Effective MSE Criterion for Combined Linear-Viterbi Equalization

Saeed A. Aldosari, Saleh A. Alshebeili and Abdulhameed M. Al-Sanie

Department of Electrical Engineering
King Saud University, Riyadh, Saudi Arabia

Abstract— **Combined linear-Viterbi equalizer (CLVE) is a class of receivers employing a linear pre-filter before the Viterbi algorithm (VA) in order to shape the original channel impulse response to some shorter desired impulse response which allows a lower complexity VA. In this paper, we present a new approach for optimizing CLVEs by making use of the Genetic Algorithm to minimize the effective mean squared error (EMSE). Unlike the conventional MSE approach, the EMSE approach takes into account the effect of noise correlations resulting from the pre-filtering process. Simulation results indicate that the proposed EMSE optimization algorithm has an advantage over the conventional MSE algorithms in that better system performance can be achieved.**

I. Introduction

It is well known that maximum likelihood sequence estimation (MLSE) is the optimum method for the detection of data in the presence of intersymbol interference (ISI) and additive white Gaussian noise (AWGN) provided that the channel impulse response (CIR) is known or can be precisely estimated [1]. Viterbi algorithm (VA) provides an efficient way of performing MLSE recursively when the length of CIR is finite. The symbol error rates of VA are often much lower than error rates of the symbol by symbol detectors. However, the use of MLSE is limited to channels having short delay spread. This is caused by the high computational demand, which is growing exponentially with the length of CIR. For real life channels, the use of MLSE may become impractical.

A considerable amount of research has been undertaken to reduce the complexity of MLSE using various techniques. Combined linear-Viterbi equalizer (CLVE) is a class of receivers employing a linear pre-filter before the VA in order to shape the original CIR to some shorter desired impulse response (DIR) which allows a lower complexity VA. Several optimization criteria were used to optimize the linear pre-filter and DIR. Minimization of the mean-squared-error (MMSE) is among the simplest and most practical CLVE design approaches [2], [3], [4]. The objective of the MMSE design approach is to minimize the MSE between the output of the pre-filter and the output of the DIR.

Although the MSE criterion plays an important role in the design of CLVEs, there are other important factors that should be taken into account such as the effect of noise correlations and the Euclidean distance property of the overall impulse response of the system. Neglecting such factors in the design process may lead to severe degradation in the overall performance of CLVE systems. In contrast to the MSE, the effective MSE (EMSE) criterion, suggested by Fredricsson [5], includes the most important factors affecting the performance of CLVEs. Hence, optimization of CLVE parameters via the EMSE criterion has an advantage in that better system performance can be achieved.

Despite the fact that it incorporates most of the factors related to the performance of CLVE systems, the EMSE criterion is nonlinear which makes the optimization process a non-trivial task. It is mostly due to this nonlinearity that the EMSE criterion has been overlooked in most of the literature related to this problem. To the best of our knowledge, there are no attempts have been yet reported to solve this optimization problem under the constraint of finite length DIR (which is necessary to limit the complexity of VA).

In this paper, we develop a new algorithm for the optimization of EMSE criterion using the genetic algorithm (GA); an elegant numerical optimization technique that has recently found a wide spread use in many fields including signal processing and communications (see for example [6] and the references therein). There are a number of advantages that have motivated the adoption of the GA as opposed to other numerical optimization techniques. One of the main advantages of the GA is that it is not susceptible to problems with local minima that arise with multimodal error surfaces, and the GA can be guaranteed to approach the global minimum under suitable circumstances. This suggests that the performance achieved by the GA can be used as a reference against which the other optimization techniques can be compared with. Another advantage of the GA is that it can be successfully used for optimizing systems with nonlinear or even discontinuous cost functions. Holland showed the power of the GA by relating it to the multiarmed bandit problem and by guaranteeing that the minimum expected cost can be achieved [10].

The GA has the disadvantage of slow convergence, which may prevent its use in some communication systems that are characterized by fast time-varying channels. However, the GA can be used successfully in slow time-varying channels. In addition, because of the parallel nature of the GA, there is an ongoing research toward implementing the GA in modern parallel architectures [7], [8]. Recent developments in massively parallel computer architectures demonstrated the potential to eliminate the computational bottleneck in implementing evolution techniques. In this regard, Twardowski [9] has shown the feasibility of GA real-time implementation using parallel associative architectures.

In the rest of this paper, we first introduce the basic concept of EMSE criterion and present the background material necessary for its computation. Second, we develop a new algorithm for the optimization of EMSE criterion using the genetic algorithm (GA). Third, we evaluate the performance of the proposed algorithm by considering practical communication channels.

II. CLVE Model

The block diagram of a CLVE system is shown in Figure 1. The transmitted sequence a_k is as-

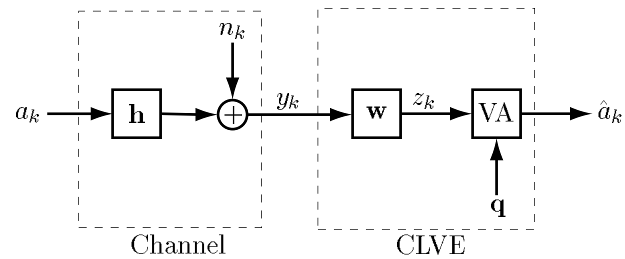

Fig. 1. Combined linear-Viterbi equalizer (CLVE).

sumed to be a zero mean M-ary sequence selected from the set $\{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. The communication channel is modeled by a transversal filter with impulse response $\mathbf{h}$ of length $L{+}1$ and corrupted with AWGN sequence n_k with variance σ^2. The received samples y_k are processed by a linear pre-filter $\mathbf{w}$ and the output sequence z_k is, then, passed to the VA to perform MLSE. The function of the linear pre-filter $\mathbf{w}$ is to shape the original CIR $\mathbf{h}$ to some shorter DIR $\mathbf{q}$. With this arrangement, the complexity of the VA can be made lower by considering the shorter DIR $\mathbf{q}$ instead of the original longer CIR $\mathbf{h}$. From Figure 1, the input to the VA is given by

$$\mathbf{z} = \mathbf{w} * (\mathbf{h} * \mathbf{a} + \mathbf{n})$$

where $*$ denotes convolution. In the following section, we describe the error events and error bounds of CLVEs since it is required to define and understand the EMSE.

III. Error Bounds for CLVE

In VA, errors tend to occur in small groups known as error events. When the metric of a certain incorrect path happens to be less than the metric of the correct path, the VA decides in favor of the incorrect path. This makes the following VA decisions diverge from the correct path until they remerge again after certain time. An error event that extends from k_1 to k_2 has the property,

$$\hat{s}_{k_1} = s_{k_1},\ \hat{s}_{k_2} = s_{k_2}, \text{but } \hat{s}_k \neq s_k \text{ for } k_1 < k < k_2 \tag{1}$$

where s_k and $\hat{s}_k$ are the states of the correct and incorrect decoding paths at time k, respectively. The length of an error event is defined as $\ell = k_2 - k_1 - L$ and is always greater than or equal to 1. It is convenient to define an error vector e

corresponding to this error event as

$$e = [\epsilon_{k_1}, \epsilon_{k_1+1}, \ldots, \epsilon_{k_1+\ell}] \tag{2}$$

where the components of e are defined as

$$\epsilon_k = a_k - \hat{a}_k \tag{3}$$

and $\hat{a}_k$ represents the estimated sequence. The error vector is characterized by the property that $\epsilon_{k_1} \neq 0$, $\epsilon_{k_1+\ell} \neq 0$, and there is no sequence of L consecutive elements that are zero.

Since the length of error events is generally not limited, there are an infinite number of possible distinct error events with different probabilities of occurrence. The probability of occurrence for a certain error event is governed by the closeness of that error event to the true sequence which, in turn, could be measured by the Euclidean weight d_e. The Euclidean weight of an error event e is defined as the Euclidean length of the error signal vector associated with that error event, i.e., [1], [4]

$$d_e = \sqrt{\sum_{k=k_1}^{k_2-1} \left(\sum_{j=0}^{L} h_j \epsilon_{k-j} \right)^2}. \tag{4}$$

Assuming that the information symbols are equally probable and that the symbols in the transmitted sequence are statistically independent, an upper-bound on the overall symbol error probability P_M of CLVEs is given by [4]

$$P_M \leq \sum_{e \in \mathcal{E}} w_e \Pr(\zeta_e \geq \frac{1}{2} d_e) \prod_{i=k_1}^{k_1+\ell} \frac{M - |\epsilon_i|/2}{M}. \tag{5}$$

where $\mathcal{E}$ denotes the complete set of valid error events and w_e is the number of nonzero components (Hamming weight or number of symbol errors) in each error event e. Note that it is not practical to consider all error events since the set is infinite. However, the probability of the very long error events is negligible (as can be seen from Equation (5)), consequently, the set of error events can be truncated with little effect. In Equation (5), ζ_e represents the length of the projection of noise terms on the direction of the error event e. For CLVEs, the probability term $\Pr(\zeta_e \geq \frac{1}{2} d_e)$ is hard to find. This is because the equalized signal in CLVEs is distorted by correlated noise in addition to some residual ISI. The correlation of the noise is a direct result of the pre-filtering process, while the residual ISI is caused by the truncation of the overall impulse response. For this reason, ζ_e will generally be non-Gaussian, and its distribution is hard to find. Qureshi in [4] has derived an upper-bound of $\Pr(\zeta_e \geq \frac{1}{2} d_e)$.

IV. EMSE Criterion

The EMSE, originally suggested by Fredricsson [5], is defined as

$$\text{EMSE} \triangleq \sup_{e \in \mathcal{E}'} \frac{\sigma_e^2}{d_e^2} \tag{6}$$

where σ_e^2 denotes the variance of ζ_e and $\mathcal{E}'$ is a subset of $\mathcal{E}$ containing the most effective error events. The subset $\mathcal{E}'$ is defined according to [5]

$$\mathcal{E}' \triangleq \left\{ e : d_e \leq d' \text{ and } \prod_{i=k_1}^{k_1+\ell} \frac{M - |\epsilon_i|/2}{M} \geq p' \right\} \tag{7}$$

where d' and p' are arbitrary positive numbers properly selected such that

$$\sum_{e \in \mathcal{E}'} w_e \Pr(e) \simeq \sum_{e \in \mathcal{E}} w_e \Pr(e) \tag{8}$$

and $\Pr(e)$ is the probability of occurrence of error event e.

In what follows, we describe the evaluation of EMSE for CLVE systems. From Figure 1, the input to the VA is given by

$$\begin{aligned} \mathbf{z} &= \mathbf{w} * (\mathbf{h} * \mathbf{a} + \mathbf{n}) \\ &= \mathbf{w} * \mathbf{h} * \mathbf{a} + \mathbf{w} * \mathbf{n} \end{aligned} \tag{9}$$

where $\mathbf{h}$ and $\mathbf{w}$ are the impulse responses of the channel and the pre-filter, respectively. Let $\mathbf{q}$ denote the finite length DIR of the system, i.e., the detector assumes that the impulse response of the system is equal to $\mathbf{q}$. Therefore, Equation (9) can be expressed as

$$\begin{aligned} \mathbf{z} &= \mathbf{q} * \mathbf{a} + (\mathbf{w} * \mathbf{h} - \mathbf{q}) * \mathbf{a} + \mathbf{w} * \mathbf{n} \\ &= \mathbf{q} * \mathbf{a} + (\mathbf{b} * \mathbf{a} + \mathbf{w} * \mathbf{n}) \end{aligned} \tag{10}$$

where $\mathbf{b} = \mathbf{w} * \mathbf{h} - \mathbf{q}$ is the impulse response of the residual ISI. It can be seen from Equation (10) that the input to the VA is composed of two components: the desired term $\mathbf{q} * \mathbf{a}$ and the undesired term $\mathbf{b} * \mathbf{a} + \mathbf{w} * \mathbf{n}$ which will be denoted by ζ throughout this section. The undesired term ζ itself is composed of two components: the correlated Gaussian noise $\chi = \mathbf{w} * \mathbf{n}$ and the residual ISI $\psi = \mathbf{b} * \mathbf{a}$ which is not handled by the VA due to truncation of the overall impulse response $\mathbf{w} * \mathbf{h}$.

The evaluation of the EMSE requires computing $\frac{\sigma_e^2}{d_e^2}$ for each error event e in a finite set of error events $\mathcal{E}'$. The length d_e of the signal error sequence, when considered as a vector in the Euclidean space, is given by

$$\begin{aligned} d_e^2 &= \|e * \mathbf{q}\|^2 \\ &= \sum_i \left(\sum_j q_j \epsilon_{i-j} \right)^2 . \end{aligned} \tag{11}$$

The variance σ_e^2 of ζ_e, the projection of the distortion in the direction of error event e is given by

$$\sigma_e^2 = \sigma_{\chi e}^2 + \sigma_{\psi e}^2 \tag{12}$$

where $\sigma_{\chi e}^2$ and $\sigma_{\psi e}^2$ are the variances of the projection of the correlated Gaussian noise χ and the residual ISI ψ in the direction of error event e, respectively. Let $\mathbf{v}$ denotes the unit vector in the direction of the error signal $e * \mathbf{q}$, i.e.,

$$\mathbf{v} = \frac{e * \mathbf{q}}{\|e * \mathbf{q}\|}. \tag{13}$$

It can be shown that the variance of the correlated noise χ along the unit vector $\mathbf{v}$ in the Euclidean space is given by the quadratic form [4]

$$\begin{aligned} \sigma_{\chi e}^2 &= \mathbf{v}\ \mathbf{C}_{\chi\chi}\ \mathbf{v}^T \\ &= \sigma^2\,\mathbf{v}\ \mathbf{C}_{\mathbf{ww}}\ \mathbf{v}^T \end{aligned} \tag{14}$$

where $\mathbf{C}_{\chi\chi}$ is the noise covariance matrix and the elements of $\mathbf{C}_{\mathbf{ww}}$ are given by

$$C_{\mathbf{ww}}(i,j) = \sum_i \sum_j w_i w_{i-j}. \tag{15}$$

Note that in computing the desired variance we need only consider a noise covariance matrix of dimension equal to the length of the error event. Therefore, if $N_{\mathbf{v}}$ is the size of $\mathbf{v}$, we need to consider a noise covariance matrix of size $N_{\mathbf{v}} \times N_{\mathbf{v}}$. Similarly, the variance of the residual ISI ψ along the unit vector $\mathbf{v}$ is given by

$$\begin{aligned} \sigma_{\psi e}^2 &= \mathbf{v}\ \mathbf{C}_{\psi\psi}\ \mathbf{v}^T \\ &= \sigma_a^2\,\mathbf{v}\ \mathbf{C}_{\mathbf{bb}}\ \mathbf{v}^T \end{aligned} \tag{16}$$

where $\mathbf{C}_{\psi\psi}$ is the covariance matrix of the residual ISI, σ_a^2 is the variance of the input sequence, and the elements of $\mathbf{C}_{\mathbf{bb}}$ are given by

$$C_{\mathbf{bb}}(i,j) = \sum_i \sum_j b_i b_{i-j}. \tag{17}$$

In summary, the EMSE of a CLVE system is first evaluated by constructing the subset $\mathcal{E}'$ such that it contains the most effective error events. For each error event, $\frac{\sigma_e^2}{d_e^2}$ is computed using Equations (11), (12),(13), (14), and (16). The EMSE of the system is the maximum value of $\frac{\sigma_e^2}{d_e^2}$.

V. Optimization Algorithm

Our objective in this paper is to optimize CLVEs with respect to the EMSE criterion. Like the MSE, the EMSE criterion can be used to optimize both the pre-filter and DIR of CLVEs. It has been shown in [5] that the optimum pre-filter obtained by minimizing the EMSE is almost identical to that obtained by minimizing the MSE provided that filters with reasonable high complexity are allowed. In contrast, the DIRs optimized via the EMSE and MSE are generally different. Therefore, we can use the genetic algorithm to seek the optimum setting of the DIR while the linear pre-filter is designed according to the conventional MSE criterion.

The GA, introduced by John Holland in 1975, is a type of directed random search that mimics the process of biological evolution [6]. Instead of a naive search and select mechanism in traditional optimization techniques, GA uses cross-over to exchange information among existing solutions to locate better solutions. Usually, a simple GA consists of three operations: Selection, Genetic Operation, and Replacement.

Initially, a population is generated randomly. The population comprises a group of chromosomes from which candidates can be selected for the solution of the problem. It should be noted that each chromosome $x_i, (i = 1, 2, \ldots , N_c)$ represents a trial solution to the problem setting. Since we are searching for the optimum setting of the DIR, each gene represents one of the coefficients of the DIR. We use bit-string encoding with arbitrary fixed range to represent each gene. The number of bits for each gene should be selected such that it maintains higher optimization accuracy while not increasing the GA complexity considerably. The fitness values of all chromosomes are evaluated by calculating the EMSE for each chromosome in a decoded form. Using the Roulette Wheel Selection procedure [6], a particular group of chromosomes (parents) is selected from the population to generate the offspring. The offspring are generated by using two genetic operations: uniform crossover and mutation. Probability terms P_c and P_m are set to determine the crossover and mutation rates, respectively. After generating the subpopulation (offspring), the fitness of the offspring is evaluated in a similar fashion to what has been done to their parents. The chromosomes in the current population are then replaced by their offspring based on the Generational-Replacement strategy in which each population of size N_p generates an equal number of new chromosomes to replace the old population [6]. To increase the GA conversion speed, this strategy will be combined with an elitist strategy where a number of the best chromosomes are copied into the succeeding generations [6]. Such a GA cycle is repeated until the desired termination criterion is reached.

It is worth mentioning that the convergence speed of the GA is largely affected by the initialization procedure of the first generation. For example, starting from a completely random population may have the effect of reducing the convergence speed of GA. Therefore we suggest to initialize some of the chromosomes in the initial population to non random values (in our case, this could be done by initializing them to values computed from the MSE solution) for the purpose of increasing the convergence speed of the optimization algorithm.

VI. Performance Evaluation

Computer simulation was conducted to evaluate the performance of the proposed EMSE optimization approach. In the simulation, we have considered a 2-PAM data transmitted through a typical urban area channel which has been generated according to the COST207 model [11], [12]. The frequency response of this channel is shown in Figure 2. This channel is characterized by its long impulse response ($L+1=21$) resulting from multiple reflections from far objects. The implementation of full complexity MLSE to equalize ISI resulting from such a channel would be impractical due to the large number of states required by the VA ($=2^{20}$).

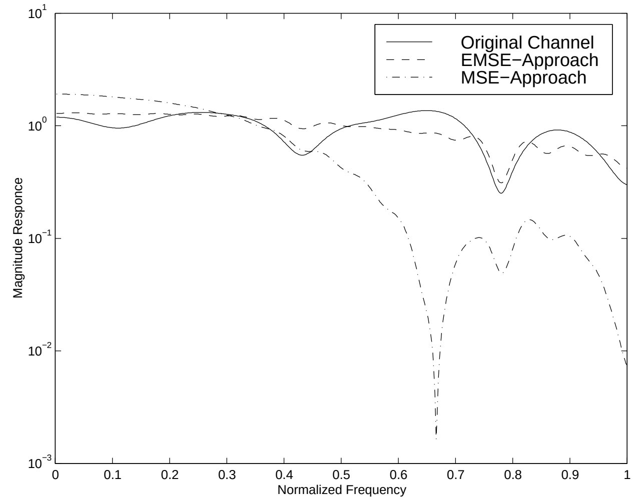

Fig. 2. Frequency response of the original and filtered channels.

In order to limit the complexity of the VA, a 51-tap linear pre-filter was used to equalize the received sequence prior to passing it to the VA in order to reduce the complexity of the VA. The VA implementation, therefore, will be based on the knowledge of the equivalent short DIR $\mathbf{q}$ with 4 taps only which means that only 2^3 states are required for VA computations. The tap gains of the linear pre-equalizer were designed such as to minimize the MSE between the output of the linear pre-filter and the output of a predefined DIR. The DIR itself was selected according to two different criteria: the conventional MSE criterion optimized using the approach of [3] and the EMSE criterion optimized using the approach of this paper. The BER performance curves are shown in Figure 3. From Figure 3, it can be seen that the EMSE approach outperforms the conventional MSE by about 5 dB at a BER of 10^{-5}. This is because, unlike the conventional MSE approach, the EMSE approach takes into account the effect of noise correlations resulting from the pre-filtering

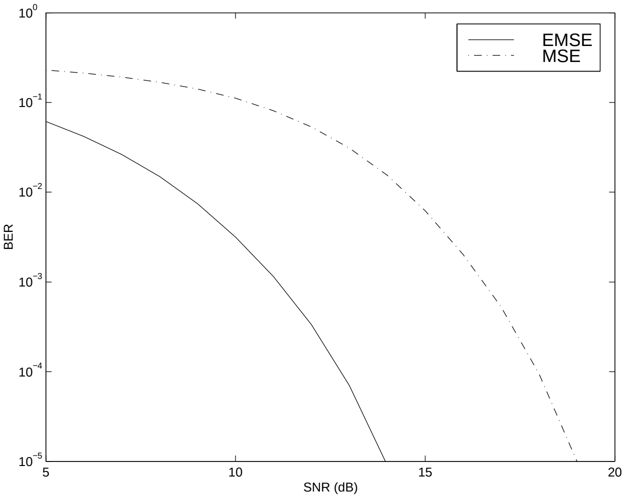

Fig. 3. BER performance comparison between MSE and EMSE optimization approaches.

process. This means that the amount of correlations introduced by EMSE-optimized filters will be considerably less than that of the MSE approach and, therefore, the performance degradation will be less. To illustrate the last point, the frequency responses of the original channel and the filtered channels using the MSE and EMSE approaches are presented in Figure 2. Figure 2 clearly shows that the spectrum of the EMSE-optimized channel closely matches that of the original channel. On the other hand, the MSE-optimized pre-filter introduces a spectral null at the normalized frequency of 0.68 causing a considerable performance loss.

VII. Conclusion

A new approach for optimizing CLVEs has been presented. This approach is based on using the GA to optimize the EMSE criterion. The main advantage of the EMSE approach is that, unlike conventional MSE approaches, it takes into account the effects of noise correlations and the Euclidean distance property of the overall impulse response. Simulation results of the new technique have shown that a significant improvement in system performance can be achieved over that of MSE approaches. This may be due to the fact that the DIR computed using the EMSE criterion is much closer to the original CIR than the DIR computed using the MSE criterion.

References

[1] G. D. Forney, Jr., “Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference,” *IEEE Trans. Inform. Theory*, vol. IT-18, no. 3, pp. 363-378, May 1972.

[2] N. Al-Dhahir and J. M. Cioffi, “Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach,” *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 903-915, May 1996.

[3] D. D. Falconer and F. R. Magee, Jr., “Adaptive channel memory truncation for maximum likelihood sequence estimation,” *The Bell System Technical Journal*, vol. 52, no. 9, pp. 1541-1562, November 1973.

[4] S. U. H. Qureshi and E. E. Newhall, “An adaptive receiver for data transmission over time-dispersive channels,” *IEEE Trans. Inform. Theory*, vol. IT-19, no. 4, pp. 448-457, July 1973.

[5] S. A. Fredricsson, “Joint optimization of transmitter and receiver filters in digital PAM systems with a Viterbi detector,” *IEEE Trans. Inform. Theory*, vol. IT-22, no. 2, pp. 200-210. March 1976.

[6] K. S. Tang, K. F. Man and Q. He, “Genetic algorithms and their applications,” *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 22-37, November 1996.

[7] A. J. Chipperfield and P. J. Flemming, “Parallel genetic algorithms: A survey,” *ACSE Research Report No. 518*, University of Sheffield, May 1994.

[8] M. Tomassini, “Parallel and distributed evolutionary algorithms: A review”, In *Evolutionary Algorithms in Engineering and Computer Science*, editors: K. Miettinen, M. Makela, P. Neittaanmaki, and J. Periaux, J. Wiley and Sons, Chichester, pp. 113-133, 1999.

[9] K. Twardowski, “An associative architecture for genetic algorithm-based machine learning,” *IEEE Comput.*, vol. 27, no. 11, pp. 27-38, November 1994.

[10] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann arbor University of Michigan Press, Ann Arbor, MI, 1975.

[11] COST207, *Final Report, COST207: Digital Land Mobile Radio Communication,* 1989.

[12] P. Hoeher, “A statistical discrete-time model for the WSSUS multipath channel,” *IEEE Trans. Vehic. Tech.*, vol. 41, no. 4, pp. 461-468, November 1992.