The Effect of a Precoder on Serially Concatenated Coding Systems with an ISI Channel

Inkyu Lee, *Member, IEEE*

***Abstract*—The performance of a serially concatenated system which includes a channel with memory preceded by a precoder as a rate-1 inner coder is presented. The effect of different precoders on the maximum-likelihood bit-error performance is analyzed. The *precoder weight gain*, which explains the good bit-error rate (BER) performance, is identified through a union bound analysis. Precoders are divided into two groups based on an analysis of the Euclidean distance and its multiplicity, and each precoder group shows a distinct BER curve behavior. It is shown that the BER curves for two precoder groups cross over each other. Convolutional codes are considered as outer codes in simulations on various intersymbol interference channels. Several important design considerations for the choice of precoders are derived based on the analysis and these are confirmed through simulations with an iterative decoding algorithm.**

***Index Terms*—Concatenated coding system, ISI channel, iterative decoding.**

I. Introduction

SINCE TURBO codes [1] were first introduced in 1993, concatenated coding systems in conjunction with iterative decoding have attracted great interest in the communications area. The impressive bit-error rate (BER) performance of parallel concatenated coding (PCC) systems employing a random interleaver has inspired people to consider several variations on its structure [2]–[4].

Benedetto *et al.* [4] proposed a serially concatenated coding (SCC) system, where two component encoders are connected serially through a random interleaver, and showed that the performance of the SCC is comparable to that of the PCC. In some situations, it was shown that SCCs do not exhibit an "error floor," the flattening region of the BER curve, which is normally observed in PCCs. In most of studies related to PCCs and SCCs, it is assumed that an encoded bit sequence is transmitted through a memoryless channel.

Recently, several researchers have proposed replacing the inner code of SCCs by other recursive structures [5]–[8]. In particular, Souvignier *et al.* [5], Öberg *et al.* [7], and Ghrayeb *et al.* [8] investigated the application of SCCs which view a channel with memory as a rate-1 inner code. In order for this system to provide the required recursive structure for the inner code, a precoder is placed in front of the channel. Narayanan [9] provides an interesting explanation on the effect of precoding on the convergence of the iterative process. However, these studies are mainly focused on particular partial response channels and are limited to applications in magnetic recording systems where code rates are usually higher than 8/9. Thus, they do not fully exhibit the effect of precoders which normally become distinctive with lower code rates. Moreover, in general communication systems, some analyses made in previous publications are no longer valid. For example, the "weight-two error events" analysis in [12] is only applicable to $\mathrm{SCC_{ch}}$'s with free distance equal to two.

In this paper, we investigate the effect of different precoders on general intersymbol interference (ISI) channels in binary modulation with various coding rates. Throughout this work, the serially concatenated system with the ISI channel will be referred to as $\mathrm{SCC_{ch}}$ to distinguish it from SCCs for memoryless channels. It can be noted that SCCs may also be applied to the ISI channel after converting the channel into a memoryless one using techniques such as equalization. However, this increases the decoder complexity considerably.

In general, the outer encoder in $\mathrm{SCC_{ch}}$'s could represent any encoding scheme. "Turbo equalization" [10], [11] that includes the channel demodulation in a decoding iteration can be viewed as an $\mathrm{SCC_{ch}}$, which takes a turbo code as an outer code. Because of the increased complexity of the outer decoder, an $\mathrm{SCC_{ch}}$ with a turbo code as an outer code is not considered in this paper. Instead, we consider only convolutional codes as an outer code.

The objective of this paper is to provide insights into the choice of precoders so as to derive design guidelines that are useful in more general and practical system designs, and not limited to magnetic recording applications. Through a union bound analysis, we introduce the *precoder weight gain* to characterize the performance gain due to the precoder. This gain is analogous to the "interleaver gain" observed in PCCs and SCCs [4]. This will explain the good BER performance of $\mathrm{SCC_{ch}}$'s, compared to the case when no precoder is used.

Later in the simulation section, it will be shown that the BER curves for some precoders cross over those for other precoders. So, one precoder reaches an error floor at a low signal-to-noise ratio (SNR), while other precoders result in a lower error floor at a higher SNR. Following the analytic approaches, precoders are divided into two groups where each group exhibits a distinctive BER curve. In this paper, rather than attempting to find a precoder which achieves the lowest error floor as was done in [12], we characterize each precoder group such that a proper precoder can be selected depending on the desired system performance. While the union bound analysis offers a good match for high

Paper approved by T. Aulin, the Editor for Coding and Communications Theory of the IEEEE Communications Society. Manuscript received May 24, 2000; revised December 13, 2000 and January 15, 2001. This paper was presented in part at the IEEE International Conference on Communications, New Orleans, LA, June 2000.

The author is with the Wireless Systems Research Laboratory, Agere Systems, Murray Hill, NJ 07974 USA (e-mail: inkyu@agere.com).

Publisher Item Identifier S 0090-6778(01)05782-8.

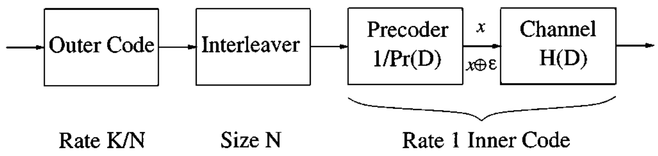

Fig. 1. Serially concatenated code system employing the channel with precoder.

SNRs, a heuristic analysis is still required to study BER curve behaviors for low SNRs. So, we derive design considerations based on both an analytical approach and actual simulations.

This paper is organized as follows. The following section analyzes the BER performance of $\mathrm{SCC_{ch}}$'s which adopt the ISI channel with a precoder as a rate-1 inner code. An ensemble average maximum-likelihood (ML) BER upper bound is derived using a union bound approach. In Section III, the asymptotic performance of $\mathrm{SCC_{ch}}$'s both with and without precoders is derived based on an actual random interleaver, not on an abstract uniform interleaver. Section IV investigates the effect of precoders on the $\mathrm{SCC_{ch}}$ performance and addresses the issues of the precoder choice which can better serve the performance requirement for a given ISI channel. Through analysis and computer simulations using iterative decoding techniques, we explain different BER behaviors depending on precoder groups and present some design considerations for the choice of precoders. Finally, Section V contains concluding remarks.

II. Performance of Serially Concatenated Codes with an ISI Channel

For simplicity, we consider a binary modulation scheme in this paper. Our derivation can be easily extended to higher-level modulation schemes. Consider a serially concatenated code system which takes the ISI channel $H(D)$ with a precoder $1/\Pr(D)$ as a rate-1 inner code as shown in Fig. 1. It is assumed that the discrete-time ISI channel model is obtained from the whitened matched-filter output [13]. Here, the outer encoder with free distance d^o_{free} has rate $R = K/N$, where K and N represent the length of input words and codewords, respectively. Thus, the size of the interleaver is equal to N. Also, note that $H(D)$ represents the transfer polynomial in the field of real numbers, whereas $\Pr(D)$ is a polynomial in a finite field. Between the precoder and the channel, a signal mapper is assumed.

As for the precoder structure, many different forms of precoders such as $\Pr_1(D)/\Pr_2(D)$ or Tomlinson–Harashima precoders [14], [15] could be used to provide the recursive nature for the inner code. However, these precoder structures require increased complexity in the inner decoder. Therefore, in this work, we limit our focus to $1/\Pr(D)$ precoders whose memory is equal to or smaller than the channel memory. With this condition, for a given channel response $H(D) = \sum_i h_i D^i$, the number of states in the inner decoder remains the same. This structure is illustrated in Fig. 2, which uses a $1/1 \oplus D \oplus D^3$ precoder as an example, where $\oplus$ represents addition modulo 2. A maximum *a posteriori* (MAP) detector serves as an inner decoder based on this channel description. Given the channel polynomial, we can define the norm of the channel $H(D)$ as $\|h\|^2 = \sum_i |h_i|^2$.

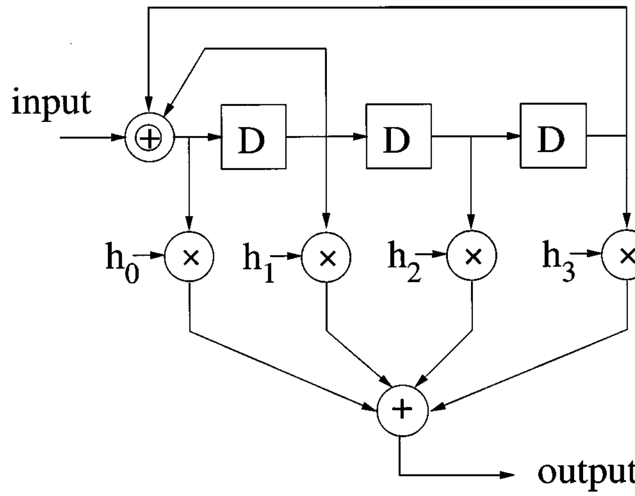

Fig. 2. Structure of channel $H(D)$ with $1/1 \oplus D \oplus D^3$ precoder.

For simplicity of the presentation, convolutional outer codes are viewed as their equivalent block code by terminating sequences of convolutional codes. In this system, codewords of $\mathrm{SCC_{ch}}$'s are defined as the precoder output words. Let x and ϵ be a correct codeword and an error codeword of length N, respectively. Then, the erroneous codeword $x' = x \oplus \epsilon$ at the precoder output generates the input error event $e = x - x'$ at the inner decoder. The probability of bit error caused by choosing $x \oplus \epsilon$ over x in the ISI channel corrupted by Gaussian noise with two-sided noise power spectral density σ^2 with an ML detector is $(w/K)\Pr\{\text{ML chooses } x \oplus \epsilon \text{ over } x\}$, where w denotes the input weight to the outer code.

For a given error codeword ϵ with weight h, 2^h different error events are possible depending on the transmitted codeword x. For example, assuming binary modulation with the input alphabet $\{0, 1\}$, an error codeword $\epsilon = \cdots 0110 \cdots$ can generate four error events $(\ldots 0, 1, 1, 0 \ldots)$, $(\ldots 0, 1, -1, 0 \ldots)$, $(\ldots 0, -1, 1, 0 \ldots)$, $(\ldots 0, -1, -1, 0 \ldots)$ depending on x. Because of the random interleaver, we can assume that each of the 2^h error events are equally probable for low weight h's which are of importance in our analysis. Now we can compute the probability of error in the ISI channel for all possible codewords x caused by the kth error codeword ϵ_k as

$$\Pr\{\text{ML chooses } x \oplus \epsilon_k \text{ over } x\} = \frac{1}{2^{h_k}} \sum_{n=1}^{2^{h_k}} Q\left(\frac{d(e_{k,n})}{2\sigma}\right)$$

where h_k represents the weight of the kth error codeword ϵ_k, $e_{k,n}$ specifies one of 2^{h_k} error events generated by the kth codeword, and $d(e_{k,n})$ denotes the Euclidean distance generated by the input error event $e_{k,n}$ in the ISI channel $H(D)$.

Since there are $2^K - 1$ nonzero error code words ϵ_k, using the union bound approach, the probability of bit error for $\mathrm{SCC_{ch}}$'s under ML decoding can be shown to have an upper bound of

$$P_b \leq \sum_{k=1}^{2^K-1} \frac{w_k}{K} \frac{1}{2^{h_k}} \sum_{n=1}^{2^{h_k}} Q\left(\frac{d(e_{k,n})}{2\sigma}\right). \tag{1}$$

Now we want to rearrange the above expression in terms of all error events. Let us first define $\mathcal{E}$ to be a set of all possible input error events for the ISI channel $H(D)$ where each error event takes on values $\{+1, 0, -1\}$. For notational convenience, $+$ and $-$ will be used instead of $+1$ and -1 for error event descriptions. An error event e can be uniquely decomposed into a concatenation of disjoint error subevents $e_i, i = 1, 2, \ldots m$. We will denote such an error event as $e = (e_1, e_2, \ldots, e_m)$. So, the squared Euclidean distance of e can be obtained by summing all the squared Euclidean distance of each subevent ($d^2(e) = \sum_{i=1}^{m} d^2(e_i)$). Note that the order of error subevents e_i's in e does not affect $d^2(e)$. For a given error event $e \in \mathcal{E}$, let us denote $E(e)$ as a set of codewords ϵ, which produce e. For simplicity, we assume every error event starts with $+$. If a codeword ϵ belongs to $E(e)$, then any shifted codewords are also members of the same set. For example, codewords generating an error event $+-$ are $E(+-) = \{110\cdots 0, 0110\cdots 0, \ldots, 0\cdots 011\}$.

Reorganizing the above expression with respect to error events $e \in \mathcal{E}$ yields

$$P_b \leq \sum_{e\in\mathcal{E}} Q\left(\frac{d(e)}{2\sigma}\right) \sum_{\epsilon\in E(e)} \frac{1}{2^{h(\epsilon)}} \frac{w(\epsilon)}{K}$$

where $d(e)$ is the Euclidean distance of the error event e, $w(\epsilon)$ represents the weight of input words which generate a codeword ϵ, and $h(\epsilon)$ denotes the weight of the codeword ϵ. We will use the notation $h(e)$ instead of $h(\epsilon)$, since the weight is the same for all codewords $\epsilon \in E(e)$.

From the above expression, we refer to the coefficient of the Q function, $\sum_{\epsilon\in E(e)} w(\epsilon)$, as $W(e)$. This represents the total weight of all input words which generate all codewords $\epsilon \in E(e)$. Combining the notations defined above, we now obtain a compact expression for the upper bound on the probability of bit error as

$$P_b \leq \sum_{e\in\mathcal{E}} \frac{W(e)}{K\cdot 2^{h(e)}} Q\left(\frac{d(e)}{2\sigma}\right). \qquad (2)$$

One important coefficient which affects the above expression is $1/2^{h(e)}$, and we will refer to the performance gain due to this factor as the *precoder weight gain*. This is unique to $\mathrm{SCC_{ch}}$'s which employ the ISI channel as an inner structure. This precoder weight gain is achieved when a precoder is employed before the channel and this will be clearly explained in Section III.

In order to evaluate the upper bound in (2), we need to obtain the total information weight $W(e) = \sum_{\epsilon} w(\epsilon)$ for all codewords $\epsilon \in E(e)$, and this is closely related to the choice of an interleaver. To evaluate the "average" random interleaver, we can adopt the uniform interleaver argument developed in [4].

We first start with an outer codeword with weight l. Exploiting the properties of the uniform interleaver, which maps a codeword with weight l at the output of the outer encoder into all of its distinct $\binom{N}{l}$ permutations, we obtain the probability that a random interleaver maps this codeword into any one of the codewords $\epsilon \in E(e)$ as $|E(e)|/\binom{N}{l}$ where $|E(e)|$ denotes the cardinality of a set $E(e)$. Also, for N much larger than the channel memory n_h of $H(D)$, neglecting the length of single error events compared to N, the size of the set $E(e)$ is upper bounded by $|E(e)| \leq \binom{N}{m(e)}$, where $m(e)$ represents the number of concatenated error subevents e_i's in e.

Let $l(e)$ denote the weight of the precoder input words which generate $\epsilon \in E(e)$. Then the total input weight $W(e)$ is upper bounded by

$$W(e) \leq \frac{\binom{N}{m(e)}}{\binom{N}{l(e)}} A^o_{l(e)}$$

where $A^o_{l(e)}$ represents the total input weight of outer codewords with weight $l(e)$.

Finally, using the binomial approximation $\binom{N}{n} \approx N^n/n!$, the probability of bit-error expression (2) becomes

$$P_b \leq \sum_{e\in\mathcal{E}} \frac{N^{m(e)-l(e)} \cdot l(e)! \cdot A^o_{l(e)}}{K \cdot m(e)! \cdot 2^{h(e)}} Q\left(\frac{d(e)}{2\sigma}\right).$$

This upper bound for the bit-error probability is obtained by applying a uniform interleaver argument, which represents the expected performance of the "average" interleaver. However, depending on the choice of specific interleavers, the BER performance of $\mathrm{SCC_{ch}}$'s may differ by an order of magnitude, and this was also reported in [7]. More importantly, this analysis based on the uniform interleaver assumptions fails to recognize the precoder weight gain factor $1/2^{h(e)}$ since $N^{m(e)-l(e)}$ appears to make bigger contributions to the upper bound. However, this applies only to an abstract uniform interleaver, and with a randomly generated actual interleaver, it is this factor $2^{h(e)}$ which plays a more important role in the BER performance. Therefore, rather than attempting to compute the upper bound based on the uniform interleaver, we will analyze the BER performance based on a randomly generated interleaver by actually counting $W(e)$ in (2) for the rest of this paper.

III. Union Bound Analysis Based on a Random Interleaver

This section computes the BER performance based on an actual random interleaver both with and without a precoder.

A. Error Rate Performance with no Precoder

We first present the asymptotic BER performance analysis for the case where there is no precoder ($\Pr(D) = 1$). Consider a nonsystematic convolutional code $[G_1(D) \quad G_2(D)]$ with generating polynomials $G_1(D) = 1 \oplus D^2$ and $G_2(D) = 1 \oplus D \oplus D^2$ as an outer code. The puncturing pattern listed in [16] is used to achieve a rate of 2/3 and the channel response $H(D) = 1 + 2D - 2D^3 - D^4$ is assumed with white Gaussian noise. Error events for this channel response $H(D)$ are tabulated in [17] and its minimum Euclidean distance is found to be $d^2(+-+) = 6$. With this outer encoding scheme, the outer code sequences corresponding to the free distance $d^o_{\mathrm{free}} = 3$ are $D^{3j}(D^2 \oplus D^4 \oplus D^5)$, $j \geq 0$. Given a particular interleaver of length $N = 1023$, it is found that outer codewords with weight $d^o_{\mathrm{free}} = 3$ generate four interleaver output words corresponding to the error event $e = (+, +00+)$ and 336 output words corresponding to the error event $e = (+, +, +)$. Therefore, their cor-

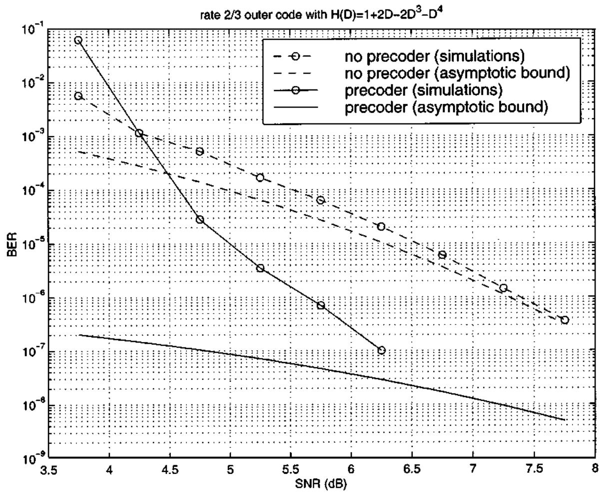

Fig. 3. Performance comparison between no precoder and $1/1 \oplus D$ precoder.

responding Euclidean distances are equal to $d^2(+,\, +00+) = d^2(+) + d^2(+00+) = 10 + 12 = 22$ and $d^2(+,\, +,\, +) = 3d^2(+) = 30$, respectively. Then, we obtain the asymptotic bit-error probability as

$$P_b \approx \frac{4}{682}\,\frac{2^2}{2^3}\,Q\left(\frac{\sqrt{22}}{2\sigma}\right) + \frac{336}{682}\,Q\left(\frac{\sqrt{30}}{2\sigma}\right)$$

where 682 is the input word length K.[1]

This asymptotic performance is dominated by the second term which is determined by three single error events $+$, because of its large multiplicity. This indicates that without a precoder, the error event e which consists of the error event $+$ dominates the asymptotic performance with $d^2(e) = d^o_{\rm free} d^2(+)$. Therefore, when the error event $+$ produces a small Euclidean distance for a given channel $H(D)$, the slope of the minimum distance asymptote becomes lower in $\mathrm{SCC_{ch}}$'s with no precoder.

This can be generalized to $\mathrm{SCC_{ch}}$'s with an outer code with rate $R = k_0/(k_0+1)$. Since we assume the input alphabet $\{0,\, 1\}$, the user bit energy E_b is equal to $1/2R$. Also, noting that $W(e) \approx K/k_0$ and $d^2(+) = \|h\|^2$, the asymptotic error rate for $\mathrm{SCC_{ch}}$'s with no precoder becomes

$$P_b \approx \frac{1}{k_0}\,Q\left(\sqrt{d^o_{\rm free}\,\frac{\|h\|^2}{\sigma^2}}\right) = \frac{1}{k_0}\,Q\left(\sqrt{\frac{k_0}{k_0+1}\,d^o_{\rm free}\cdot \mathrm{SNR}}\right)$$

where SNR is defined as $\|h\|^2(E_b/N_0)$ with $N_0 = 2\sigma^2$.

We plot the simulation results obtained by applying the iterative decoding techniques described in [4], [10], and [18] with ten iterations in Fig. 3. To incorporate the energy in the ISI channel $H(D)$, we define the SNR as $\|h\|^2(E_b/N_0)$ with $N_0 = 2\sigma^2$ and is used in the x axis. It should be noted that since the weight of channel input words $h(e)$ with no precoder is the same as that of outer codewords, $\mathrm{SCC_{ch}}$'s without a precoder are unable to generate the precoder weight gain $1/2^{h(e)}$.

[1]The 2^2 term accounts for the fact that there are four cases which have the same distance: $(+,\, +00+)$, $(+,\, -00-)$, $(-,\, +00+)$, and $(-,\, -00-)$.

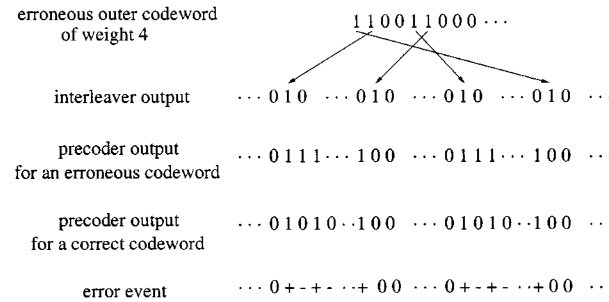

Fig. 4. Codeword illustration for $1/1 \oplus D$ precoder.

B. Error-Rate Performance with a Precoder

Let us consider an $\mathrm{SCC_{ch}}$ with a precoder $1/(1 \oplus D)$. Since input words with odd weights to $1/(1 \oplus D)$ precoder generate very large weights, only even weights input words are of our interest. Considering the same generating polynomials, puncturing pattern and $H(D)$ as in the no precoder case, we can easily find that there are four outer codewords of weight 4 and they are $D^{3j}(1 \oplus D \oplus D^3 \oplus D^5)$, $D^{3j}(D^2 \oplus D^3 \oplus D^6 \oplus D^7)$, $D^{3j}(D^2 \oplus D^4 \oplus D^7 \oplus D^8)$, $D^{3j}(D^2 \oplus D^3 \oplus D^9 \oplus D^{11})$, $j \geq 0$. These weight 4 sequences are permuted by the random interleaver, resulting in precoder input words with the form of $D^{i_1} \oplus D^{i_2} \oplus D^{i_3} \oplus D^{i_4}$ for $0 \leq i_1 < i_2 < i_3 < i_4$. Then, most of precoder output sequences are in the form of $D^{i_1} \oplus D^{i_1+1} \cdots \oplus D^{i_2-1} \oplus D^{i_3} \oplus D^{i_3+1} \cdots \oplus D^{i_4-1}$ and these sequences can support error events consisting of two error subevents with lengths $i_2 - i_1$ and $i_4 - i_3$. Among all error events e of length l, one with alternating signs (i.e., $e = +-+-\cdots+$) produces the smallest Euclidean distance for most channels $H(D)$. This process is illustrated in Fig. 4. The Euclidean distance for $+-+-\cdots$ is found to be 8 as long as the length of the error event is greater than 3 [17]. Therefore, the overall Euclidean distance for the precoder $1/(1 \oplus D)$ is equal to $d^2(e) = 2 \cdot 8 = 16$ as long as error subevents are longer than three symbols. Through an exhaustive search, it is found that this particular interleaver generates one dominant error event $e = (+-+,\, +-+-+-+)$ and the corresponding Euclidean distance is $d^2(e) = d^2(+-+) + d^2(+-+-+-+) = 6 + 8 = 14$. Note again that the probability that the precoder output codeword $\cdots 011 \cdots 10 \cdots$ of weight h supports the error event $+-+\cdots+$ is $1/2^h$. The same codeword can support other error events such as $--\cdots-$, but the corresponding Euclidean distance would be much higher, thus its contribution to the BER becomes negligible compared to the dominant error event. So, combining with $W(e) = 2$, the asymptotic performance approaches

$$P_b \approx \frac{2}{682}\,\frac{2^2}{2^{10}}\,Q\left(\frac{\sqrt{14}}{2\sigma}\right). \tag{3}$$

The simulation results for this case are also plotted in Fig. 3. Compared to no precoder case, the system with a precoder exhibits a larger separation between the asymptotic bound and simulation results. This is due to a fact that only one error event case is included when computing the asymptotic performance in (3). It is also interesting to see that at low SNRs, $\mathrm{SCC_{ch}}$ with no precoder achieves better performance than one with a precoder.

This becomes pronounced with more powerful outer codes with higher free distance d^o_{free}, as observed in [5] and [7].

Compared to the no precoder case, it is clear that SCC_{ch}'s with a precoder perform better even with a smaller Euclidean distance. This is mainly due to the coefficient of the Q function, $1/2^{h(e)}$, in (2), which is determined by weight of the precoder output codewords. Normally, a precoder can generate codewords with large weights even when the weight of the precoder input words is small. So, SCC_{ch}'s with a precoder can achieve a high precoder weight gain. This effect of the weight of h on the BER performance has not been addressed in other literatures on SCC_{ch}'s before. Also, this precoder weight gain is not observed in other serially concatenated structures. For example, the performance of SCCs in [4] is determined by the codeword weights and not by the weight distribution in a codeword. In other words, two different codewords with the same weight contribute the same amount to the BER of SCCs in [4], but could have different contributions to the BER in SCC_{ch}'s since they support different error events. A similar effect of the weight distribution in a codeword on the performance of a coded system is found in the continuous phase modulation (CPM) [19].

As shown in this example, for a given interleaver, the precoder $1/\Pr(D)$ affects the multiplicity while the ISI channel $H(D)$ determines the Euclidean distance of the overall case. In other words, the asymptotic slope of the BER curve is determined by the channel $H(D)$, and multiplicities of error events are lowered by the precoder weight gain. We can further improve the asymptotic BER performance by having precoders whose precoder output words yield a high Euclidean distance for the channel $H(D)$. The same goal can be achieved by enhancing the interleaver to avoid certain output streams which might generate a small Euclidean distance [18], [20].

IV. Effect of Different Precoders

We will now analyze the effect of the choice of different precoders. Assuming the memory of precoder $\Pr(D)$ does not exceed that of the channel $H(D)$, the inner decoder complexity is determined by the channel memory n_h of $H(D)$ and is independent of the precoder memory. Therefore, we focus on precoders whose memory is smaller than or equal to the channel memory. We now present SCC_{ch}'s with convolutional codes with various code rates. In this section, all simulations use the iterative decoding algorithm [4] with ten iterations. An interleaver can be designed to eliminate certain precoder input words which could result in small Euclidean distances. To this end, an S-random interleaver proposed in [21] is used in this simulation section to avoid precoder input words which can yield small Euclidean distances.

A. SCC_{ch} *with Convolutional Codes*

Consider a channel $H(D) = (1+D)(1-0.5D)(1+0.25D) = 1+0.75D-0.375D^2-0.125D^3$ and assume the same convolutional outer code polynomials used in Section III-A. The rate 4/5 puncturing pattern in [16] is used which yields $d^o_{\text{free}} = 2$. Since the dominant error events come from precoder output words with low weight, we focus on precoder output words with input weight two in this case. An S-interleaver with S set to 14 is generated. This means that the interleaver transforms input sequences with a form of $D^i(1 \oplus D^{S_1})$, $i \geq 0$ where $S_1 < 15$ into output sequences with the form of $D^j(1 \oplus D^{S_2})$, $j \geq 0$ where $S_2 \geq 15$.[2]

TABLE I
CODEWORDS GENERATED BY WEIGHT 2 INPUTS WITH $S = 14$ CONSTRAINT AND THEIR EUCLIDEAN DISTANCES IN $H(D)$

$Pr(D)$		Precoder Output/Euclidean distance
weight-	$1 \oplus D$	$1(1)^{14}/2.16, 1(1)^{15}/2.16, \cdots$
two	$1 \oplus D^2$	$1(01)^7/7.19, 1(01)^8/7.97, \cdots$
precoders	$1 \oplus D^3$	$1(001)^4/7.59, 1(001)^5/9.06, \cdots$
multi-	$1 \oplus D \oplus D^2$	$11(011)^4/10.28, 11(011)^5/12.25, \cdots$
weight	$1 \oplus D^2 \oplus D^3$	$10111(0010111)^2/9.06, 10111(0010111)^3/12, \cdots$
precoders	$1 \oplus D \oplus D^3$	$11101(0011101)^2/9.06, 11101(0011101)^3/12, \cdots$
	$1 \oplus D \oplus D^2 \oplus D^3$	$11(0011)^3/8.88, 11(0011)^4/11.03, \cdots$

Table I lists precoder output words corresponding to weight-two input words where two ones are separated by at least $S = 14$ zeros. Here a parenthesized string $(s)^i$ denotes i repetitions of the string s. For purpose of illustration, we divide precoders into two groups: weight-two precoders and multiweight precoders where weight-two precoders $\Pr(D)$ take the form of $1 \oplus D^n$ for $n > 1$. For each precoder output, Euclidean distances for all possible error events are considered and the smallest ones are listed among them. For example, a precoder output $1(01)^7$ could support both error events $+(0+)^7$ and $+(0-)^7$ which generate Euclidean distance 7.19 and 20.31 for $H(D)$, respectively. A similar analysis can be done by viewing the ISI channel with a precoder as a trellis code [12]. However, due to computational difficulties, this technique can become unmanageable for longer channel memories. Also, by treating the precoder and $H(D)$ separately as is done here, we can gain a clear view on how each error event associated with a channel $H(D)$ affects the performance in SCC_{ch}'s.

The bit-error performance of SCC_{ch}'s is dependent upon the actual interleaver used, since multiplicities corresponding to precoder output codewords are determined by the interleaver. Note that the Euclidean distances listed in Table I represent the case where a general S-interleaver with $S = 14$ is assumed. To get a more accurate estimate for the BER, we need to count the multiplicities based on the actual interleaver. One brute force way is to identify all the interleaver input words corresponding to the precoder output words listed in Table I. Then finally check if the interleaver input words are the valid codewords for the outer code, and compute the total input weight $W(e)$ and the corresponding precoder weight gain. Using this search algorithm, it is found that multiplicities for precoder output words corresponding to the precoders $\Pr(D) = 1 \oplus D$ and $1 \oplus D^2$ are much lower than those corresponding to the precoders $\Pr(D) = 1 \oplus D^2 \oplus D^3$ and $1 \oplus D \oplus D^3$ for a

[2]It should be noted that an S interleaver can also generate output sequences with the form of $D^j(1 \oplus D^{S_2})$, $j \geq 0$, where $S_2 < 15$ for input sequences with a form of $D^i(1 \oplus D^{S_1})$, $i \geq 0$, where $S_1 \geq 15$.

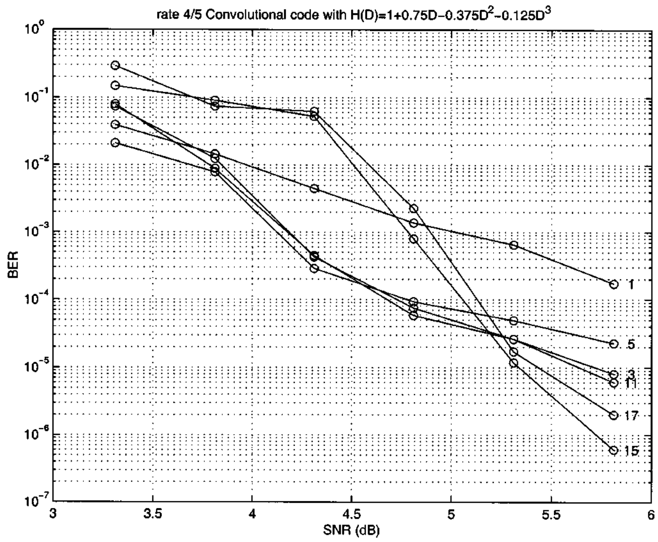

Fig. 5. Rate-4/5 convolutional outer code with different precoders.

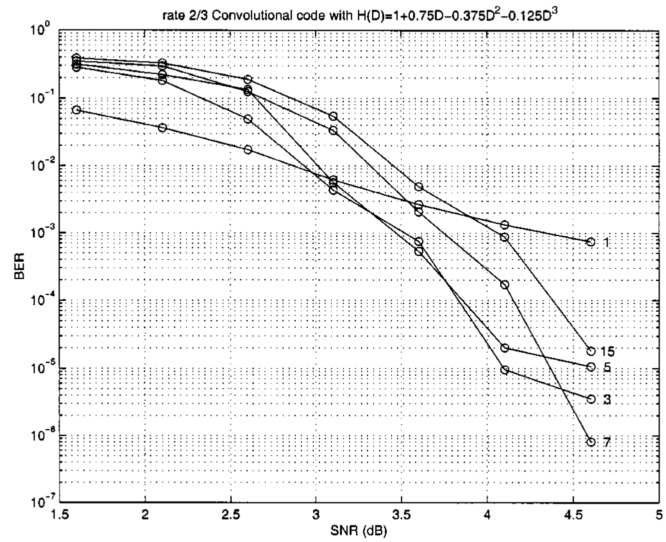

Fig. 6. Rate-2/3 convolutional outer code with different precoders.

given interleaver. Since Euclidean distances for the former are also smaller than those for the latter as shown in Table I, we expect that the BER curves for these precoders will cross over each other. This is confirmed by the simulation results shown in Fig. 5. Let pr denote the octal representation of $\Pr(D)$ (for example, $pr = 13$ in octal indicates $\Pr(D) = 1 \oplus D^2 \oplus D^3$). Numbers on each curve in figures indicate pr. The size of the random interleaver is set to $N = 500$. This small N is chosen in order to have error floors higher than $\mathrm{BER} = 10^{-7}$, which is often a simulation limit. Also, in many practical systems, a small size interleaver is employed to avoid a latency problem.

Two interesting observations can be made from this plot. First, for low to moderate SNRs, weight-two precoders $pr = 3, 5, 11$ perform better than multiweight precoders, but as SNR increases, the BERs of multiweight precoders such as $pr = 15, 17$ decrease rapidly and eventually outperform the others. This is exactly what we expect from the Euclidean distance analysis of Table I which determines the asymptotic BER slope. Second, the slopes of weight-two precoders are similar to that of the no precoder case ($pr = 1$). This indicates that the minimum Euclidean distances for these cases are all similar, while the multiplicities of the minimum distance error event for $pr = 3, 5, 11$ are lowered by the precoder weight gain. Therefore, the BER curves of weight-two precoders appear to be shifted down by that gain compared to the no precoder curve.

Figs. 6 and 7 show more simulation plots with code rates set to 2/3 and 1/2, respectively. For the rate-2/3 outer code in Fig. 6, again the puncturing pattern in [16] is used. These plots show similar BER curve patterns as in Fig. 5. The BER curves of $\mathrm{SCC_{ch}}$'s with weight-two precoders $pr = 3, 5$ start to flatten as SNR increases, while multiweight precoders $pr = 7, 13, 15$ exhibit quite a steep slope for high SNRs. Another interesting point which we can observe is that the crossover points of the BER curves at which multiweight precoders $pr = 13, 15, 17$ start to outperform weight-two precoders $pr = 3, 5$ occur at lower BERs as the code rate decreases. This means that as more powerful codes are used in outer codes, multiplicities of error

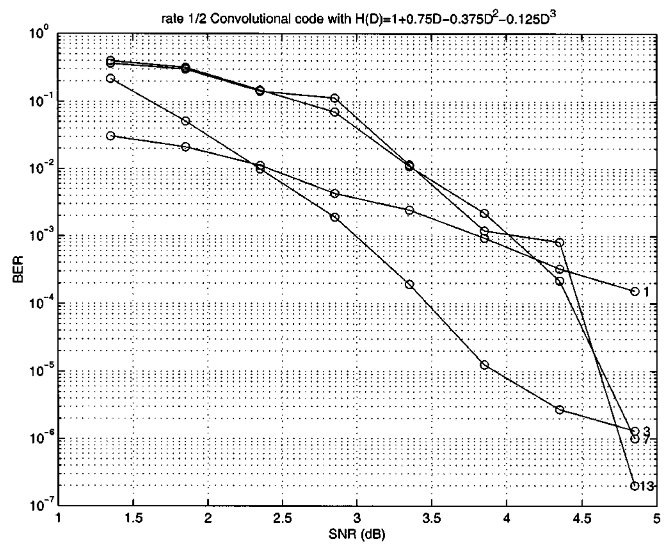

Fig. 7. Rate-1/2 convolutional outer code with different precoders.

events realized in weight-two precoders $pr = 3, 5$ become much smaller because of the increased precoder weight gain.

In applications in previous work [5], [12], it may be sufficient to consider only the error floor regions, since SNRs required for the "cliff," or "waterfall," regions for various precoders do not differ by much. However, this is not the case for general communication systems, and it is obvious that depending on the precoder, SNRs to achieve $\mathrm{BER} = 10^{-4}$ could be different by one decibel, as shown in Fig. 7. There are no precoders which are the optimum choice for both the cliff region and the error floor region. Therefore, we need to consider the overall BER curve to determine the best choice of precoders for the desired system performance.

It has been observed that as the interleaver size increases, differences in the BER curve behaviors in two precoder groups become less distinct, especially in the "cliff" region, since the overall slope is getting steeper. Note that this curve pattern analysis for each precoder group holds for a given channel $H(D)$.

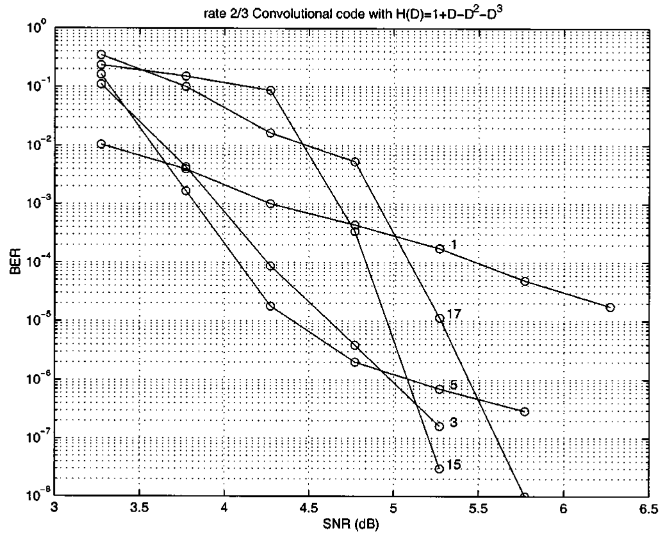

Fig. 8. Rate 2/3 convolutional outer code with different precoders.

With different choices of the channel, one can analyze its Euclidean distance and multiplicity for each precoder as done in this example. However, this peculiar BER behavior of one group of precoders crossing over the other group of precoders for high SNRs is observed in other choice of channels as well. This is again due to the fact that multiplicities of one group of precoders are larger than those of the other group of precoders in general, while the minimum Euclidean distance of the former is greater than that of the latter. Therefore, a cross over point always exists in the BER curves.

Fig. 8 shows simulation results for the channel $H(D) = 1 + D - D^2 - D^3$ and an interleaver size of $N = 2000$. For this simulation, 100 000 000 symbols were processed for the Monte Carlo BER count to show very low BER region. Again, similar BER behaviors for different precoders are obtained when compared to previous simulation plots. For example, the BER curve of the precoder $pr = 5$ clearly exhibits the precoder weight gain over the no precoder case, while both cases show the same slope at high SNRs. Based on this plot, we can note that for system applications aiming for a BER higher than 10^{-6}, the precoder $pr = 5$ is better suited, while a precoder $pr = 15$ is a proper choice to provide very low error rate performance. The reason why $pr = 15$ has such a steep asymptote can be explained in line with observations made in the previous examples. When low weight precoder input words are divided by $\Pr(D) = 1 \oplus D \oplus D^3$, many codewords would contain $11\,101(0\,011\,101)^i$, $i \geq 0$, and codewords containing $11\,101(0\,011\,101)^i$ generate very high Euclidean distance for the corresponding error events. In contrast, $\Pr(D) = 1 \oplus D^2$ generates codewords which contain $1(01)^i$ and this supports the error event $e = +(0+)^i$ with $d^2(e) = 4$, which is the minimum Euclidean distance for $H(D)$ with no precoder, thus resulting in an asymptotically lower slope.

The channel responses assumed in this example contain nulls in the frequency domain at DC and the Nyquist frequency, thus resulting in "zero-cycles" in their error sequence [17]. Because of the "zero-cycles," each error event may contain infinite repetitions of a string without changing the overall distance properties. For example, with $H(D) = 1 + D - D^2 - D^3$, the error events $e = +(-+)^i$ yield the same Euclidean distance regardless of the number of repetition i, and this behavior is called *quasi-catastrophic* [22]. Therefore, even if an S-interleaver with a very high S is employed, some precoders could still get a small Euclidean distance. If the channel $H(D)$ has no frequency nulls, then "zero-cycles" would be removed from its error sequences, and longer error sequences would tend to produce larger Euclidean distances.

B. Summary of Observations

Based on a few observations made in this simulation section, we can draw some design considerations as follows.

- The performance of $\mathrm{SCC_{ch}}$'s with no precoder is dominated by d^o_{free} error events $+$. In other words, the asymptotic slope with no precoder is determined by $d^2(+) = \|h\|^2$ for $H(D)$.
- When a precoder is employed, the asymptotic BER performance is improved by the precoder weight gain, which results from transforming low weight outer codewords into ones with high weights by precoders. This precoder weight gain lowers the multiplicity of the minimum Euclidean distance error event. The more powerful the outer codes $\mathrm{SCC_{ch}}$'s employ, the greater the precoder weight gain becomes in comparison to no precoder case.
- Weight-two precoders and multiprecoders exhibit distinct BER curve behaviors. Note that weight of precoder polynomials does not necessarily determine the BER curve behavior. Criteria of dividing precoders into two groups are dependent upon $H(D)$. Also, multiplicities for error events are determined by the actual interleaver. It has been found that, in general, weight-two precoders yield lower multiplicities than multiweight precoders and that Euclidean distances for the former are smaller than those for the latter. Therefore, the BER curves of weight-two and multiweight precoders normally cross over each other, and for high SNRs, multiweight precoders outperform weight-two precoders.
- The crossover point in the BER curve at which multiweight precoders start to outperform weight-two precoders becomes lower as more powerful outer codes are used. Therefore, the choice of precoders depends upon the target BER region relative to this crossover point. For example, when one needs to use simple outer codes for decoder complexity reasons or needs to employ a high rate system, multiweight precoders are suitable, since the BER crossover point would normally be higher than the target BER. On the other hand, weight-two precoders are preferred in applications which require an error floor at a low SNR, or in systems with a low coding rate.
- It is clear from simulations that primitive polynomials are not necessarily the best choice for precoders, which was originally suggested in [23].

In conclusion, we can better serve the system performance requirement by the careful choice of precoders. Since the BER curves for weight-two and multiweight precoders are quite different, one should choose a precoder depending on the performance requirement. To this end, one first needs to identify Euclidean distances and multiplicities of minimum distance error events for each precoder for a given $H(D)$ and an interleaver. After that, the target BER should be considered to determine the best precoder. In general, when powerful outer codes are employed such that the crossover point in the BER is much lower than the target BER, precoders with higher error floors could be used. In contrast, precoders which yield much larger Euclidean distance are more suitable when trying to provide near error-free performance, because of the steep asymptote.

V. Conclusions

We have presented a serially concatenated code system that takes a general ISI channel as a rate-1 inner code. Through the union bound analysis, we have identified the precoder weight gain which explains much smaller multiplicities of error events in precoders compared to no precoder case.

The effect of the choice of different precoders is also analyzed. We showed that precoders are generally divided into two groups based on their BER curve behaviors. Due to differences in Euclidean distances and multiplicities, the BER curves for two groups cross over each other. Through several simulations and Euclidean distance analysis, some important design considerations regarding the choice of precoders are drawn. These are new observations which are not available in previously studied results. Based on analysis and guidelines derived in this paper, we can better understand and predict the BER behavior of the serially concatenated code systems and can choose precoders which better serve the system performance requirement.

Acknowledgment

The author would like to thank T. Richardson for helpful discussions and for proofreading the first draft of this paper. The author would also like to thank the anonymous reviewers and the Editor, Dr. T. Aulin, for valuable comments that led to a greatly improved manuscript.

References

[1] A. G. C. Berrou and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Communications (ICC'93)*, May 1993, pp. 1064–1070.

[2] P. Robertson and T. Wörz, "A novel bandwidth efficient coding scheme employing turbo codes," in *Proc. IEEE Int. Conf. Communications (ICC'96)*, June 1996, pp. 962–967.

[3] R. Lucas, M. Bossert, and M. Breitbach, "Iterative soft-decision decoding of linear binary block codes," in *Proc. IEEE Int. Symp. Information Theory Applications*, Sept. 1996, pp. 811–814.

[4] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 909–926, May 1998.

[5] T. Souvignier *et al.*, "Turbo decoding for PR4: Parallel vs. serial concatenation," in *Proc. IEEE Int. Conf. Communications (ICC'99)*, June 1999, pp. 1638–1642.

[6] K. R. Narayanan and G. L. Stüber, "A serial concatenation approach to iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 47, pp. 956–961, July 1999.

[7] M. Öberg and P. H. Siegel, "Performance analysis of turbo-equalized dicode partial-response channel," in *Proc. 36th Annu. Allerton Conf. Communication, Control and Computing*, Sept. 1998, pp. 230–239.

[8] A. Ghrayeb and W. E. Ryan, "Concatenated code system design for storage channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 709–718, Apr. 2001.

[9] K. R. Narayanan, "Effect of precoding on the convergence of turbo equalization for partial response channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 686–698, Apr. 2001.

[10] D. Raphaeli and Y. Zarai, "Combined turbo equalization and turbo decoding," *IEEE Commun. Lett.*, vol. 2, pp. 107–109, Apr. 1998.

[11] W. Ryan, L. McPheters, and S. McLaughlin, "Combined turbo coding and turbo equalization for PR4 equalized Lorentzian channels," in *Proc. Conf. Information Systems and Sciences*, 1998.

[12] W. E. Ryan, "Concatenated codes for class IV partial response channels," *IEEE Trans. Commun.*, vol. 47, pp. 445–454, Mar. 2001.

[13] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.

[14] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electron. Lett.*, vol. 7, pp. 138–139, Mar. 1971.

[15] H. Harashima and H. Miyakawa, "Matched-transmission techniques for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. COM-20, pp. 774–780, Aug. 1972.

[16] D. Haccoun and G. Begin, "High-rate punctured convolutional codes for Viterbi and sequential decoding," *IEEE Trans. Commun.*, vol. 37, pp. 1113–1125, Nov. 1989.

[17] S. A. Altekar *et al.*, "Error events characterization on partial response channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 241–247, Jan. 1999.

[18] W. E. Ryan. A turbo code tutorial. [Online]. Available: http://www.ece.-arizona.edu/~ryan.

[19] P. Moqvist, "Serially concatenated systems: An iterative decoding approach with applications to continuous phase modulation," Ph.D. dissertation, Chalmers Univ. Technol., Göteborg, Sweden, 1999.

[20] O. F. Açikel and W. E. Ryan, "Punctured turbo-codes for BPSK/QPSK channels," *IEEE Trans. Commun.*, vol. 47, pp. 1315–1323, Sept. 1999.

[21] D. Divsalar and F. Pollara, "Turbo codes for PCS applications," in *Proc. IEEE Int. Conf. Communications (ICC'95)*, June 1995.

[22] G. D. Forney and A. R. Calderbank, "Coset codes for partial response channels; or, coset codes with spectral nulls," *IEEE Trans. Inform. Theory*, vol. 35, pp. 925–943, Sept. 1989.

[23] W. Ryan, "Design considerations for concatenating convolutional codes with partial response channels," in *Proc. Annu. Allerton Conf. Communication, Control and Computing*, 1999.

Inkyu Lee (S'91–M'95) was born in Seoul, Korea, in 1967. He received the B.S. degree (with honors) in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1990. He received the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1992 and 1995, respectively.

From 1991 to 1995, he was a Research Assistant at the Information Systems Laboratory at Stanford University. In 1995, he joined AT&T Bell Laboratories, which later become Bell Laboratories, Lucent Technologies. He is currently with the Wireless Systems Research Laboratory, Agere Systems (formerly Microelectronics Group at Lucent Technologies), Murray Hill, NJ. He has nine patents pending. His research interests include digital communications, signal processing and coding techniques applied to wireless systems, digital subscriber lines and magnetic recording channels.

Dr. Lee is an Editor for Wireless Communication Theory of the IEEE Transactions on Communications.