

The History of Echo Cancellation

EDITORS' INTRODUCTION

Man Mohan Sondhi (mms@research.avayalabs.com) was born on 18 December 1933 in Ferozepore, India. His undergraduate studies in Delhi and Bangalore, India, were completed in 1953, and his graduate studies at the University of Wisconsin were completed in 1957 when he received his Ph.D. degree. He was with Bell Laboratories for 39 years as a member of the technical staff and later as a distinguished member of technical staff before joining Avaya Labs Research as a consultant four years ago. Dr. Sondhi's research work has focused on echo cancellation, speech signal processing, and automatic speech recognition. He coauthored the book *Advances in Network and Acoustic Echo Cancellation* (2001) and coedited *Advances in Speech Signal Processing* (1991). He was elected a Bell Laboratories Fellow (1998) and was awarded the Bell Laboratories Arnold trophy (1997), the IEEE Eric Sumner prize (1998), and the IEEE Signal Processing Society Best Paper Award (2001).

His nonprofessional interests include listening to folk music, traveling, and reading. *Brave New World* by Aldous Huxley, *God's Politics* by Jim Wallis, *1984* by George Orwell, and *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design* by Richard Dawkins are some of his favorite books. One of his favorite quotations is from Daniel J. Boorstin, award-winning author and librarian of Congress from 1975 to 1987: "The greatest obstacle to discovery is not ignorance—it is the illusion of knowledge." In the story told here, Dr. Sondhi recalls his work and events that led to the discovery of echo cancellation. In his smooth story telling, you may recognize qualities such as honesty and understatement, which he also appreciates in his collaborators.

—Adriana Dumitras and George Moschytz
"DSP History" column editors
adrianad@ieee.org
moschytz@isi.ee.ethz.ch

The history of echo cancellation begins on 10 July 1962. On that date, Bell Laboratories ushered in the era of satellite communications with the launching of Telstar, the first telecommunications satellite. (In those days Bell Labs used to be called the Bell Telephone Labs and later, AT&T Bell Labs, to emphasize the fact that the labs were owned by AT&T. Still later, when AT&T Bell Labs split up in 1996, the labs became the research arm of Lucent Technologies.) By a happy coincidence, I had joined Bell Labs just eight days earlier, unaware of the fact that the launching of Telstar was destined to have a fairly significant effect on my career. This is because it was soon to lead to the development of the echo canceller, which is one of the projects on which I spent a considerable amount of time. In what follows, I will give an account of my per-

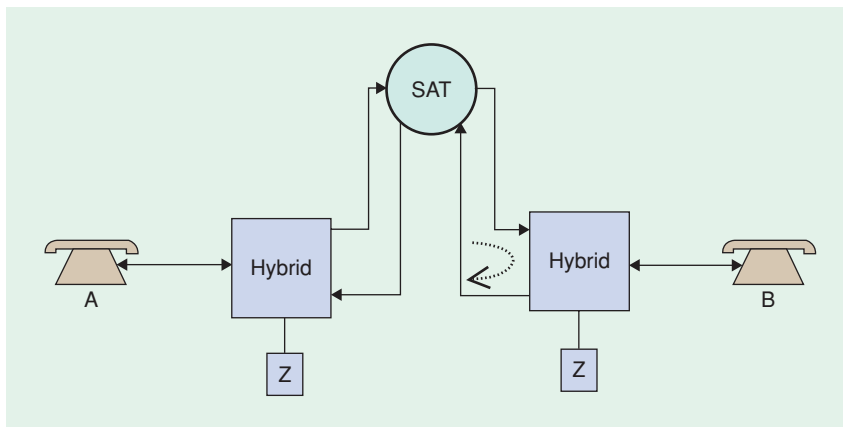
sonal involvement in the development of echo cancellers.

THE ECHO PROBLEM IN TELECOMMUNICATIONS

To understand the connection between satellite communications and echo cancellation, we need to digress a bit and look at the problem of echo in telecommunications, the methods used to alleviate its harmful effects on telephone conversations, and the need for a brand new solution with the advent of satellite communications. To begin with, let us ignore telephony involving wireless and Internet connections. These will be mentioned briefly towards the end of the article.

The local circuit (called the subscriber's loop) that connects every conventional analog telephone to the local central office is a two-wire line. To set up a local telephone connection, the central

office essentially connects the two wires of one party to the two wires of the other party, thus creating a single loop circuit on which the telephones at the two ends inject their signals. This strategy does not work for long-distance connections whenever the signals need to be amplified to compensate for circuit losses. Since amplifiers are one-way devices, separate paths are needed for transmission and reception. Hence, long-distance telephone circuits are four-wire circuits: one pair of wires for transmission and one for reception. The interface between a two-wire circuit and a four-wire circuit is a device called a *hybrid*. Thus, a long distance telephone connection consists of a four-wire circuit and two hybrids—one at each end to connect to the local two-wire loop, as shown in Figure 1. Let us restrict the discussion to the right-hand end of the connection (B). We will call this end the near end, and the other



[FIG1] A long-distance connection.

end (A) the far end. Everything discussed next will apply to the other end as well, with the roles of A and B interchanged.

The function of the hybrid shown on the right-hand side of Figure 1 is to allow signals to flow between subscribers A and B along the paths labeled with solid arrows but to prevent a signal from being reflected back to subscriber A along the path labeled with a dotted arrow. (Similarly, the hybrid on the left should ideally prevent a reflected signal from returning to subscriber B.) A hybrid is essentially a balanced bridge network and would behave ideally if a balancing impedance (shown as Z in Figure 1) could be synthesized so as to exactly match the impedance of the two-wire circuit. Unfortunately, this is not possible in practice because there are far fewer four-wire circuits than there are two-wire circuits. Therefore, a hybrid may get connected to any of the subscriber loops served by the central office. And subscriber loops by their very nature have a wide variety of characteristics—various lengths, types of wire, type of telephones, etc. Also, the impedance may change during a conversation, for example if an extension phone is switched on or off. So the best that can be done is to provide some compromise balancing impedance. The compromise impedance that is used in the United States provides a reflected signal that is (on average) 11 dB below the input signal. The reflection at a hybrid is the main source of reflections in a telephone network, although in principle, a reflection

can occur anywhere where there is an impedance mismatch.

If the leading edge of the reflected signal reaches the far end subscriber A with a round trip delay of a few milliseconds, then it is perceived as a reverberation. If, however, the round trip delay exceeds a few tens of milliseconds, the reflection is perceived as a distinct echo. Even at low levels, such echoes are very disruptive of conversations, and the level of tolerable echo decreases as the delay increases.

Echoes from hybrids have been around ever since long distance telephony was introduced many decades ago, and methods have been devised to combat their effects. The most successful of these methods are based on a device known as an *echo suppressor*. It is basically a voice-operated switch that disconnects (or introduces a very large transmission loss) in the return path whenever a decision mechanism decides that the signal on that path is an echo. As seen in Figure 1, this return path is also shared by the speech of the near-end subscriber B. Therefore, errors made by the decision mechanism might sometimes allow an echo and sometimes chop low-level portions of the near-end speech. Nevertheless, well-designed echo suppressors provide a very good solution to the echo problem on terrestrial telephone lines.

MOTIVATION FOR ECHO CANCELLATION

How can the echo problem be handled on other than terrestrial telephone lines?

Telephony via Telstar (mentioned earlier), which was a low-flying satellite at about 200 mi above the surface of the Earth, would not have introduced any major problems beyond those of terrestrial telephony. However, even in the early 1960s, it was anticipated that telephony via satellites would eventually use so-called geostationary satellites. These are satellites in the earth's equatorial plane that take exactly 24 h to complete an orbit, so they stay at a fixed position relative to the earth. Such satellites need to be about 23,000 mi above the earth's surface. The round trip echo path is therefore approximately $4 \times 23,000 = 92,000$ mi. That corresponds to about 500 ms of propagation time for a radio wave. With additional delays at the end point, the total round-trip delay of the echo can be as high as 600 ms. Experiments showed that a delay of this length in the absence of echoes can be acceptable for telephony. However, such a long delay changes the dynamics of a conversation to the extent that echo suppressors no longer provide a satisfactory solution to the echo problem. Bursts of unsuppressed echo and clipping of soft segments of the near-end speech become much too frequent to be acceptable. Improvements in the design of echo suppressors were not able to deal with this problem. A new technique was required that did not interrupt the echo return path, and that technique was *echo cancellation* as opposed to *echo suppression*.

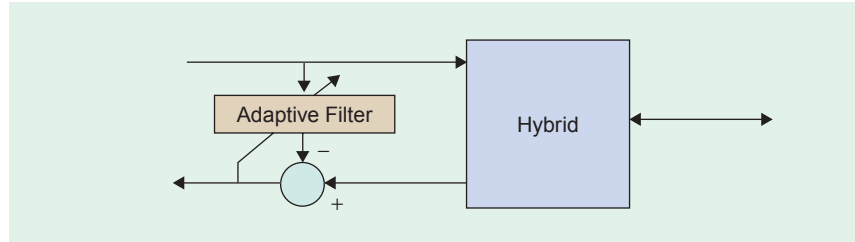
IDEA OF THE ECHO CANCELLER

The idea of echo cancellation, which is illustrated in Figure 2, is to generate a synthetic replica of the echo and to subtract it from the return signal. Note that this would eliminate the echo without interrupting the speech from the near end. If the path of the echo is linear, then its impulse response may be measured, and a filter with that impulse response may be synthesized. The far end signal filtered by such a filter gives the desired replica of the real echo. This is in fact what was done in early attempts at echo cancellation, and it was shown that the echo can be effectively cancelled. This so-called "open-loop"

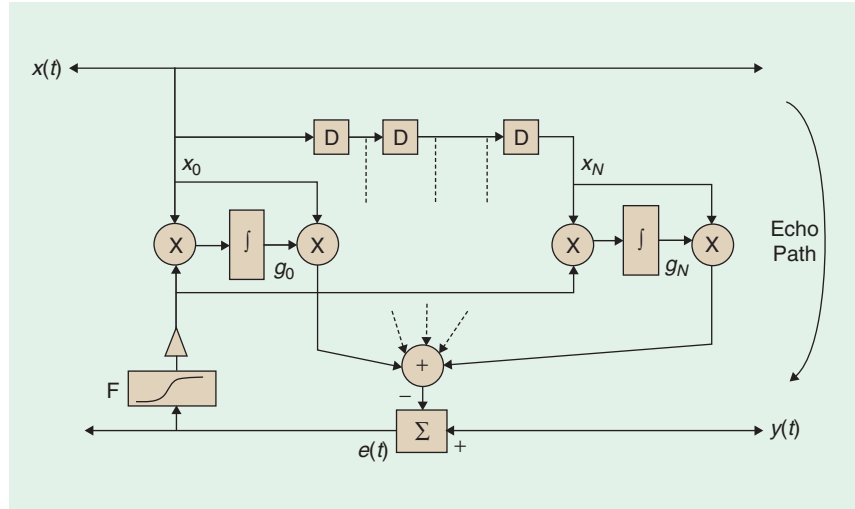
method of echo cancellation has the disadvantage that test signals need to be introduced, and such test signals would be needed intermittently throughout a conversation because echo paths are not perfectly stationary. These considerations led to the idea of *adaptive* echo cancellation, first proposed by John L. Kelly, Jr. of Bell Labs. He invited Benjamin F. Logan and me to join him in exploring this idea. John Kelly unfortunately died of a sudden heart attack very soon after this, and the task of writing the first full paper [1] summarizing this work fell on me. Two patents were filed simultaneously in 1966 and simultaneously issued in 1970 [2], [10]. The idea was to use the speech signal itself as the test signal and to adapt a filter so as to automatically drive itself to match its characteristics to whatever echo path it was connected.

IMPLEMENTATIONS OF THE ECHO CANCELLER

The echo canceller was originally conceived of as an analog device, and its first analog implementation was done in 1966 by A.J. Presti with my help [3]. The block diagram of the circuit is shown in Figure 3, and the basic theory of its operation is shown in “The Algorithm of the Echo Canceller.” As shown there, under fairly reasonable conditions, the weights of the tapped-delay line filter adjust themselves to drive the error signal to zero. Given the technology of the time, even a 50 tap (compared to several thousand taps in some modern implementations) echo canceller was estimated to cost about US\$1,500—far too expensive to find any application in the telephone network. Even digital implementations with discrete components (first implemented, as far as I can remember, by Comsat Corporation in the mid 1970s) were not feasible. The commercial use of echo cancellers had to wait till the 1980s and was initiated by the development of an integrated circuit chip by Duttweiler and Chen [4] in 1980. Today, of course, echo cancellers are deployed by the millions on the telephone network, and there are single chips with cancellers on



[FIG2] Block diagram of the echo canceller.



[FIG3] Block diagram of the acoustic echo canceller, where D is a unit delay, g_i are adaptable gains, x_i are delayed versions on the input signal x , and μ is the loop gain.

THE ALGORITHM OF THE ECHO CANCELLER

Let us explain the operation of the echo canceller of Figure 3 in the ideal case when the echo path is constant and the return signal contains just the echo, without any near-end signal or noise. Each box marked D in Figure 3 represents a delay of T s and $x(t)$ is the time signal from the far end. We assume that the echo, $y(t)$, can be modeled as $y(t) = \sum_{i=0}^N h_i x(t - iT)$, where the coefficients h_i are unknown but constant.

Define the vectors $\mathbf{x} \equiv [x(t), x(t - T), \dots, x(t - NT)]$ and $\mathbf{h} \equiv [h_0, h_1, \dots, h_N]$. Then $y(t) = \mathbf{h}^T \mathbf{x}$, where the superscript T denotes matrix transpose.

Define the vector $\mathbf{g} \equiv [g_0, g_1, \dots, g_N]$, and the vector $\mathbf{r} = \mathbf{h} - \mathbf{g}$. Then the error signal $e(t) = \mathbf{r}^T \mathbf{x}$.

Examining the inputs to the integrators in Figure 3, we see that $(d\mathbf{g})/(dt) = \mu F(e)\mathbf{x}$. Since the vector \mathbf{h} is assumed to be constant, $(d\mathbf{g})/(dt) = -(d\mathbf{r})/(dt)$. Therefore $(d\mathbf{r})/(dt) = -\mu F(e)\mathbf{x}$, where μ is the loop gain.

Premultiplying both sides of this equation by $2\mathbf{r}^T$ we get:

$$\frac{d}{dt} \mathbf{r}^T \mathbf{r} = -2\mu F(e) \mathbf{r}^T \mathbf{x} = -2\mu F(e) e.$$

If the nonlinear function F is chosen such that its sign is the same as that of its argument, then clearly the norm of the vector \mathbf{r} is strictly decreasing as long as the error is not zero. Integrating the last equation between zero and some time then shows that $|e| \rightarrow 0$. Under some weak assumptions about the properties of the signal $x(t)$, it then follows that $\mathbf{r}^T \mathbf{r} \rightarrow 0$. Thus $\mathbf{r} \rightarrow 0$, i.e., $\mathbf{g} \rightarrow \mathbf{h}$. In other words, the canceller adapts itself to the echo path.

them, each with hundreds of taps, capable of cancelling echoes simultaneously on hundreds of channels.

A digital equivalent of the circuit of Figure 3 is obtained by replacing the time derivative by a first difference. It then becomes the famous least-mean square (LMS) algorithm, whose discovery in the late 1950s was described in a recent "DSP History" column by Bernard Widrow [5]. A normalized form of that algorithm (NLMS) is almost exclusively the algorithm used in echo cancellers today. In recent times, a variation has been proposed by my colleague Don Duttweiler, in which the loop gain μ is different for different taps and μ is proportional to the magnitude of the current estimate of the tap weight. This algorithm called proportional NLMS (PNLMS) converges much faster than NLMS in practice. However, a rigorous proof of its convergence has never been found, except in the limiting case of slow convergence.

During the 1970s, my interest in echo cancellers was dormant because, as I mentioned earlier, they did not seem to be commercially viable. However, Debasis Mitra and I did some mathematical analyses [7] of the echo canceller that were interesting though not of much practical importance. One of the problems we looked at is the estimation of rate of convergence. As noted in "The Algorithm of the Echo Canceller," it is relatively easy to prove convergence of

the algorithm. Surprisingly, however, to this day, it is not possible to derive rigorous estimates of the rate of convergence except at the limit of slow convergence. Debasis and I were able to find some crude bounds on the rate of convergence. Other properties that we investigated were the effects of noise and other perturbations and the effect of nonlinearities in multipliers.

INITIAL APPLICATIONS

As mentioned in [5], the LMS algorithm (a digital equivalent of the circuit illustrated in Figure 3) was first proposed for adaptive switching circuits. As far as I can tell, widespread applications of the algorithm (with some modifications) are echo cancellation and automatic equalization of modems, the latter introduced by R.W. Lucky [6] in the 1960s. Noise cancellation, proposed during the 1970s, is another application that has found some limited use. (The echo in the telephone network is a disturbance of the type that noise cancellers deal with. Therefore, one may in fact characterize the echo canceller as the original noise canceller!)

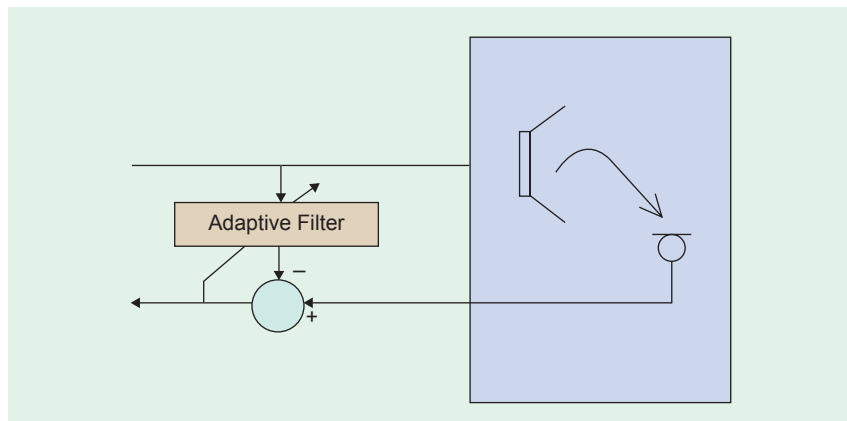
The echo cancellers described above are called network or line echo cancellers. The advent of fiber optic cables has eliminated satellite delay, which was the original motivation for echo cancellers. However, with increased use of coding to reduce bit rate, especially in wireless telephony, satellite

delay has been replaced by coding delay. Also, delays are introduced due to packetization in Internet-based telephony (VoIP). So the need for echo cancellers promises to remain unabated for a long time to come. Another very widespread application of echo cancellers is in data transmission. But they are very different in detail from cancellers for voice signals and are outside the scope of this article.

SOME NEWER APPLICATIONS

Since the 1990s, there has been considerable interest in acoustic echo cancellers for teleconferencing. In this application, illustrated in Figure 4, the echo is generated acoustically by the coupling between the loudspeaker and microphone via the impulse response of a conference room. In principle, the problem is similar to that of network echo cancellers, but there are some differences in practice because the impulse responses in the acoustic case are much longer and more variable. The subband echo canceller structure proposed by Walter Kellermann has proved useful for this application because it reduces the amount of computations required per second.

Finally, let us mention the modern application of echo cancellation for stereophonic (in general, multichannel) teleconferencing, in which a pair of microphones and a pair of loudspeakers are used in each conference room. Such conferences are not yet common, but they will most likely be the preferred type of teleconference in the not-too-distant future. The stereophonic case turns out to be qualitatively different from the single channel case in that the solution of the stereophonic case is not unique. In collaboration with my colleagues Dennis Morgan and Jacob Benesty, I have been seeking good solutions to this problem. Jacob's clever suggestion of using a nonlinearity to partially decorrelate the two channels is so far the best solution that has been found.



[FIG4] Acoustic echo cancellation. The box represents a room. The arrow inside the room is meant to represent the acoustic path from the loudspeaker to the microphone. The actual coupling from loudspeaker to microphone is a superposition of a large number of paths reflected off walls and objects in the room.

(continued on page 102)

In such a case it is usually advisable to represent the coefficients using 16 b and accumulate the sum of the polynomial terms to a 32-b variable. When we multiply the input x with itself, we have to discard half of the bits of the result to fit it into 16 b. This introduces a quantization error, which becomes progressively larger in proportion to the result with higher powers of x .

Therefore, on fixed-point systems the most viable alternative is usually to utilize low-degree (e.g., quadratic or cubic) polynomials and increase the number of subintervals, if the accuracy must be improved. Notice, however, that at some point the word length of the multiplication operands starts to limit the accuracy.

EXAMPLE

To summarize all of the above, let's study a complete example of function approximation using polynomials. We will approximate the function $\ln(x)$ and use $[0.5, 1]$ as the approximation interval. We select $S = 4$ and $D = 3$ and thus have to find the coefficients for the polynomial $p(x) = ax^3 + bx^2 + cx + d$ in the subintervals $[0.5, 0.625]$, $[0.625, 0.75]$, $[0.75, 0.875]$, and $[0.875, 1]$. Our optimal approximation polynomial search routine returns the polynomial coefficients shown in Table 1.

[TABLE1] POLYNOMIAL COEFFICIENTS VERSUS SUBINTERVAL RANGE.

| Subinterval | a | b | c | d |
|-----------------------|--------|---------|--------|---------|
| $0.5 \leq x < 0.625$ | 1.8951 | -4.7898 | 5.3674 | -2.4163 |
| $0.625 \leq x < 0.75$ | 1.0339 | -3.1954 | 4.3822 | -2.2131 |
| $0.75 \leq x < 0.875$ | 0.6250 | -2.2834 | 3.7035 | -2.0446 |
| $0.875 \leq x < 1$ | 0.4062 | -1.7130 | 3.2073 | -1.9006 |

Let's now see what happens with the input $x = 987$. First x is normalized to the interval $[0.5, 1]$ by moving the binary point ten places to the left. (In the fixed-point case we don't really move anything, we just imagine there is a point on the left side of the leftmost 1 b. Then we take this into account when doing the final denormalization.) The normalized input becomes $987 \cdot 2^{-10} = 0.9638671875$. This value falls into the last subinterval $[0.875, 1]$. With the values given in the table we get $p(x) = -0.036802$. Denormalization yields $-0.036802 + 10 \cdot \ln(2) = 6.89466998$. The precise value for $\ln(987)$ is 6.89467004.

CONCLUSIONS

We explained the use of polynomial approximations for calculating function values and demonstrated their viability in a fixed-point numerical environment. Because of their fast computation and capability of efficient parallelization, polynomial approximations are the algorithm of choice for the realization of elementary functions in

modern, pipelined DSP architectures. (Listings of various MATLAB routines modeling the proposed function approximation schemes described in this article are available at <http://apollo.ee.columbia.edu/spm/?i=external/tips>.)

ACKNOWLEDGMENTS

Associate Editor Rick Lyons is gratefully acknowledged for his numerous helpful comments and suggestions. Appreciation is extended to Cleve Moler (MathWorks Inc.) for his review of the manuscript.

AUTHOR

Jyri Ylösto (Jyri-Matti.Ylostalo@nokia.com) is a research engineer at Nokia Research Center in Helsinki, Finland. He received his M.Sc. degree in 1993 and licentiate degree in 1999 from the Helsinki University of Technology.

REFERENCES

- [1] J. Muller, *Elementary Functions, Algorithms and Implementation*. Boston, MA: Birkhäuser, 1997.
- [2] M. Powell, *Approximation Theory and Methods*. Cambridge, UK: Cambridge Univ. Press, 1981. **SP**

LAST BUT NOT LEAST

This personal history of the development of echo cancellation would not be complete without mentioning other related works. Needless to say, several laboratories, especially in Australia and Japan, have investigated the problem of echo cancellation. A comprehensive description of the early development of echo cancellers may be found in the review article [8]. Another description especially summarizing the progress in the past decade or so may be found in a book [9] recently published in collaboration with my colleagues Jacob Benesty, Tomas Gaensler, Dennis Morgan,

and Steve Gay. References [8] and [9] also discuss other very important components of an echo canceller (besides the adaptation algorithm) such as double-talk detectors and nonlinear processors.

REFERENCES

- [1] M.M. Sondhi, "An adaptive echo canceler," *Bell Syst. Tech. J.*, vol. 46, no. 3, pp. 497-511, Mar. 1967.
- [2] M.M. Sondhi, "Closed loop adaptive echo canceler using generalized filter networks," U.S. Patent 3 499 999, 1970.
- [3] M.M. Sondhi and A.J. Presti, "A self-adaptive echo canceler," *Bell Syst. Tech. J.*, vol. 45, no. 12, pp. 1851-1854, Dec. 1966.
- [4] D.L. Duttweiler and Y.S. Chen, "A single chip VLSI echo canceler," *Bell Syst. Tech. J.*, vol. 59, no. 2, pp. 149-160, 1980.

- [5] B. Widrow, "Thinking about thinking: the discovery of the LMS algorithm," *IEEE Signal Process Mag.*, vol. 22, no. 1, pp. 100-106, Jan. 2005.

- [6] R.W. Lucky, "Automatic equalization for digital communication," *Bell Syst. Tech. J.*, vol. 44, no. 4, pp. 547-588, Apr. 1965.

- [7] M.M. Sondhi and D. Mitra, "New results on the performance of a well-known class of adaptive filters," *Proc. IEEE*, vol. 64, no. 11, pp. 1583-1597, 1976.

- [8] M.M. Sondhi and D. Berkley, "Silencing echoes on the telephone network," *Proc. IEEE*, vol. 68, no. 8, pp. 948-963, 1980.

- [9] J. Benesty, T. Gaensler, D.R. Morgan, M.M. Sondhi, and S.L. Gay, *Advances in Network and Acoustic Echo Cancellation*. New York: Springer-Verlag, 2001.

- [10] J.L. Kelly and B.F. Logan, "Self-adaptive echo canceler," U.S. Patent 3 500 000, 1970.