# Effects of Clock Skew in Iterative Neural Network and Optical Feedback Processors

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

***Abstract:***

Although optical processor implementations of artificial neural networks promise algorithmic convergence at the speed of light, little attention is normally given to the consequences of different optical path lengths required within the processor on the resulting performance. Such clock skew can have significant degrading effects on the predicted accuracy and speed of the processor. A similar problem occurs in iterative electronic 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 steady state solutions which 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 mapping in the feedback path. Feedback using hard nonlinearities, on the other hand, can result in a steady state solution which depends on the clock skew.

***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. In certain cases, disregarding this *clock skew* in processor analysis can lead to drastically different implementation results. A similar problem occurs in iterative asynchronous artificial neural networks where the communication time delay between two neurons is proportional to their physical separation.

In this paper, we show that, in certain feedback algorithms, clock skew does not affect 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 the steady state result when hard nonlinearities are used in the feedback path [14-16].

***Preliminaries***:

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

Let a field of $N$ states, $\{S_n \mid 1 \le n \le N\}$, be altered by feedback in a skewed processor. Let $\vartheta_n$ denote the instantaneous operator that maps the previous states into the current $n$th state at time $t$. We can then write:

$$S_n(\mathrm{t}) = \vartheta_n[\ \{S_m(\mathrm{t}-\tau_{nm})\ | 1 \le m \le N\}\ ]\ ;\ 1 \le n \le N \tag{1}$$

where $\tau_{nm}$ 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 $t \to \infty$ and assume a stable steady state, then (1) becomes:

$$S_n(\infty) = \vartheta_n[\ \{\ S_m(\infty)\ | 1 \le m \le N\}]\ ;\ 1 \le n \le N \tag{2}$$

Although not explicitly indicated, this steady state solution is, in general, a function of the clock skew. If, however, (2) has only a single solution for all $S_n(\infty)$, then the clock skew has no effect on the steady state solution.

Some specific examples of iterations unaffected by clock skew will now be given.

***Solution of Simultaneous Equations:***

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

$$\mathbf{S}_{n+1} = \mathbf{f} + \mathbf{A}\mathbf{S}_n \tag{3}$$

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 $\mathbf{A}$ does not exceed one [17]. With reference to (1), if performed on a skewed processor, (3) would be implemented as*

$$S_n(t) = \sum_{m=1}^{N} a_{nm} S_m(t-\tau_{nm}) + f_n \mu(t-\gamma_n); \quad 1 \le n \le N \tag{5}$$

where $\gamma_n$ denotes the time skew of the input's contribution, $\mu(\cdot)$ is the unit step function and $f_n$ is the $n$th element of f. Letting $t \to \infty$ and assuming a stable result gives

$$S_n(\infty) = \sum_{m=1}^{N} a_{nm} S_m(\infty) + f_n ; \quad 1 \le n \le N \tag{6}$$

or, equivalently, in matrix- vector form

$$S_\infty = \mathbf{A}\, S_\infty + f \tag{7}$$

If $\mathbf{I} - \mathbf{A}$ is not singular, then the solution to this equation is (4). Clock skew therefore does not effect the solution. The *alternating projection neural network* (without the sigmoid nonlinearity) when interpreted either homogeneously [18-21] or, in layered form [20-21] from the hidden to output layer, is a special case of this example.

***Contractive Operators:***

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

$$S(\infty) = \vartheta\, S(\infty) \tag{8}$$

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

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

where

$$|| a ||^2 = a^T a$$

and $0 \le r < 1$ . If $0 \le r \le 1$, then $\vartheta$ is said to be *nonexpansive*. If $\vartheta$ is contractive, then the solution of (8) is unique [17-19] and there is no contribution of clock skew to the steady state result. When $\vartheta$ is nonexpansive, (8) can have a number of solutions.

---

* Alternately, the forcing vector can be time varying due to, say, the source's rise time. If the forcing vector approaches a steady state value of $f_n$ and the system remains stable, then the steady state result remains the same.

*Example:*
Our previous example is a special case of a contractive mapping since, from (6),

$$\vartheta\, S(\infty) = \mathbf{A}\;\; S(\infty)\; + f$$

The operator is contractive if

$$\| (\mathbf{A}\, x - f) - (\mathbf{A}\, y - f) \| = \| \mathbf{A}(x-y) \| < r \| x-y \|$$

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

*Example:*

We can nonlinearly generalize (7) to

$$S_\infty = \eta(A\, S_\infty + f_1) + f_2 \tag{10}$$

where both $f_1$ and $f_2$ are forcing functions and $\eta$ is a pointwise nonlinearity. By a pointwise nonlinearity, we mean that, if $z = \eta\, w$, then the *n*th element of z is equal to $\eta_n(w(n))$ where $\eta_n$ is a given function. In the parlance of neural networks, $\eta$ would be referred to as a *sigmoid* operator [25-26]. Using (9), the corresponding operator is contractive if

$$\| \eta(Ax + f_1) - \eta(Ay + f_1) \| \le r \| x - y \| \tag{11}$$

We will show that the operator is contractive if the spectral radius of A does not exceed one and $\eta$ contains *soft nonlinearities*. A nonlinearities is said to be soft if

$$| d\,\eta_n(\phi) / d\phi | \le 1\,;\; 1 \le m \le N$$

for all $\phi$. As is illustrated in Figure 1, this constraint has the property that

$$| \eta_n x - \eta_n y | \le | x-y |$$

As a result,

$$\| \eta(Ax + f_1) - \eta(Ay + f_1) \| \le \| (Ax + f_1) - (Ay + f_1) \|$$

Using the results of the previous example, the operator corresponding to (10) is clearly contractive and clock skew has no effect on the final result.

This unique convergence constraint can be generalized to the requirement that

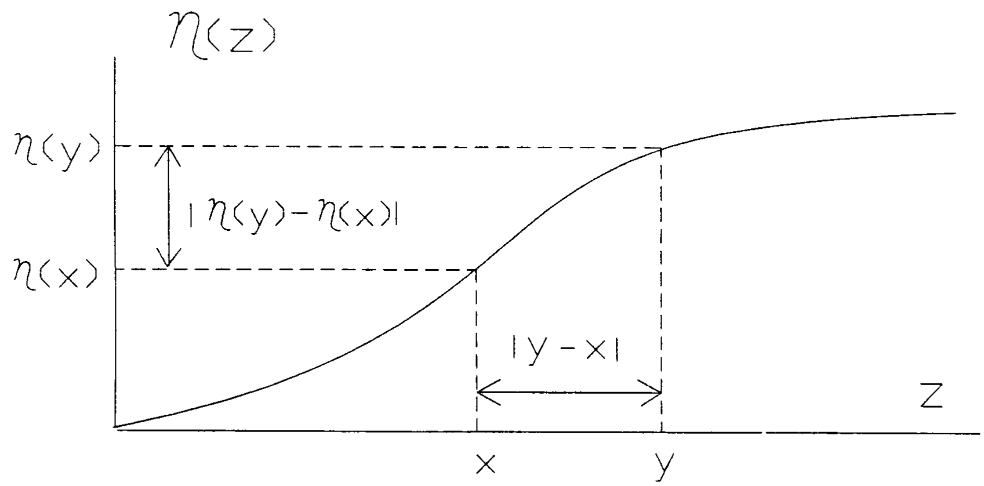


***Figure Caption:***
Figure 1: An example of a soft nonlinearity. Any interval on the z axis maps to a smaller interval.

---

$$\left| \, d\,\eta_n(\phi) \,/\, d\phi \, \right| < \rho(\mathbf{A}) \,;\, 1 \leq m \leq N$$

where $\rho(\mathbf{A})$ is the spectral radius of **A**.

***Hard Nonlinearities:***

Clock skew can be a factor when implementing an iterative algorithm with hard nonlinearities. Conwell [22] discusses such effects in certain neural networks which use no forcing functions and a unit step nonlinearity:

$$\eta_n(\phi) = \begin{cases} 1 & ;\ \phi \geq 0 \\ 0 & ;\ \text{otherwise} \end{cases}$$

Different clock skews produced different steady state results.

***Conclusions:***

We have shown that clock skew does not affect the steady state solution of iterative algorithms when the feedback operation is contractive. Such skew can occur in in asynchronously operated neural networks and in optical processors with feedback.

*Acknowledgements*:
This work was supported by the SDIO/IST administered by ONR through the Optical Systems Lab at Texas Tech University in Lubbock and by the Washington Technology Center at the University of Washington. A discretionary gift from the PhysioControl Corp. for the authors' work in neural networks is also gratefully appreciated. In addition, L.E. Atlas was supported by an NSF Presidential Young Investigators Award.

***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, 3073-3076 (1981).
3. R. J. Marks, II, "Coherent optical extrapolation of 2-D band-limited signals: processor theory," *Appl. Opt.* 19, 1670-1672 (1980).
4. R. J. Marks, II and D. K. Smith, "An iterative coherent processor for bandlimited signal extrapolation," *Proc. SPIE* 231, 106-111 (1980).
5. Marion O. Hagler and Steven 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. Mustafa A. G. Abushagur, H. John Caulfield, Peter M. Gibson, and Mohammad Habli, "Superconvergence of hybrid optoelectronic processors," *Appl. Opt.* 26, 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, 3123 (1986).
8. M. A. G. Abushagur and H. J. Caulfield, "Speed and Convergence of Bimodal Optical Computers," *Opt. Eng.* 26, 022 (1987).
9. H. J. Caulfield and M. A. G. Abushagur, "Hybrid Analog-Digital Algebra Processors," *Proc. Soc. Photo-Opt. Instrum. Eng.* 634, 86 (1986).
10. R. J. Marks, II, L. E. Atlas and K. F. Cheung, "A class of continuous level neural nets," *Proceedings of the Fourteenth Congress of the International Commission for Optics*, pp. 29-30, Quebec City, Quebec, Canada, August 24-28, 1987.
11. R. J. Marks, II, L. E. Atlas and K. F. Cheung, "Architectures for a continuous level neural network based on alternating orthogonal projections," *Proceedings of O-E/LASE '88 Conference on Neural Network Models for Optical Computing,* Los Angeles, January 1988.
12. R.J. Marks II, L.E. Atlas and K.F. Cheung "Optical processor architectures for alternating projection neural networks", *Optics Letters* (in press).

13. Joseph Shamir, "Fundamental speed limitations on parallel processing," *Appl. Opt.* 26, 1567 (1987).
14. J.J. Hopfield "Neural networks and physical systems with emergent collective computational abilities", *Proc. of the National Academy of Sciences*, USA, 81, 1982, pp.3088-3092.
15. K.F. Cheung, L.E. Atlas and R.J. Marks II "Synchronous versus asynchronous behavior of Hopfield's content addressable memory", *Applied Optics*, vol. 26, pp.4808-4813 (1987).
16. D. W. Tank and J. J. Hopfield,"Simple neural optimization networks: an A/D converter, signal decision circuit and a linear programming circuit," *IEEE Trans. Circuits Syst.*, vol. CAS-33, p. 533 (1986).
17. R. J. Marks, II, and D. K. Smith, "Gerchberg - type linear deconvolution and extrapolation algorithms" in *Transformations in Optical Signal Processing*, edited by W. T. Rhodes, J. R. Fienup and B. E. A. Saleh, SPIE vol. 373, pp. 161-178 (1984).
18. R.J. Marks II "A class of continuous level associative memory neural nets", *Applied Optics*, vol. 26, pp.2005-2009 (1987).
19. K.F. Cheung, R.J. Marks II and L.E. Atlas "Neural net associative memories based on convex set projections", *Proceedings of the IEEE First International Conference on Neural Networks*, San Diego, June 1987.
20. R.J. Marks II, L.E. Atlas, J.A. Ritcey, S. Oh and K.F. Cheung "Alternating projection neural networks", (in review).
21. R.J. Marks II, L.E. Atlas, S. Oh and J.A. Ritcey "The performance of convex set projection based neural networks", *Proceedings of the IEEE Conference on Neural Information Processing Systems - Natural and Synthetic*, Boulder Colorado, November 1987.*
22. Ronald W. Schafer, Russell M. Mersereau and Mark A. Richards, "Constrained Iterative Restoration Algorithms," *Proceedings of the IEEE*, 69, 432-450 (1981).
23. V. T. Tom, T. F. Quatieri, M. H. Hayes and J. H. McClellan, "Convergence of iterative nonexpansive signal reconstruction algorithms," *IEEE Trans. Accoust. Speech and Signal Processing*, ASSP-29, 1052-1058 (1981).
24. Arch W. Naylor and George R. Sell, *Linear Operator Theory in Engineering and Science*, Springer-Verlag, New York (1982).
25. R. P. Lippman, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, p. 7 (April 1987).
26. J. G. McDonnel, R. J. Marks, II, and L. E. Atlas "An introduction to neural networks for solving combinatorial search problems," *IEEE Expert*, (in press)
27. Peter R. Conwell, "Effects of Connection Delays in Two State Model Neural Circuits," *International Conf. on Neural Networks*, San Diego, p. III-95-104.