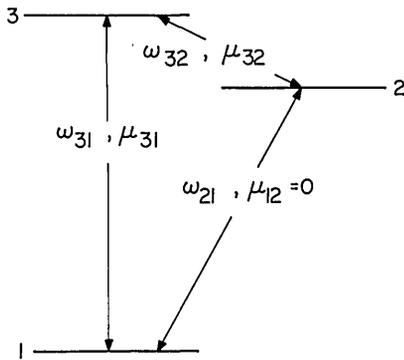
Fig. 1. Three level system interacting with a bichromatic field. One field ω_1 tuned to the 1–3 resonance, the other ω_2 to the 3–2 resonance.

Perot contains a three-level system irradiated by two beams, one acting as the forward signal, and the other, of a different frequency, as the backward error correcting probe.

The layout of the paper is as follows. Section II presents the model of the three-level system in an optical cavity, and coupled input–output equations are derived. In Sec. III the input–output equations are solved and the desired response is shown to result.

II. Model

Consider the system, shown in Fig. 1, which consists of three energy levels. The system is irradiated by two fields, E_1 and E_2, given by

$$E(t,z) = E_1 \exp(i\omega_1 t) + E_2 \exp(i\omega_2 t), \qquad (1)$$

where ω_1 is close to the resonant frequency for the 1–3 transition and ω_2 is tuned to the resonant frequency for the 2–3 transition. Using the density matrix formalism, the time development of the system can be written as

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} [\rho, H], \qquad (2)$$

where ρ is the density operator, H is the Hamiltonian, and the square brackets represent the commutator. After making the rotating wave approximation and writing Eq. (2) explicitly, we then write the optical Bloch equations for this system as follows[3]:

$$\frac{\partial \rho_{12}}{\partial t} = \frac{iE(t)}{\hbar} [\mu_{13}\rho_{32} - \mu_{32}\rho_{13}] - \rho_{12}(\gamma - i\omega_{21}), \qquad (3)$$

$$\frac{\partial \rho_{13}}{\partial t} = \frac{-iE(t)}{\hbar} [\mu_{13}r_{13} + \mu_{23}\rho_{12}] - \rho_{13}(\gamma - i\omega_{31}), \qquad (4)$$

$$\frac{\partial \rho_{32}}{\partial t} = \frac{-iE(t)}{\hbar} [\mu_{32}r_{32} - \mu_{31}\rho_{12}] - \rho_{32}(\gamma - i\omega_{32}), \qquad (5)$$

$$\frac{\partial r_{13}}{\partial t} = \frac{iE(t)}{\hbar} [2(\mu_{13}\rho_{31} - \mu_{31}\rho_{13})$$
$$- (\mu_{32}\rho_{23} - \mu_{23}\rho_{32})] - (r_{13} - r_{13}^{eq})\gamma, \qquad (6)$$

$$\frac{\partial r_{32}}{\partial t} = \frac{iE(t)}{\hbar} [(\mu_{31}\rho_{13} - \mu_{13}\rho_{31})$$
$$+ 2(\mu_{32}\rho_{23} - \mu_{23}\rho_{32})] - (r_{32} - r_{32}^{eq})\gamma, \qquad (7)$$

where r_{13} and r_{32} are the population differences given by $r_{13} = \rho_{11} - \rho_{33}$ and $r_{32} = \rho_{33} - \rho_{22}$. The superscript eq represents equilibrium, and so refers to the equilibrium population differences. In these equations μ_{ij} represents the dipole moment connecting level i to j and ω_{ij} is given by

$$\omega_{ij} = \frac{(\varepsilon_i - \varepsilon_j)}{\hbar},$$

where ε_i is the energy of level i. The population decay rate and the phase decay rate are taken as equal and have been written as γ. The above equations have been solved for steady state in Refs. 3 and 4 and, for exactly resonant fields, the optical susceptibilities for the two frequencies ω_1 and ω_2 are given by

$$\chi(\omega_1) = -\frac{2\alpha_{oa}}{k_1} \frac{4 + X_1^2 + 4X_2^2}{(1 + X_1^2 + X_2^2)(4 + X_1^2 + X_2^2)}, \qquad (8)$$

$$\chi(\omega_2) = \frac{2\alpha_{oe}}{k_2} \frac{3X_1^2}{(1 + X_2^2 + X_1^2)(4 + X_2^2 + X_1^2)}. \qquad (9)$$

In the above equations we have scaled the fields by the square root of the saturation intensities for the respective transitions so that we can write

$$X_1 = \frac{|E_1||\mu_{13}|}{\hbar\gamma\sqrt{2}}, \qquad X_2 = \frac{|E_2||\mu_{32}|}{\hbar\gamma\sqrt{2}},$$

$$\alpha_{oa} = \frac{N}{V} \frac{\mu_{13}^2}{\gamma} \frac{k_1}{2\varepsilon_0\hbar}, \qquad \alpha_{oe} = \frac{N}{V} \frac{\mu_{32}^2}{\gamma} \frac{k_2}{2\varepsilon_0\hbar},$$

where k_1 and k_2 are the magnitudes of the wave vectors at frequencies ω_1 and ω_2. Following Ref. 5 and making the mean field assumption, the cavity input–output steady state equations are written as

$$Y_1 = X_1 \left[1 + 2\frac{C_1}{F} (4 + X_1^2 + 4X_2^2) \right], \qquad (10)$$

$$Y_2 = X_2 \left[1 + 2\frac{C_2}{F} (3X_1^2) \right], \qquad (11)$$

where F is given by

$$F = (1 + X_1^2 + X_2^2)(4 + X_1^2 + X_2^2), \qquad (12)$$

where Y_1 and Y_2 are the scaled input fields given by

$$Y_1 = |E_{1in}| \frac{|\mu_{13}|}{\hbar\gamma\sqrt{2}}, \qquad Y_2 = |E_{2in}| \frac{|\mu_{32}|}{\hbar\gamma\sqrt{2}}.$$

Here C_1 and C_2 are given by

$$C_1 = \frac{\alpha_{oa}LR_1}{2(1 - R_1)}, \qquad C_2 = \frac{-\alpha_{oe}LR_2}{2(1 - R_2)},$$

where R_i is the reflection coefficient for the cavity dielectric mirrors at frequency ω_i, L is the length of the cavity, α_{oa} and α_{oe} are the absorption and emission coefficients as previously defined. The factor C is basically equal to the absorption or gain per pass through the cavity divided by the transmission coefficient. A large C_1 implies more absorption (until saturation), and a large $|C_2|$ leads to increased gain of the input with frequency ω_2. To make the mean field
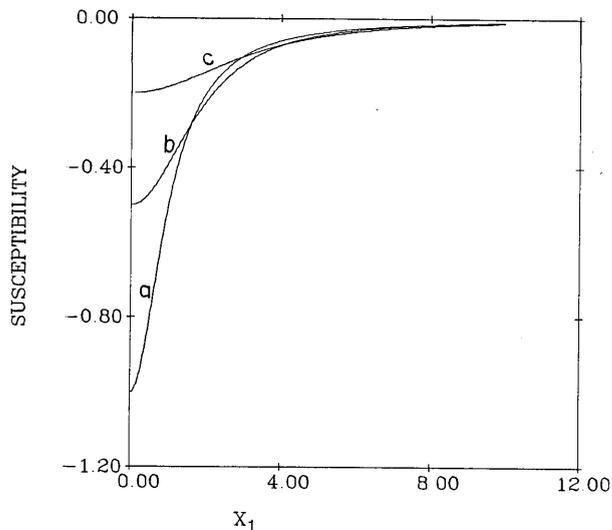
Fig. 2. On-resonance susceptibility of ω_1 as a function of X_1 for different values of X_2: a, $X_2 = 0.0$; b, $X_2 = 2.0$; c, $X_2 = 4.0$. As X_1 increases the susceptibility saturates.
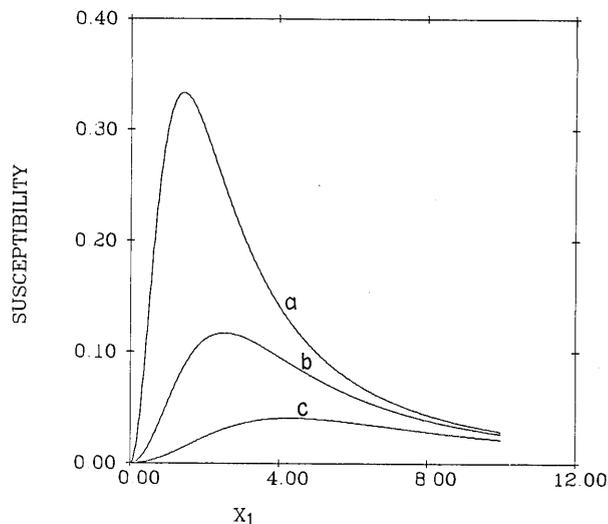

Fig. 3. On-resonance susceptibility of ω_2 as a function of X_1 for different values of X_2: a, $X_2 = 0.0$; b, $X_2 = 2.0$; c, $X_2 = 4.0$. As X_1 increases the susceptibility first reaches a maximum then falls to zero because of detuning caused by splitting of level 3.

assumption the gain or absorption per pass through the cavity should not be large. Since our model is only intended to give qualitative results this assumption has been made for simplicity.

It is instructive at this point to plot the on-resonance absorption and emission as given by Eqs. (8) and (9) as functions of the scaled fields X_1 and X_2. This is shown in Figs. 2 and 3. In Fig. 2 one notes that raising the pump intensity X_1 at first increases the susceptibility of frequency ω_1 until it saturates. For frequency ω_2, increasing X_1 increases its susceptibility and thus its gain, but as X_1 reaches saturating intensities the emission decreases as can be seen in Fig. 3. The reason for this decrease of the gain results from the splitting of level 3 by the pump such that the X_2 beam is now detuned from the 3–2 resonance. It is this property that immediately suggests the use of this system to obtain the desired response for the backpropagating neuron. As the field X_1 increases the emitted field first rises and then falls such that the maximum can be expected at the point of highest gain for X_1. The nonlinear response is obtained by putting the whole system in a cavity as usual. The cavity response is described in the following section.

III. Input–Output Cavity Response

Using Eqs. (10) and (11) the cavity response to the two beams can be evaluated. In Fig. 4 three curves have been plotted: a, the forward signal output X_1 vs input Y_1; b, the backward signal transmission X_2/Y_2; and c, the differential of the forward signal dX_1/dY_1. Let us now consider the application of this cavity as an optical neuron. If the forward propagating signal has the frequency ω_1, output X_1 has a typical nonlinear response to the input, as in curve a in Fig. 4. If an error signal is detected and sent back to the system at frequency ω_2, as input Y_2, transmission X_2/Y_2, shown as b, is then proportional to the differential of the forward
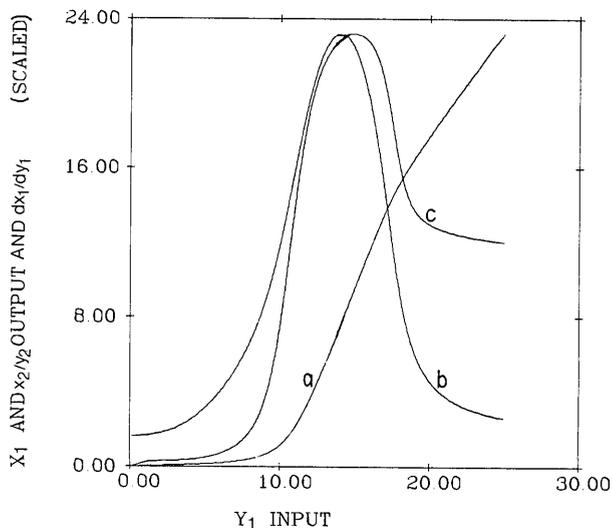

Fig. 4. Three signals: a, X_1 vs Y_1 the forward output response; b, the backward propagating signal transmission response, scaled approximately ($\times 1.67$), X_2/Y_2 vs Y_1; and c, the gain of the forward signal dX_1/dY_1. The parameter values are $C_1 = 20$, $C_2 = -35$, and $Y_2 = 0.5$.

propagating signal c. It should be noted that the output of the forward signal X_1 is not a sigmoidal nonlinearity, the gate being transparent at saturating intensities. This is generally the case for the optical etalons suggested for implementation of nonlinear gates. When a true sigmoidal output in which the transmission levels off at high intensities is required, an optical limiter is used in conjunction with the gate. In the system we propose, the backward transmission is proportional to the differential of an ideal sigmoidal output, falling to zero as the intensity increases. In other words, the backward signal has a zero response after the forward signal has passed the decision threshold, and the weights need not be changed.
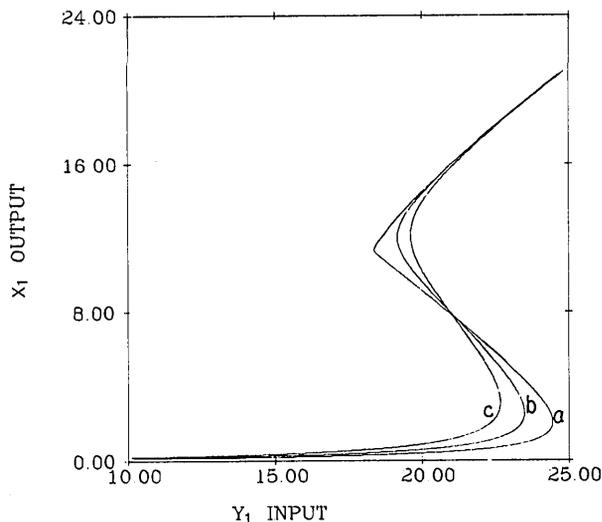
Fig. 5. Optical bistability response curves of X_1 vs Y_1 for varying values of Y_2: a, $Y_2 = 0.01$; b, $Y_2 = 0.5$; c, $Y_2 = 1.0$. As Y_2 increases the switch-up point is shifted to the left. The parameter values are $C_1 = 40$ and $C_2 = -22$.

Fig. 6. Optical bistability response curves of X_2 vs Y_1 for varying values of Y_2: a, $Y_2 = 0.01$; b, $Y_2 = 0.5$; c, $Y_2 = 1.0$. As Y_2 increases the switch-up point is shifted to the left. The parameter values are $C_1 = 40$ and $C_2 = -22$.
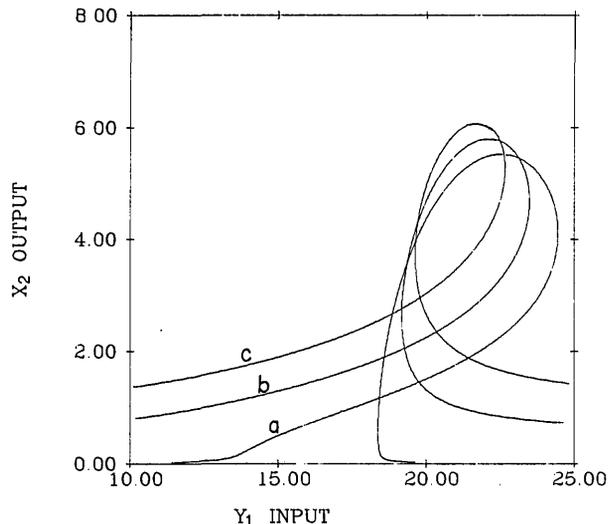
The three-level system also displays the possibility of tunability. If the control beam intensity Y_2 is charged, the output response will be shifted somewhat and the width and slope of the high gain will be changed. This is more clearly seen when parameters which give optical bistable output of X_1 are chosen. In Figs. 5 and 6 these outputs are plotted for different values of the control Y_2 input. We note that the tuning is accomplished by low power inputs of the control beam since both beams are allowed to resonate in the cavity. The idea of using a control beam to shift the width and slope of the output curves can be utilized to some advantage. Specifically, if one wishes to correct errors by simply moving the threshold, the control beam Y_2 could act as an input in the forward direction. Other applications of the tunability of optically bistable outputs have been presented elsewhere.[6-8] The case of tunable optical bistability was also pointed out in Ref. 9, but there the control was of the same order of intensity as the main output, since it was not allowed to resonate in the cavity.

IV. Conclusion

The three-level system analyzed here shows the nonlinear optical properties required for backpropagating neurons as explained. The reason for the decrease of the gain of the frequency tuned to the 3–2 transition is the splitting of level 3 by the intense saturating pump. In a cavity, a resonant mode of the transition from the split level to level 2 might be excited if the gain is large. A difficulty in implementing this neuron is the requirement of generating an error signal with a different frequency. This would have to be done using a nonlinear interaction, such as a phase conjugate frequency shifter, at the output level. Although we recognize this limitation, it does not seem too much of a price to pay for an all-optical neuron which by its physical properties implements the backpropagating algorithm. It should also be pointed out that whatever system one is using, the backpropagating signal must be distinguishable from the forward beam.

The three-level system also shows the possibility of tunable optical bistability and adaptive thresholding by low power control beams, which could well be of use in optical computing.

References

1. D. Psaltis and N. Farhat, "Optical Information Processing Based on an Associative-Memory Model of Neural Nets with Thresholding and Feedback," Opt. Lett. **10**, 98 (1985).
2. K. Wagner and D. Psaltis, "Multilayer Optical Learning Networks," Appl. Opt. **26**, 5061 (1987).
3. R. L. Panock and R. J. Temkin, "Interaction of Two Laser Fields with a Three-Level Molecular System," IEEE J. Quantum Electron. **QE-13**, 425 (1977).
4. R. G. Brewer and E. L. Hahn, "Coherent Two-Photon Processes: Transient and Steady-State Cases," Phys. Rev. A **11**, 1641 (1975).
5. G. P. Agrawal and H. J. Carmichael, "Optical Bistability Through Nonlinear Dispersion and Absorption," Phys. Rev. A **19**, 2074 (1979).
6. D. Kagan and H. Friedmann, "Optical Logic Elements from Four-Level Systems Interacting with Two Fields," Proc. Soc. Photo-Opt. Instrum. Eng. **700**, 71 (1986).
7. D. Kagan and H. Friedmann, "Adaptive Optical Logic and Switching Devices," Proc. Soc. Photo-Opt. Instrum. Eng. **835**, 354 (1987).
8. D. Kagan and H. Friedmann, "Tunable Optical Bistability of Bichromatic Fields Interacting with Four-Level Systems," IEEE J. Quantum Electron. **QE-25**, (June 1989), to be published.
9. N. M. Lawandi and W. S. Rabinovich, "Absorptive Bistability in a Three-Level System Interacting with Two Fields," IEEE J. Quantum Electron. **QE-20**, 458 (1984).