S. Oh, D.C. Park , R.J. Marks II and L.E. Atlas, "Nondispersive propagation skew in iterative neural networks and optical feedback processors", Optical Engineering, vol.28, pp.526-532 (1989)

Nondispersive propagation skew in iterative neural network and optical feedback processors

Seho Oh
Dong Chul Park
Robert J. Marks II
Les E. Atlas
University of Washington
Interactive Systems Design Laboratory
Mail Stop FT-10
Seattle, Washington 98195

Abstract. Although certain iterative optical processors promise algorithmic convergence at the speed of light, little attention is normally given to the consequences of different path lengths required within the processor on the processor performance. The resulting clock skew can have significant degrading effects on the predicted accuracy, stability, and speed of the processor. A similar problem occurs in iterative asynchronous artificial neural networks when, for example, the time delay between two neurons is proportional to their physical separation. In this paper, we show that in the absence of temporal dispersion, certain iterative algorithms have stable steady-state solutions that are independent of clock skew. Examples include stable linear feedback and feedback using soft (slowly varying) nonlinearities. Both are special cases of using a contractive operation in the feedback path. Such processing algorithms can have stable steady-state solutions that are independent of clock skew. Feedback using hard nonlinearities, on the other hand, can result in either an oscillatory or a steady-state solution that depends on the clock skew.

Subject terms: optical signal processing; neural networks; propagation skew; optical feedback.

Optical Engineering 28(5), 526-532 (May 1989).

CONTENTS

Invited Paper IA-108 received Nov. 21, 1988; revised manuscript received Dec. 12, 1988; accepted for publication Feb. 10, 1989. Portions of this paper were published in S. Oh, L. E. Atlas, R. J. Marks II, and D. C. Park, "Effects of clock skew in iterative neural network and optical feedback processors," in *Proc. IEEE International Conference on Neural Networks,* Vol. II, pp. 437–442 (1988).

1. INTRODUCTION

A number of analog[1–5] and discrete[6–12] optical processors have been proposed that use feedback. Shamir[13] has noted that in such systems, the time required for feedback can vary significantly due to the variation of optical path lengths.[14] In certain cases, disregarding this *clock skew* in processor analysis can lead to either unstable or drastically different implementation results. A similar problem occurs in iterative asynchronous artificial neural networks when the communication time delay between two neurons is proportional to their physical separation.

Our analysis is restricted to temporally nondispersive systems. For such systems, a temporal impulse stimulus at any input coordinate can appear later only as a single temporal impulse at any specified output coordinate. Thus, for each input/output coordinate pair, there exists a single temporal delay. If this delay varies from coordinate pair to coordinate pair, the system is skewed.

In this paper, we show that in certain feedback algorithms, temporally nondispersive clock skew does not affect the stability or the steady-state solution of the processor. When an iterative algorithm uses a (possibly nonlinear) contractive operation in the feedback path, the resulting steady-state solution is shown to be unaffected by clock skew. Clock skew is shown, however, to have an effect on systems such as Hopfield artificial neural networks[15,16] when hard nonlinearities are used in the feedback path.

2. PRELIMINARIES

In this section, we develop a general description for temporally nondispersive clock skew in a feedback processor and then show specific instances in which that model can be used to determine whether the steady-state solution is affected by the skew. We consider only a discrete model, although the concepts can be applied to analog processors.

Let a field of N states, $\{s_n | 1 \leq n \leq N\}$, be altered by feedback in a temporally nondispersive skewed processor. Let ϑ_n denote the instantaneous operator that maps the previous states into the current n*th* state at time t. We can then write

$$s_n(t) = \vartheta_n[\{s_m(t - \tau_{nm}) | 1 \leq m \leq N\}] \ , \quad 1 \leq n \leq N \ , \tag{1}$$

where τ_{nm} is the clock skew corresponding to the time required for the state s_m to make a contribution to the state s_n. If we let $t \rightarrow \infty$ and assume a stable steady state, then Eq. (1) becomes

$$s_n(\infty) = \vartheta_n[\{s_m(\infty) | 1 \leq m \leq N\}] \ , \quad 1 \leq n \leq N \ . \tag{2}$$

Although not explicitly noted, this steady state may depend on the clock skew. If, however, Eq. (2) has but a single solution for all $s_n(\infty)$, then the clock skew has no effect on the steady-state solution.

Some specific instances of such processors are now given.

3. SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS

Let $\mathbf{s}_n$ denote a vector of states at time n, $\mathbf{f}$ a like-dimensioned forcing vector, and $\mathbf{A}$ a square matrix. The linear difference equation

$$\mathbf{s}(n + 1) = \mathbf{A}\mathbf{s}(n) + \mathbf{f} \tag{3}$$

is known to converge to the steady-state solution

$$\mathbf{s}(\infty) = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{f} \tag{4}$$

if the spectral radius (magnitude of the maximum eigenvalue) of $\mathbf{A}$ (denoted $\|\mathbf{A}\|$) does not exceed one.[17] The effects of clock skew on convergence and stability of this iteration are now considered.

3.1. Convergence

With reference to Eq. (1), if performed on a skewed processor, Eq. (3) would be implemented as*

$$s_n(t) = \sum_m a_{nm} s_m(t - \tau_{nm}) + f_n(t) \ , \quad 1 \leq n \leq N \ , \tag{5}$$

where f_n is the n*th* element of $\mathbf{f}$ and $f_n(t) \rightarrow f_n$ allows explicitly for input rise time. Letting $t \rightarrow \infty$ and assuming a stable result gives

$$s_n(\infty) = \sum_m a_{nm} s_m(\infty) + f_n \ , \quad 1 \leq n \leq N \ , \tag{6}$$

or equivalently, in matrix-vector form,

$$\mathbf{s}(\infty) = \mathbf{A}\mathbf{s}(\infty) + \mathbf{f} \ . \tag{7}$$

If $(\mathbf{I} - \mathbf{A})$ is not singular, then the solution to this equation is unique and is given by Eq. (4). Clock skew therefore does not affect the solution. The *alternating projection neural network* when interpreted either homogeneously[18,19] or in layered form[20,21] from the hidden to output layer is a special case of this example.

3.2. Stability

The above analysis is conditioned on the stability of the skewed iterations. By letting $n \rightarrow \infty$, for example, we might predict that the iteration $x(n) = 2x(n-1) + 1$ would converge to $x(\infty) = -1$. The difference equation, however, is clearly unstable and $x(\infty) = \pm\infty$ if $x(0) \neq -1$. From the viewpoint of z-transform analysis, the pole of this difference equation lies outside the unit circle.

A sufficient condition for stability of skewed iteration is given by the following:

Lemma 1—Let $\mathbf{A} = (a_{ij})$ denote a square matrix of complex numbers. Define $\mathbf{A}(s) = [a_{ij}\exp(-s\tau_{ij})]$, where $s = \sigma + j\omega$. If $\|\mathbf{A}(s)\| < 1$ for $\mathrm{Re}(s) \geq 0$, then Eq. (5) converges to Eq. (4).

A proof is given in Sec. 9.1. Note that as a special case, we conclude that an iteration without skew converges if $\|\mathbf{A}\| < 1$ since

$$\|\mathbf{A}(s)\| = |\exp(-s\tau)| \ \|\mathbf{A}\| \leq \|\mathbf{A}\| < 1 \ ,$$

where $\tau_{nm} = \tau$ for all (n,m). Two important results built on this lemma follow:

Lemma 2—let $\mathbf{B} = (|a_{ij}|)$. If $\|\mathbf{B}\| < 1$, then Eq. (5) converges stably to Eq. (4) for any $\tau_{ij} \geq 0$.

Therefore, varying the phase terms in the matrix does not affect the convergence stability if the zero phase iteration is stable.

Lemma 3—If $\tau_{ij} = u_i + v_j$ and $\|\mathbf{A}\| < 1$, then Eq. (5) converges stably to Eq. (4) for any $u_i \geq 0$ and $v_j \geq 0$.

Proofs of lemmas 2 and 3 are in Secs. 9.2 and 9.3, respectively. The last lemma will be applied to optical feedback systems later in the paper.

4. CONTRACTIVE OPERATORS

In this section, we explore a more general criterion for which clock skew does not affect steady-state results. We may write Eq. (2) in vector form as

$$\mathbf{s}(\infty) = \boldsymbol{\vartheta}\mathbf{s}(\infty) \ . \tag{8}$$

If the vector operator $\boldsymbol{\vartheta}$ is a *contractive* operator, then

*All summations (Σ) in this paper are from 1 to N.

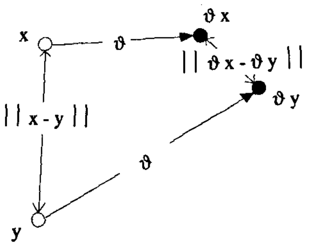


Fig. 1. Geometrical illustration of a contractive operator. After the contractive operation, the signals are closer together (solid dots) than originally (hollow dots.)

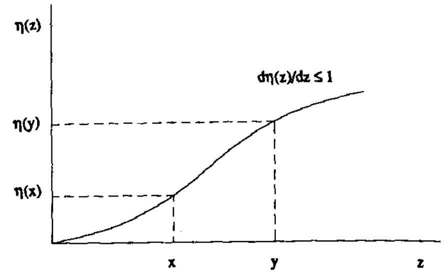


Fig. 2. Example of a soft nonlinearity. Any interval on the z axis maps to a smaller interval.

$$\|\vartheta \mathbf{x} - \vartheta \mathbf{y}\| \leq r\|\mathbf{x} - \mathbf{y}\| , \tag{9}$$

where the norm for a vector **a** is defined by

$$\|\mathbf{a}\|^2 = \mathbf{a}^T \mathbf{a}$$

and $0 \leq r < 1$. If $0 \leq r \leq 1$, then ϑ is said to be *nonexpansive*. The reason for the terminology is evident from the geometry in Fig. 1. Operating on two signals, **x** and **y**, by the operator ϑ results in two signals closer together (contractive) or at least not as far apart (nonexpansive).

4.1. Convergence

If ϑ is contractive, then Eq. (8) has a unique solution[17–19,22] and there is no contribution of clock skew to the steady-state result. When ϑ is nonexpansive, Eq. (8) can have a number of solutions.

Example 1: The linear iteration discussed in the previous section is a special case of a contractive mapping since, from Eq. (6),

$$\vartheta \mathbf{s}(\infty) = \mathbf{A}\mathbf{s}(\infty) + \mathbf{f} .$$

The operator is contractive if

$$\|(\mathbf{A}\mathbf{x} + \mathbf{f}) - (\mathbf{A}\mathbf{y} + \mathbf{f})\| = \|\mathbf{A}(\mathbf{x} - \mathbf{y})\| \leq r\|\mathbf{x} - \mathbf{y}\| .$$

This is clearly true if the spectral radius of **A** does not exceed one.

Example 2: We can nonlinearly generalize Eq. (5) to

$$s_n(t) = \eta_n \left[\sum_m a_{nm} s_m(t - \tau_{nm}) + g_n(t) \right] + h_n(t) , \quad 1 \leq n \leq N , \tag{10}$$

where $f_n(t) \rightarrow f_n$ and $g_n(t) \rightarrow g_n$ are forcing functions. Assuming stability, the steady-state solution in vector form is

$$\mathbf{s}(\infty) = \eta[\mathbf{A}\mathbf{s}(\infty) + \mathbf{g}] + \mathbf{h} , \tag{11}$$

where η is a pointwise nonlinear vector operator; i.e., if $\mathbf{w} = \eta\mathbf{z}$, then the *nth* element of **w** is equal to $\eta_n(z_n)$, where η_n is a given function. In the parlance of neural networks, η_n could be referred to as a *sigmoid* operator.[23,24] Using Eq. (9), the corresponding operator is contractive if

$$\|\eta(\mathbf{A}\mathbf{x} + \mathbf{f}) - \eta(\mathbf{A}\mathbf{y} + \mathbf{f})\| \leq r\|\mathbf{x} - \mathbf{y}\| . \tag{12}$$

We show that the operator is contractive if the spectral radius of **A** does not exceed one and η contains *soft nonlinearities*. That is,

$$\left| \frac{d\eta_n(z)}{dz} \right| \leq 1 , \quad 1 \leq n \leq N$$

for all z. As illustrated in Fig. 2, this constraint has the property that

$$|\eta_n x - \eta_n y| \leq |x - y| .$$

As a result,

$$\|\eta(\mathbf{A}\mathbf{x} + \mathbf{g}) - \eta(\mathbf{A}\mathbf{y} + \mathbf{g})\| \leq \|(\mathbf{A}\mathbf{x} + \mathbf{g}) - (\mathbf{A}\mathbf{y} + \mathbf{g})\| .$$

Using the results of the previous example, the operator corresponding to Eq. (11) is therefore contractive and clock skew has no effect on the final result.

This unique convergence constraint can be generalized to the requirement that for all z,

$$\|\mathbf{A}\| \left| \frac{d\eta_n(z)}{dz} \right| < 1 , \quad 1 \leq n \leq N . \tag{13}$$

4.2. Stability

The following lemma establishes a sufficient condition for stability of the skewed operation in Eq (10).

Lemma 4—For a given matrix **A** and time delays $\{\tau_{nm}\}$, if Eq. (5) converges for every $f_n(t)$ and η is nonexpansive, then Eq. (10) is stable.

A proof is given in Sec. 9.4.

5. HARD NONLINEARITIES

Clock skew can be a factor when implementing an iterative algorithm with hard nonlinearities. Consider the following example of Hopfield's content addressable memory neural network.[15,23,25–27]

Example 1: From the three library vectors

$$\mathbf{v}_1 = [0010\ 0101\ 0110]^T ,$$

$$\mathbf{v}_2 = [1011\ 0001\ 0001]^T ,$$

$$\mathbf{v}_3 = [1101\ 0110\ 1000]^T ,$$

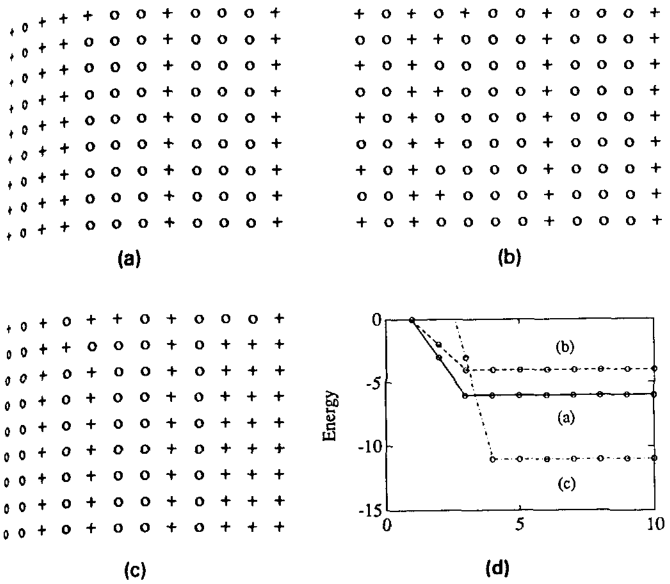


Fig. 3. Examples of stability and convergence in Hopfield's model. Iteration time in (a), (b), and (c) are in rows top to bottom (+ denotes +1). The first row is the result of the first iteration. We see that s(0) (a) converges to v_2 without skew, (b) oscillates with skew, and (c) converges to a vector with a different skew that is not in the library. (d) Energy transitions with iteration for each case.

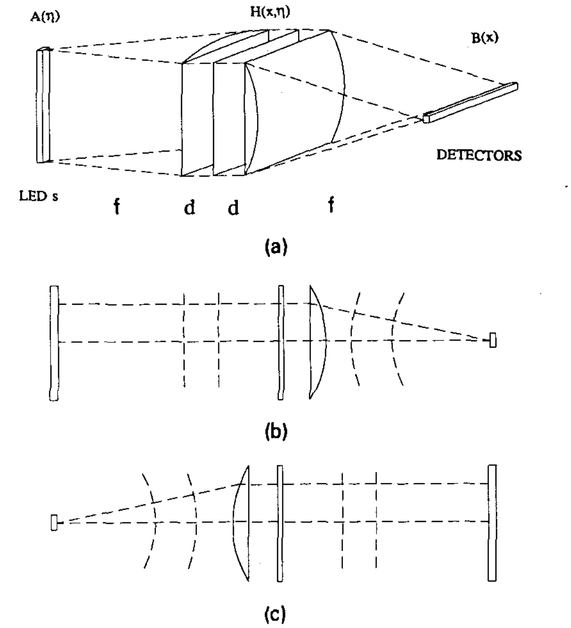


Fig. 4. (a) Diagram of a matrix-vector multiplier; (b) side view; (c) top view.

we form in accordance with Hopfield's recipe the interconnection matrix

$$A = (BB^T - NI) ,$$

where $N = 3$, $V = [v_1 : v_2 : v_3]$, $B = 2V - 1$, and 1 is a matrix of ones. We form the iteration $S(n+1) = \eta[AS(n)]$, where, for $1 \le n \le N$, $\eta_n(\bullet)$ is the unit step function [$\eta_n(x) = 1$ if $x \ge 0$ and is 0 otherwise]. If we initialize with $S(0) = [1011\ \ 0001\ \ 1010]^T$ and iterate synchronously, the solution of the operation converges to v_2. However, if this operation has a skew of three clocks delays for $\tau_{4,3}$ and $\tau_{4,11}$ and two clocks delays for those remaining, then the iteration oscillates. The iteration converges to $[0010\ \ 1001\ \ 0111]^T$ in the case that the skew is three clocks delays for $\tau_{4,3}$, $\tau_{4,11}$, $\tau_{6,9}$, and $\tau_{6,12}$ and two clocks delays for those remaining. These three examples are respectively illustrated in Figs. 3(a) through 3(c). The energy of the neural network at the nth iteration is defined as

$$E(n) = -\frac{s(n)^T As(n)}{2} .$$

A plot of the energy for these three examples is in Fig. 3(d).

6. SKEW IN OPTICAL PROCESSORS

The major source of skew in optical processors is the time delay resulting from the differing optical lengths (OL)[28] between input and output. For nondispersive clock skew, the time delay from the input point (ξ, η) to the output point (x, y) can be written as

$$\tau(x,y;\ \xi,\eta) = \frac{OL}{c} ,$$

where c is the speed of light in free space. In this section, we show some examples of optical processor operations that are not affected by skew.

Example 1: A commonly used processor for performing matrix-vector multiplication is shown in Fig. 4(a). The top view of this processor, shown in Fig. 4(c), resembles a point source collimator. Since we are interested in only a single point at the output, there is no clock skew due to OL differences from this perspective. The side view of the processor, shown in Fig. 4(b), is equivalent to that in Fig. 4(c) except that the input and output are reversed. Since there is no skew from this view, the composite processor has no clock skew due to OL differences. The total OL from $(0, \eta)$ at the input plane to $(x, 0)$ at the output plane is given by

$$OL_t = OL_1 + 2d + OL_2 = 2f + 2\Delta_0 + 2d ,$$

where OL_1 is the OL from $(0, \eta)$ to (x, η), OL_2 is from (x, η) to $(x, 0)$, and $OL_1 = OL_2 = f + \Delta_0$, with Δ_0 denoting the OL through the center of the lens. This equation states that OL_t is constant for all (x, η) pairs.

Example 2: Figure 5(a) illustrates the Stanford matrix-vector multiplier[29] that is more light efficient than the one in Fig. 4(a). Since the performance is similar to that in Fig. 4(b), there is no skew apparent in the side view in Fig. 5(b). From the perspective of the top view in Fig. 5(c), however, the apparent point source input is incident on the detector as a cylindrical wave. Under a Fresnel approximation, the skew is therefore quadratic. The total OL for this processor is given by

$$OL_t = 5f + \Delta + \frac{\eta^2}{f} = K + \frac{x^2}{f} ,$$

where Δ denotes the sum of the OLs through the center of the lenses and K is a constant. The time delay for this processor is thus

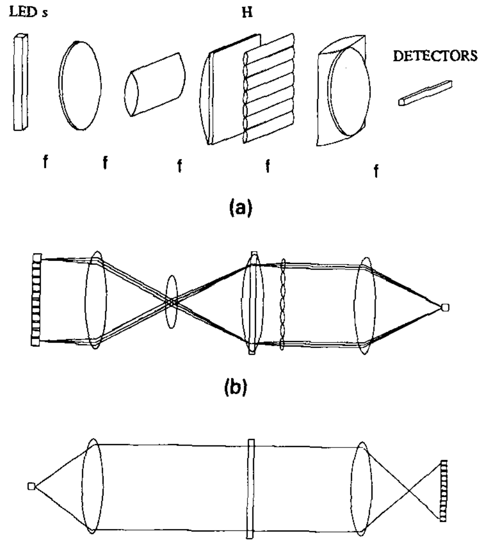


Fig. 5. (a) Illustration of the Stanford matrix-vector multiplier; (b) side view; (c) top view.

$$\tau(x,\eta) = \frac{OL_t}{c} = \frac{K}{c} + \frac{x^2}{fc} \ .$$

Therefore, this processor is temporally skewed, but the skew is separable. By lemmas 3 and 4, if Eq. (13) is true, any iterative processors that employ this Stanford matrix-vector multiplier using a pointwisely soft nonlinearity in the feedback path that satisfies Eq. (13) will converge independent of this skew. An example of such a processor is the alternating projection neural network (APNN), which uses linear feedback.[11,12]

7. FINAL REMARKS

The primary source of clock skew in optical system is differing optical lengths. We have investigated the effects of clock skew on the performance of iterative processors and have shown that clock skew does not affect the convergence and stability of the solution when the feedback is contractive. We also have shown some examples of optical systems that have no skew or are not affected by skew when used iteratively.

8. ACKNOWLEDGMENTS

This work was supported by the Strategic Defense Initiative Organization/Innovative Science and Technology administered by the Office of Naval Research through the Optical Systems Lab at Texas Tech University in Lubbock and by the Washington Technology Center at the University of Washington. In addition, L. E. Atlas was supported by a National Science Foundation Presidential Young Investigator's Award.

9. APPENDIXES

9.1. Proof of lemma 1

We take the Laplace transform of Eq. (5)

$$s_n(s) = \sum_m a_{nm} s_m(s) \exp(-s\tau_{nm}) + f_n(s) \ ,$$

or equivalently, in matrix form,

$$\mathbf{s}(s) = \mathbf{A}(s)\mathbf{s}(s) + \mathbf{f}(s) \ .$$

Since $\|\mathbf{A}(s)\| < 1$, $\det[\mathbf{I} - \mathbf{A}(s)] \neq 0$, and $\mathbf{s}(s)$ becomes

$$\mathbf{s}(s) = [\mathbf{I} - \mathbf{A}(s)]^{-1}\mathbf{f}(s) \ .$$

By applying the final value theorem, we obtain

$$\mathbf{s}(\infty) = \lim_{s\to 0} s\mathbf{s}(s) = \lim_{s\to 0}[\mathbf{I} - \mathbf{A}(s)]^{-1}[s\mathbf{f}(s)]$$

$$= [\mathbf{I} - \mathbf{A}(s)]^{-1}\mathbf{f} \ ,$$

which is our desired result.

9.2. Proof of lemma 2

Let $\mathbf{y}$ be an N dimensional vector and $\mathbf{A}(s) = [a_{ij}\exp(-s\tau_{ij})]$ and N×N matrix. Then

$$\|\mathbf{A}(s)\mathbf{y}\|^2 = \mathbf{y}^*\mathbf{A}(s)^*\mathbf{A}(s)\mathbf{y}$$

$$= \sum_i\sum_j\sum_k y_i^* a_{ki}^* \exp(-s^*\tau_{ki}) a_{kj} \exp(-s\tau_{kj}) y_j$$

$$\leq \sum_i\sum_j\sum_k |y_i|\, |a_{ki}\exp(-s\tau_{ki})|\, |a_{kj}\exp(-s\tau_{kj})|\, |y_j| \ ,$$

where the asterisks denote the complex conjugate for scalars and the complex conjugate transpose for matrices. Let $z_i = |y_i|$, $b_{ki} = |a_{ki}\exp(-s\tau_{ki})|$, and $\mathbf{z} = [z_i]$. Then

$$\|\mathbf{A}(s)\mathbf{y}\|^2 \leq \sum_i\sum_j\sum_k z_i b_{ki} b_{kj} z_j = \|\mathbf{Bz}\|^2 \ .$$

Since $\|\mathbf{y}\| = \|\mathbf{z}\|$, $\|\mathbf{A}(s)\| \leq \|\mathbf{B}\| < 1$. Since convergence is assured for $\mathbf{B}$ by assumption, we conclude that Eq. (5) converges to Eq. (4) and our proof is complete.

9.3. Proof of lemma 3

From the assumption, we have

$$\mathbf{A}(s) = \{a_{ij}\exp[-s(u_i + v_j)]\} = [a_{ij}\exp(-su_i)\exp(-sv_j)] \ .$$

By letting

$$\mathbf{D}_u = \mathrm{diag}[\exp(-su_1), \exp(-su_2), \ldots, \exp(-su_N)] \ ,$$

$$\mathbf{D}_v = \mathrm{diag}[\exp(-sv_1), \exp(-sv_2), \ldots, \exp(-sv_N)] \ ,$$

$\mathbf{A}(s)$ becomes

$$\mathbf{A}(s) = \mathbf{D}_u\mathbf{A}\mathbf{D}_v \ .$$

Since $\|\mathbf{D}_u\| \leq 1$ and $\|\mathbf{D}_v\| \leq 1$ for $\mathrm{Re}(s) \geq 0$, we have

$$\|\mathbf{A}(s)\| \leq \|\mathbf{D}_u\|\,\|\mathbf{A}\|\,\|\mathbf{D}_v\| \leq \|\mathbf{A}\| < 1 \ .$$

From this and lemma 1, we conclude that Eq. (5) converges to Eq. (4).

9.4. Proof of lemma 4

We rewrite Eq. (10) as

$$s_n(t) = \eta_n[y_n(t) + g_n(t)] + h_n(t) \ ,$$

where

$$y_n(t) = \sum_m a_{nm} s_m(t - \tau_{nm}) .$$

Let $\mathbf{y}(t)$ denote the vector of $y_n(t)$. We take the norm of $\mathbf{s}_n(t_1) - \mathbf{s}_n(t_2)$:

$$\begin{aligned}\|\mathbf{s}(t_1) - \mathbf{s}(t_2)\| &\le \|\eta[\mathbf{y}(t_1) + \mathbf{g}_n(t_1)] - \eta[\mathbf{y}(t_2) + \mathbf{g}(t_2)]\| \\ &\quad + \|\mathbf{h}(t_1) - \mathbf{h}(t_2)\| \\ &\le \|\mathbf{y}(t_1) - \mathbf{y}(t_2)\| + p(t_1, t_2) ,\end{aligned}$$

or

$$\|\mathbf{s}(t_1, t_2)\| \le \|\mathbf{y}(t_1, t_2)\| + p(t_1, t_2) , \qquad (14)$$

where

$$p(t_1, t_2) = \|\mathbf{g}(t_1) - \mathbf{g}(t_2)\| + \|\mathbf{h}(t_1) - \mathbf{h}(t_2)\| ,$$

$$\mathbf{s}(t_1, t_2) = \mathbf{s}(t_1) - \mathbf{s}(t_2) ,$$

and $\mathbf{y}(t_1, t_2)$ denotes the vector of

$$y_n(t_1, t_2) = \sum_m a_{nm}[s_m(t_1 - \tau_{nm}) - s_m(t_2 - \tau_{nm})] .$$

Substitute $\mathbf{Q}$ for $\mathbf{s}$ in the linear skewed iteration in Eq. (5) and take the difference between Eq. (5) at t_1 and t_2 to obtain

$$\mathbf{Q}(t_1, t_2) = \mathbf{x}(t_1, t_2) + \mathbf{f}(t_1, t_2) ,$$

where

$$\mathbf{f}(t_1, t_2) = \mathbf{f}(t_1) - \mathbf{f}(t_2) ,$$

$$\mathbf{Q}(t_1, t_2) = \mathbf{Q}(t_1) - \mathbf{Q}(t_2) ,$$

and $\mathbf{x}(t_1, t_2)$ denotes the vector of

$$x_n(t_1, t_2) = \sum_m a_{nm}[Q_m(t_1 - \tau_{nm}) - Q_m(t_2 - \tau_{nm})] .$$

We construct $\mathbf{f}(t_1, t_2)$ to be colinear with $\mathbf{x}(t_1, t_2)$. Let

$$\mathbf{f}(t_1, t_2) = \begin{cases} \text{any vector} , & \text{for } \mathbf{x}(t_1, t_2) = 0 , \\ C(t_1, t_2)\mathbf{x}(t_1, t_2) , & \text{otherwise} , \end{cases}$$

where the proportionality constant $C(t_1, t_2)$ is chosen so that $\|\mathbf{f}(t_1, t_2)\| = p(t_1, t_2)$. Then, by construction,

$$\|\mathbf{Q}(t_1, t_2)\| = \|\mathbf{x}(t_1, t_2)\| + p(t_1, t_2) .$$

Because $\|\mathbf{Q}(t_1, t_2)\|$ converges to 0 by the assumption, Eq. (14) also converges:

$$\|\mathbf{s}(t_1, t_2)\| \to 0, \quad \text{for } t_1 \to \infty, t_2 \to \infty ,$$

or equivalently, since our vectors are in a finite dimensional Euclidean space,

$$\lim_{t \to \infty} \mathbf{s}(t) = \mathbf{s}(\infty) ,$$

and our proof is complete.

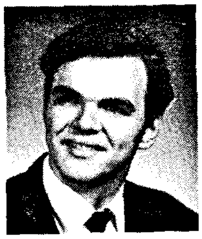
10. REFERENCES

1. J. Cederquist and S. H. Lee, "Confocal feedback systems with space-variance, time sampling, and secondary feedback loops," J. Opt. Soc. Am. 71, 643-650 (1981).
2. T. Sato, K. Saski, Y. Nakamura, M. Linzer, and S. J. Norton, "Tomographic image reconstruction from limited projections using coherent optical feedback," Appl. Opt. 20(17), 3073-3076 (1981).
3. R. J. Marks II, "Coherent optical extrapolation of 2-D band-limited signals: processor theory," Appl. Opt. 19(10), 1670-1672 (1980).
4. R. J. Marks II and D. K. Smith, "Iterative coherent processor for band-limited signal extrapolation," in *1980 International Optical Computing Conference*, Book 1, W. T. Rhodes, ed., Proc. SPIE 231, 106-111 (1980).
5. M. O. Hagler and S. V. Bell, "Nyquist stability conditions for laser feedback amplifiers with Gaussian or Lorentzian gain profiles," J. Opt. Soc. Am. A 3, 308-318 (1986).
6. M. A. G. Abushagur, H. J. Caulfield, P. M. Gibson, and M. Habli, "Superconvergence of hybrid optoelectronic processors," Appl. Opt. 26(23), 4906-4907 (1987).
7. H. J. Caulfield, J. H. Gruninger, J. E. Ludman, K. Steiglitz, H. Rabitz, J. Gelfand, and E. Tsoni, "Bimodal optical computers," Appl. Opt. 25(18), 3128-3131 (1986).
8. M. A. G. Abushagur and H. J. Caulfield, "Speed and convergence of bimodal optical computers," Opt. Eng. 26(1), 22-27 (1987).
9. H. J. Caulfield and M. A. G. Abushagur, "Hybrid analog-digital linear algebra processors," in *Optical and Hybrid Computing*, H. H. Szu, ed., Proc. SPIE 634, 86-95 (1987).
10. R. J. Marks II, L. E. Atlas, and K. F. Cheung, "A class of continuous level neural nets," in *Proc. Fourteenth Congress of the International Commission for Optics*, H. H. Arsenault, ed., Proc. SPIE 813, pp. 29-30 (1987).
11. R. J. Marks II, L. E. Atlas, S. Oh, and K. F. Cheung, "Architectures for a continuous level neural network based on alternating orthogonal projections," in *Neural Network Models for Optical Computing*, R. A. Athale and J. Davis, eds., Proc. SPIE 882, 90-92 (1988).
12. R. J. Marks II, L. E. Atlas, S. Oh, and K. F. Cheung, "Optical processor architectures for alternating projection neural networks," Opt. Lett. 13, 533-525 (1988).
13. J. Shamir, "Fundamental speed limitations on parallel processing," Appl. Opt. 26(9), 1567 (1987).
14. B. D. Clymer and J. W. Goodman, "Timing uncertainty for receivers in optical clock distribution for VLSI," Opt. Eng. 27(11), 944-954 (1988).
15. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proc. Natl. Acad. Sci. USA 81, 3088-3092 (1982).
16. K. F. Cheung, L. E. Atlas, and R. J. Marks II, "Synchronous versus asynchronous behavior of Hopfield's content addressable memory neural net," Appl. Opt. 26(22), 4808-4813 (1987).
17. R. J. Marks II and D. K. Smith, "Gerchberg-type linear deconvolution and extrapolation algorithms," in *Transformations in Optical Signal Processing*, W. T. Rhodes, J. R. Fienup, and B. E. A. Saleh, eds., Proc. SPIE 373, 161-178 (1984).
18. R. J. Marks II, "Class of continuous level associative memory neural nets," Appl. Opt. 26(10), 2005-2010 (1987).
19. K. F. Cheung, S. Oh, R. J. Marks II, and L. E. Atals, "Neural net associative memories based on convex set projections," in *Proc. IEEE First. Int. Conf. on Neural Networks* (San Diego, June 1987).
20. R. J. Marks II, S. Oh, and L. E. Atlas, "Alternating projection neural networks," IEEE Trans. Cir. Syst. (in press).
21. R. J. Marks II, L. E. Atlas, S. Oh, and J. A. Ritcey, "The performance of convex set projection based neural networks," in *Proc. IEEE Conf. on Neural Information Processing Systems—Natural and Synthetic* (Boulder, Col., Nov. 1987).
22. A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*, Springer-Verlag, New York (1982).
23. R. P. Lippman, "An introduction to computing with neural nets," IEEE Acoust. Speech Sig. Proc. ASSP-35(4), 7 (1987).
24. J. G. McDonnel, R. J. Marks II, and L. E. Atlas, "An introduction to neural networks for solving combinatorial search problems," in *Northcon/88 Conf. Record*, Vol. II, pp. 868-876, Western Periodicals Co., North Hollywood, Calif. (1988).
25. 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).
26. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," Appl. Opt. 24(10), 1469-1475 (1985).
27. R. J. Marks II and L. E. Atlas, "Geometrical interpretation of Hopfield's content addressable memory neural network," in *Northcon/88 Conf. Record*, Vol. II, pp. 964-977, Western Periodicals Co., North Hollywood, Calif. (1988).

28. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, New York (1980).
29. J. W. Goodman, A. R. Dias, and L. M. Woody, "Fully parallel, high speed incoherent optical method for performing discrete Fourier transforms," Opt. Lett. 2, 1-3 (1978).

Seho Oh was born in Daejeon, Korea, in 1957. He received his BS degree in electronics engineering from Seoul National University and the MS degree in electrical engineering from Korea Advanced Institute of Science and Technology, Seoul. From 1981 through 1986, he was with Goldstar Central Research Laboratory in Seoul. Since September 1986, he has been a doctoral candidate in the Department of Electrical Engineering at the University of Washington. His present research interests are signal processing, optical neural nets, and pattern recognition.

Dong Chul Park was born in Buchon, Korea, in 1958. He received the BS degree in electronics engineering from Sogang University, Seoul, in 1980, and the MS degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Seoul, in 1982. From 1982 to 1985 he was a research engineer in the Goldstar Central Research Laboratory, Seoul. Since September 1985 he has been working toward the Ph.D. in the Department of Electrical Engineering at the University of Washington. His research interests include fault tolerant computing, optical neural networks, signal processing, and pattern recognition.



Robert J. Marks II received his BS (1972) and MS (1973) in electrical engineering, both from Rose-Hulman Institute of Technology in Terre Haute, Ind., and his Ph.D. (1977) from Texas Tech University in Lubbock. He toiled for one year as a reliability engineer for the Department of the Navy (1973–74). Dr. Marks joined the faculty of the Department of Electrical Engineering at the University of Washington in December 1977, where he is currently a professor. He was awarded the outstanding Branch Councilor Award in 1982 by IEEE and in 1984 was presented with an IEEE Centennial Medal. He is a senior member of IEEE. Prof. Marks is also chair of the IEEE Circuits and Systems Society Technical Committee on Neural Systems and Applications and is chairman pro tem of the IEEE Neural Networks Committee. He was a cofounder and first president of the Puget Sound Section of the OSA and was recently elected that organization's first honorary member. Dr. Marks has more than 80 archival journal and proceedings publications in the areas of detection theory, signal recovery, optical computing, and artificial neural processing. He is a member of Eta Kappa Nu and Sigma Xi.

Les Atlas is currently an assistant professor at the University of Washington. He received his Ph.D. in electrical engineering from Stanford University in 1983 and has been doing research in speech processing, artificial neural networks, and optics since. He is a recipient of a National Science Foundation Presidential Young Investigator Award.