Interleaver Design for Turbo Codes

Hamid R. Sadjadpour, *Senior Member, IEEE*, Neil J. A. Sloane, *Fellow, IEEE*, Masoud Salehi, and Gabriele Nebe

***Abstract*—The performance of a Turbo code with short block length depends critically on the interleaver design. There are two major criteria in the design of an interleaver: the distance spectrum of the code and the correlation between the information input data and the soft output of each decoder corresponding to its parity bits. This paper describes a new interleaver design for Turbo codes with short block length based on these two criteria. A deterministic interleaver suitable for Turbo codes is also described. Simulation results compare the new interleaver design to different existing interleavers.**

***Index Terms*—Concatenated codes, convolutional codes, turbo codes.**

I. Introduction

TURBO codes [1] have an impressive near-Shannonlimit error correcting performance. The superior performance of Turbo codes over convolutional codes is achieved only when the length of the interleaver is very large, on the order of several thousand bits. For large block size interleavers, most random interleavers perform well. On the other hand, for some applications, it is preferable to have a deterministic interleaver, to reduce the hardware requirements for interleaving and deinterleaving operations. One of the goals of this paper is to propose a deterministic interleaver design to address this problem. For short interleavers, the performance of the Turbo code with a random interleaver degrades substantially up to a point where its bit error rate (BER) performance is worse than the BER performance of convolutional codes with similar computational complexity. For short block length interleavers, selection of the interleaver has a significant effect on the performance of the Turbo code. In many applications, such as voice, delay is an important issue in choosing the block size. For these applications, there is a need to design short block size interleavers that demonstrate acceptable BER performance. Several authors have suggested interleaver designs for Turbo codes suitable for short block sizes [2]–[5].

There are two major criteria in the design of an interleaver: 1) the distance spectrum properties (weight distribution) of the code, and 2) the correlation between the soft output of each decoder corresponding to its parity bits and the information input data sequence. Criterion 2 is sometimes referred to as the iterative decoding suitability (IDS) criterion [2]. This is a measure of the effectiveness of the iterative decoding algorithm and the fact that if these two data sequences are less correlated, then the performance of the iterative decoding algorithm improves.

The performance of Turbo codes at low BER is mainly dominated by the minimum effective free distance ($d_{\min}$) [13], [16]. It has been shown [6] that the Turbo code asymptotic performance approaches the $d_{\min}$ asymptote. The noise floor that occurs at moderate to high signal-to-noise ratios (SNRs) is the result of small $d_{\min}$ [6]. The noise floor can be lowered by increasing either the interleaver size or $d_{\min}$. Increasing interleaver block size (N) can increase $d_{\min}$. Increasing $d_{\min}$ can be achieved (when N is fixed) by appropriate choice of interleaver. In our approach, maximizing $d_{\min}$ is a goal in designing the interleaver.

Performance evaluation of Turbo codes is usually based on the assumption that the receiver is a maximum likelihood (ML) decoder. However, Turbo codes actually use a suboptimal iterative algorithm. A soft output decoding algorithm such as maximum *a posteriori* probability (MAP) [7] is used in the iterative algorithm. The performance of iterative decoding improves if the information that is sent to each decoder from the other decoders is less correlated with the input information data sequence. Hokfelt *et al.* [2] proposed the IDS criterion for designing an interleaver. In the interleaver design proposed here, we recommend the use of the IDS criterion with some modifications.

Trellis termination of Turbo codes is critical, especially when the interleaver is designed to maximize $d_{\min}$. If this problem is not addressed in the design of the interleaver, it can lead to a very small value for $d_{\min}$ because of the existence of data sequences with no trellis termination and low output weight, resulting in a degradation in the performance of the Turbo code. References [8]–[10] have addressed this question.

The paper is organized as follows. In Section II, random and S-random interleavers [11] are described. Our approach is based on S-random interleavers. The IDS [2] criterion is also briefly discussed. In Section III, a two-step S-random interleaver design is presented. Our approach requires knowing which polynomials are divisible by a primitive polynomial; this question is addressed in the Appendix. Section IV describes a deterministic interleaver design based on the results from Section III. We conclude the paper by comparing the BER performance of Turbo codes utilizing our interleaver design to other interleavers.

II. Problem Statements

An interleaver π is a permutation $i \mapsto \pi(i)$ that changes the order of a data sequence of N input symbols $d_1, d_2, \ldots, d_N$. If the input data sequence is $\mathbf{d} = [d_1, d_2, \ldots, d_N]$, then the permuted data sequence is $\mathbf{d}P$, where P is an interleaving matrix with a single one in each row and column, all other entries being zero. Every interleaver has a corresponding deinterleaver (π^{-1})

Manuscript received April 21, 2000; revised February 22, 2001.

H. R. Sadjadpour and N. J. A. Sloane are with AT&T Shannon Labs, Florham Park, NJ 07932-0971 USA(e-mail: sadjadpour@research.att.com; njas@research.att.com).

M. Salehi is with the Department of Electrical Engineering, Northeastern University, Boston, MA 02115 USA.

G. Nebe is with the Abteilung Reine Mathematik, Universität Ulm, Ulm D-89069, Germany.

Publisher Item Identifier S 0733-8716(01)02556-2.

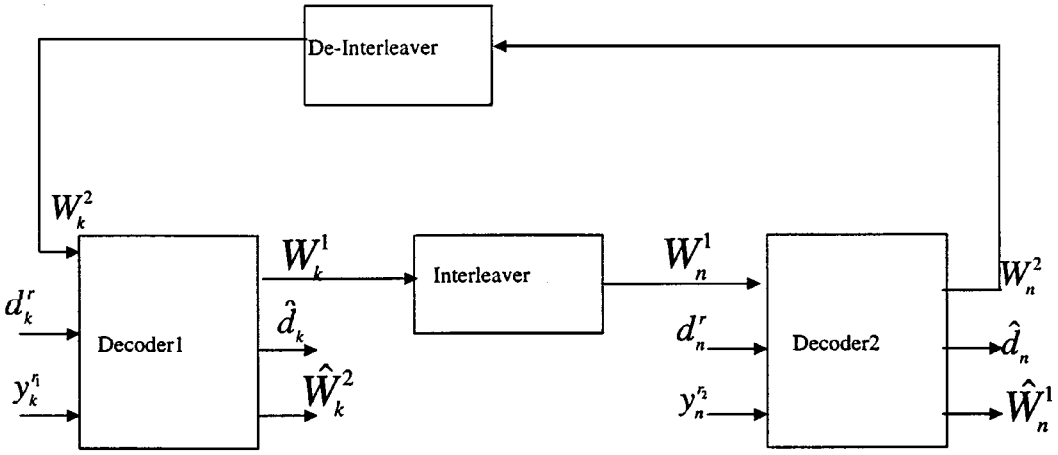

Fig. 1. Structure of a Tubo decoder.

that acts on the interleaved data sequence and restores it to its original order. The deinterleaving matrix is simply the transpose of the interleaving matrix (P^T).

A random interleaver is simply a random permutation π. For large values of N, most random interleavers utilized in Turbo codes perform well. However, as the interleaver block size decreases, the performance of a Turbo code degrades substantially, up to a point when its BER performance is worse than that of a convolutional code with similar computational complexity. Thus, the design of short interleavers for Turbo codes is an important problem [2]–[5].

An S-random interleaver (where $S = 1, 2, 3, \ldots$) is a "semirandom" interleaver constructed as follows. Each randomly selected integer is compared with S previously selected random integers. If the difference between the current selection and S previous selections is smaller than S, the random integer is rejected. This process is repeated until N distinct integers have been selected. Computer simulations have shown that if $S \leq \sqrt{N/2}$, then this process converges [11] in a reasonable time. This interleaver design assures that short cycle events are avoided. A short cycle event occurs when two bits are close to each other both before and after interleaving.

A new interleaver design was recently proposed based on the performance of iterative decoding in Turbo codes [2]. Turbo codes utilize an iterative decoding process based on the MAP or other algorithms that can provide a soft output. At each decoding step, some information related to the parity bits of one decoder is fed into the other decoder together with the systematic data sequence and the parity bits corresponding to that decoder. Fig. 1 shows this iterative decoding scheme. The inputs to each decoder are the input data sequence, d_k, the parity bits y_k^1 or y_k^2, and the logarithm of the likelihood ratio (LLR) associated with the parity bits from the other decoder (W_k^1 or W_k^2), which is used as *a priori* information. All these inputs are utilized by the decoder to create three outputs corresponding to the weighted version of these inputs. In Fig. 1, $\hat{d}_k$ represents the weighted version of the input data sequence, d_k. Also d_n in the same figure demonstrates the fact that the input data sequence is fed into the second decoder after interleaving. The input to each decoder from the other decoder is used as *a priori* information in the next decoding step and corresponds to the weighted version of the parity bits. This information will be more effective in the performance of iterative decoding if it is less correlated with the input data sequence (or interleaved input data sequence). Therefore, it is reasonable to use this as a criterion for designing the interleaver. For large block size interleavers, most random interleavers provide a low correlation between W_k^i and input data sequence, d_k. The correlation coefficient, $r^1_{W^1_{k_1}, d_{k_2}}$, is defined as the correlation between $W^1_{k_1}$ and d_{k_2}. It has been shown [2] that $r^1_{W^1_{k_1}, d_{k_2}}$ can be analytically approximated by

$$\hat{r}^1_{W^1_{k_1}, d_{k_2}} = \begin{cases} a \exp^{-c|k_1 - k_2|}, & \text{if } k_1 \neq k_2 \\ 0, & \text{if } k_1 = k_2 \end{cases} \tag{1}$$

where a and c are constants that depend on the encoder feedback and feedforward polynomials. The correlation coefficient at the output of the second decoder, $\mathbf{r}^2_{\mathbf{W}^2, \mathbf{d}}$, is approximated by

$$\hat{\mathbf{r}}^2_{\mathbf{W}^2, \mathbf{d}} = \tfrac{1}{2}\, \hat{\mathbf{r}}^1_{\mathbf{W}^1, \mathbf{d}} P \left(I + \hat{\mathbf{r}}^1_{\mathbf{W}^1, \mathbf{d}}\right) \tag{2}$$

where the two terms in the righthand side of (2) correspond to the correlation coefficients between $\mathbf{W}^2$ and the input data, i.e., $\mathbf{d}$ and $\mathbf{W}^1$ [2]. In our notation, $\hat{\mathbf{r}}^2_{\mathbf{W}^2, \mathbf{d}}$ represents the correlation coefficient matrix and $\hat{r}^2_{W^2_{k_1}, d_{k_2}}$ represents one element of this matrix.

Similar correlation coefficients can be computed for the deinterleaver. The correlation matrix corresponding to de-interleaver, $\hat{\mathbf{r}}'^2_{\mathbf{W}^2, \mathbf{d}}$, is the same as (2) except that P is replaced by P^T.

Then $\mathbf{V}_{k_1}$ is defined to be

$$\mathbf{V}_{k_1} = \frac{1}{N-1} \sum_{k_2=1}^{N} \left(\hat{r}^2_{W^2_{k_1}, d_{k_2}} - \overline{\hat{r}}^2_{W^2_{k_1}, \mathbf{d}} \right)^2 \tag{3}$$

where

$$\overline{\hat{r}}^2_{W^2_{k_1}, \mathbf{d}} = \frac{1}{N} \sum_{k_2=1}^{N} \hat{r}^2_{W^2_{k_1}, d_{k_2}}. \tag{4}$$

V'_{k_1} is defined in a similar way using $\hat{\mathbf{r}}'^2_{\mathbf{W}^2,\mathbf{d}}$. The iterative decoding suitability (IDS) measure is then defined as

$$\text{IDS} = \frac{1}{2N}\sum_{k_1=1}^{N}\left(V_{k_1}+V'_{k_1}\right). \tag{5}$$

A low value of IDS is an indication that the correlation properties between $\mathbf{W}^1$ and $\mathbf{d}$ are equally spread along the data sequence of length N. An interleaver design based on the IDS condition is proposed in [12].

III. Two-Step S-Random Interleaver Design

A new interleaver design, a two-step S-random interleaver, is presented here. The goal is to increase the minimum effective free distance, $d_{\min}$, of the Turbo code while decreasing or at least not increasing the correlation properties between the information input data sequence and W^i_k. Hokfelt *et al.* [2], [12] introduced the IDS criterion to evaluate the correlation properties. The two vectors for the computation of IDS in (5) are very similar for most interleavers. Thus, it is sufficient to only use one of them, i.e., V_{k_1}. Instead, we can define a new criterion based on decreasing the correlation coefficients for the third decoding step, i.e., the correlation coefficients between extrinsic information from the second decoder and information input data sequence. In this regard, the new correlation coefficient matrix, $\hat{\mathbf{r}}'^2_{\mathbf{W}^2,\mathbf{d}}$, is defined as

$$\begin{aligned}\hat{\mathbf{r}}'^2_{\mathbf{W}^2,\mathbf{d}} &= \tfrac{1}{2}\,\hat{\mathbf{r}}^2_{\mathbf{W}^2,\mathbf{d}}P^T\left(I+\hat{\mathbf{r}}^2_{\mathbf{W}^2,\mathbf{d}}\right)\\ &= \tfrac{1}{4}\left(\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}+\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}P\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}P^T\right)\\ &\quad\times\left(I+\tfrac{1}{2}\,\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}P+\tfrac{1}{2}\,\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}P\hat{\mathbf{r}}^1_{\mathbf{W}^1,\mathbf{d}}\right).\end{aligned} \tag{6}$$

$V'^{(\text{new})}_{k_1}$ can now be computed in a similar way to (3) by using (6). The new iterative decoding suitability (IDS_1) is then defined as

$$\text{IDS}_1 = \frac{1}{2N}\sum_{k_1=1}^{N}\left(V_{k_1}+V'^{(\text{new})}_{k_1}\right). \tag{7}$$

A small value for IDS_1 only guarantees that the correlation properties are spread equally throughout the data sequence. However, this criterion does not attempt to reduce the power of correlation coefficients, i.e., $(\hat{r}^2_{W^2_{k_1},d_{k_2}})^2$ and $(\hat{r}'^2_{W^2_{k_1},d_{k_2}})^2$. Therefore, we recommend the following additional condition as a second iterative decoding suitability criterion

$$\text{IDS}_2 = \frac{1}{2N^2}\sum_{k_1=1}^{N}\sum_{k_2=1}^{N}\left(\left(\hat{r}^2_{W^2_{k_1},d_{k_2}}\right)^2+\left(\hat{r}'^2_{W^2_{k_1},d_{k_2}}\right)^2\right). \tag{8}$$

We then use the average of these two values as a new IDS criterion, namely

$$\text{IDS}_{(\text{new})} = \tfrac{1}{2}\left(\text{IDS}_1+\text{IDS}_2\right). \tag{9}$$

Minimizing (9) is then one of our goals in optimizing the interleaver.

As we described earlier, S-random interleavers avoid short cycle events. This property guarantees that two bits close to each other before interleaving will have a minimum distance of S after interleaving. More specifically, for information input data i and j, and permuted data $\pi(i)$ and $\pi(j)$, an S-random interleaver will guarantee that if $|i-j|\le S$, then $|\pi(i)-\pi(j)|>S$. However, this does not exclude the possibility that $\pi(j)=j$, which can degrade the performance of iterative decoding of Turbo codes for this particular bit. The larger the distance between j and $\pi(j)$, the smaller the correlation between the information input data sequence and W^i_k. We therefore introduce an additional measure, S_2, which is defined to be the minimum permissible distance between j and $\pi(j)$ for all $j=1,2,\ldots,N$.

Unlike [12], where the interleaver design is based just on the IDS criterion, our interleaver is designed in two stages. In the first stage, we design an interleaver that satisfies the S-random criterion together with the S_2 condition. In the second stage, we try to increase the minimum effective free distance ($d_{\min}$) of the Turbo code while considering the $\text{IDS}_{(\text{new})}$ constraint. The design is as follows. We begin by selecting some values for S_1 and S_2.

Step 1) Each randomly selected integer $\pi(i)$ is compared with the previous selections $\pi(j)$ to check that if $i-j\le S_1$ then $|\pi(i)-\pi(j)|>S_1$. We also insist that π must satisfy $|i-\pi(i)|>S_2$.

Besides the above conditions, the last m tail bits used for trellis termination in the first decoder are chosen to satisfy $\pi(1)=N$, and if $\pi(i)=N-k$ with $k<m$ then $i<N/2$. This condition will guarantee that trellis termination for the first decoder is sufficient and there will not be any low weight sequence at the output of the second decoder caused by failure of trellis termination.

Step 2) Choose the maximum predetermined weight w_{det} for input data sequences and the minimum permissible effective free distance of the code $d_{\min,\,w_{\text{det}}}$. Find all input data sequences of length N and weight $w_l\le w_{\text{det}}$ and their corresponding effective free distance d_{w_l} for the Turbo encoder with an interleaver design based on step 1 such that $d_{w_l}\le d_{\min,\,w_{\text{det}}}$. All these input data sequences are divisible before and after interleaving by the feedback polynomial (usually a primitive polynomial) of the Turbo encoder. Consider the first input data block of weight w_1 with nonzero elements in locations $(i_1,i_2,\ldots,i_{w_1})$ and $d_{\min,\,w_1}\le d_{\min,\,w_{\text{det}}}$. Compute $\text{IDS}_{(\text{new})}$ based on (9) for the original interleaver designed in step 1. Set $j=i_1+1$ and find the pair $(j,\pi(j))$. Interchange the interleaver pairs $(i_1,\pi(i_1))$ and $(j,\pi(j))$ to create a new interleaver, i.e., $(i_1,\pi(j))$ and $(j,\pi(i_1))$. Compute the new IDS, $\text{IDS}'_{(\text{new})}$, based on the new interleaver design. If $\text{IDS}'_{(\text{new})}\le\text{IDS}_{(\text{new})}$, replace the interleaver by the new one. Otherwise, set $j=j+1$ and continue. Repeat this operation for all input data sequences with a minimum weight of $w_l\le w_{\text{det}}$ and $d_{w_l}\le d_{\min,\,w_{\text{det}}}$. After completing this operation, return to step 2 and find all input data sequences of weight $w_l\le w_{\text{det}}$ with $d_{w_l}\le d_{\min,\,w_{\text{det}}}$ for the new

interleaver. Continue this step until it converges and there is no input data sequence of weight $w_l \le w_{\rm det}$ with $d_{w_l} \le d_{\min,\, w_{\rm det}}$. Obviously, if $d_{\min,\, w_{\rm det}}$ is too large, the second step may never converge, and in this case, $d_{\min,\, w_{\rm det}}$ should be reduced.

An interleaver design proposed in [14] and [15] is based on the joint S-random criteria and elimination of all error patterns of weight w_i. However, in practice, the joint optimization criteria will not converge easily and, therefore, the value of S must be reduced and w_i restricted to only weight two inputs. For weights larger than two, the convergence of the algorithm is a problem because of the large number of possibilities. By separating these two criteria into two steps, we can easily find the appropriate interleaver satisfying each step separately. The two steps in the two-step S-random interleaver design are independent operations. The second step tries to increase the minimum effective free distance of the code (based on the interleaver design in the first step) to a predetermined value ($d_{\min,\, w_{\rm det}}$), while attempting not to increase the correlation between the information input data and the soft output of each decoder corresponding to its parity bits. Obviously, if $d_{\min,\, w_{\rm det}}$ is set to too large a value, the second stage of the design may completely change the interleaver produced by the first step and produce an inferior design. This possibility will be illustrated later by simulation.

It is shown in [13] that the feedback polynomials for the recursive systematic convolutional encoder of Turbo codes should be chosen to be primitive polynomials. When used for Turbo codes, primitive polynomials exhibit better distance spectrum properties. The Appendix describes how to find all input data sequences of weight $w_{\rm det}$ that are divisible by a primitive polynomial. This information is required for the second step in our approach.

IV. Deterministic Interleaver Design

The following theorem describes a deterministic interleaver based on step 1 in the previous section.

Theorem 1: Let α and N be relatively prime natural numbers such that $\alpha-1$ divides N, and let $S_1 = \min\{\alpha,\, \lfloor N/(\alpha+1)\rfloor\}$, $S_2 = \lfloor(\alpha-1)/2\rfloor$. Then there is a permutation $\pi \in S_N$ such that a) if $|(i-j) \bmod N| \le S_1$ and $i \ne j$ then $|\pi(i)-\pi(j)) \bmod N| \ge S_1$, and b) for all i, $|(i-\pi(i)) \bmod N| \ge S_2$.

Proof: Let $\beta = \lfloor(\alpha-1)/2\rfloor$ and define π: $\{1, \ldots, N\} \to \{1, \ldots, N\}$ by $\pi(i) = \alpha i + \beta$, where $\pi(i)$ is to be interpreted as the number $\pi(i) \in \{1, \ldots, N\}$ that is congruent to $\alpha i + \beta$ modulo N. Since $\gcd(\alpha, N) = 1$, π is indeed a permutation. If α^{-1} denotes the inverse of $\alpha \bmod N$, then $\pi^{-1}(j) = \alpha^{-1}(j-\beta)$ is the inverse permutation to π.

a) Note that $S_1 \le \alpha$ and $S_1 \le \lfloor N/(\alpha+1)\rfloor$. Let i and j be elements of $\{1, \ldots, N\}$ with $i \ne j$ and $|(i-j) \bmod N| \le S_1$. Then either i) $1 \le i-j \le S_1$ or ii) $1 \le N-(i-j) \le S_1$.

In case i) we have $|(\pi(i)-\pi(j)) \bmod N| = |\alpha(i-j) \bmod N| = \min\{\alpha(i-j),\, N-\alpha(i-j)\}$, and we will show both terms are $\ge S_1$. In fact, since $i-j \ge 1$, $\alpha(i-j) \ge \alpha \ge S_1$. Also, since $i-j \le S_1 \le N/(\alpha+1)$, we have $N-\alpha(i-j) \ge N-\alpha N/(\alpha+1) = N/(\alpha+1) \ge S_1$.

In case ii) we have $1 \le N-(i-j) \le S_1$, so $N-S_1 \le i-j \le N-1$. However

$$N - \frac{N}{\alpha} \le N - \frac{N}{\alpha+1} \le N - S_1$$

so $\alpha N - N \le \alpha(i-j) \le \alpha N - \alpha \le \alpha N$, which means $(\alpha(i-j)$ is trapped between two successive multiples of N, namely $(\alpha-1)N$ and αN. Therefore

$$\begin{aligned}&|(\pi(i)-\pi(j)) \bmod N| \\ &\quad = |\alpha(i-j) \bmod N| \\ &\quad = \min\{\alpha N - \alpha(i-j),\, \alpha(i-j) - (\alpha-1)N\}.\end{aligned}$$

Again we show both terms are $\ge S_1$. Since we are in case ii), $\alpha N - \alpha(i-j) \ge \alpha \ge S_1$. Second, $\alpha(i-j) - (\alpha-1)N \ge \alpha(N-S_1) - (\alpha-1)N = N - \alpha S_1 \ge N - \alpha N/(\alpha+1) = N/(\alpha+1) \ge S_1$.

b) Let $i \in \{1, \ldots, N\}$. Then $|(i-\pi(i)) \bmod N| = |(\alpha-1)i + \beta \bmod N|$. Since $\alpha-1$ divides N, and $\beta = \lfloor(\alpha-1)/2\rfloor$, the last expression is at least $\lfloor(\alpha-1)/2\rfloor = S_2$.

Q.E.D.

To maximize the constants S_1 and S_2, the number α should be close to $\sqrt{N}$. Then S_1 is also about $\sqrt{N}$. The following elementary consideration shows that one cannot achieve $S_1 > \sqrt{N}$: Assume that $S_1 = \sqrt{N}$. Then the $\sqrt{N}$ values $\pi(1), \ldots, \pi(\sqrt{N})$ have pairwise distance $\ge \sqrt{N}$. Therefore, the "balls" with radius $\sqrt{N}/2$ cover the $\sqrt{N}\sqrt{N} = N$ numbers $\{1, \ldots, N\}$ completely. Thus, Theorem 1 yields a solution where S_1 is already optimal.

In some applications, such as wireless systems in Rayleigh fading channels, it has been suggested that an additional interleaver be incorporated either before the first encoder or in the path of the systematic data sequence, or alternatively over the entire data sequence (both the systematic data and the parity bits) in order to improve the performance of the system [17]. The deterministic interleaver proposed here can be used for these applications without adding too much complexity to the system.

It should be noted that there are other deterministic interleaver designs such as those provided in [18] and [19] that perform better than random interleavers. It would be of interest in future research to compare our approach with existing deterministic interleaver designs including those mentioned above.

V. Simulation Results and Conclusion

This section provides simulation results for the BER performance of Turbo codes using the new interleaver design and comparisons with S-random and random interleavers. The constituent encoders are recursive systematic convolutional codes with memory $m = 3$ and with feedback and feedforward generator polynomials $(15)_{\rm oct}$ and $(17)_{\rm oct}$, respectively. The trellis termination is applied only to the first encoder.

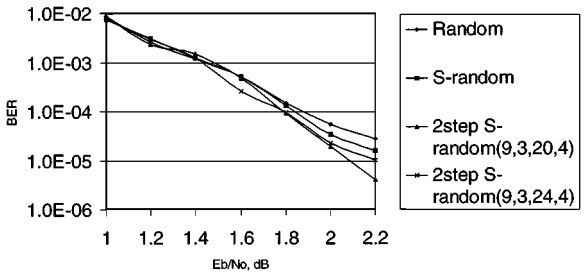

Fig. 2. Performance of Turbo code for different interleavers of size 192 bits and BPSK signal.

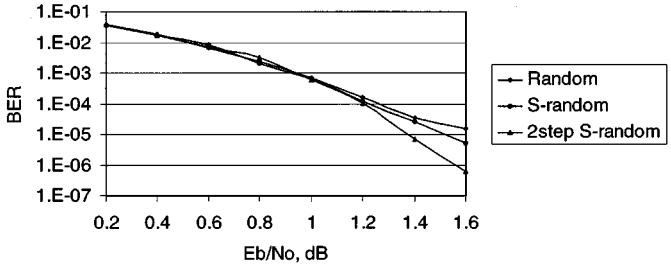

Fig. 3. Performance of Turbo code for different interleavers of size 400 bits and BPSK signal.

In all the examples, the number of iterations (using the logarithmic version of the BCJR algorithm [7]) is 18. For the first two examples, the signal is binary phase-shift keying (BPSK) with a code rate of 1/3. In the first example, the interleaver block size is 192. The BER performance of the new interleaver design is compared with S-random and random interleavers. For the new interleaver, two interleavers with design parameters $(S_1, S_2, d_{\min, w_{\det}}, w_{\det}) = (9, 3, 20, 4)$ and (9, 3, 24, 4) are chosen. For the S-random interleaver, the value of S is 9. From Fig. 2, it can be concluded that the new interleaver design performs much better than other interleavers at low BER. It is also obvious that the error floor for Turbo codes is much lower with the new interlearver design because of the larger value of $d_{\min}$. This figure also shows that choosing a very large value for $d_{\min, w_{\det}}$ can degrade the performance of the Turbo code. For this particular example, the two-step S-random interleaver with $d_{\min, w_{\det}} = 20$ performs better than that with $d_{\min, w_{\det}} = 24$. The appropriate maximum value for $d_{\min, w_{\det}}$ depends on the length of the interleaver and it is usually obtained by trial and simulations. Fig. 3 compares the BER performance of the two-step S-random interleaver design with S-random and random interleavers with a block size of 400. For the new interleaver, the design parameters are $(S_1, S_2, d_{\min, w_{\det}}, w_{\det}) = (14, 6, 26, 4)$ and for the S-random interleaver $S = 14$. The two-step S-random interleaver has much better BER performance than the S-random interleaver at low BER and results in a lower error floor for Turbo codes. In practice, because the correlation properties of the input data and the parity information are decreasing exponentially, it is sufficient to choose a small value for S_2.

We have also compared the two-step S-random interleaver with Hokfelt's interleaver design. Hokfelt's approach results in many interleavers for each run of the algorithm with different BER performance. If we choose a random instance of these designs, it may perform worse than the S-random or two-step S-random interleaver design. However, if we choose the best resulting interleaver among them, its performance can be as good as the two-step S-random interleaver design. For the interleavers of length 192 and 400 bits, the best interleavers found by Hokfelt's approach can perform as well as the two-step S-random interleavers that were used in examples 1 and 2.

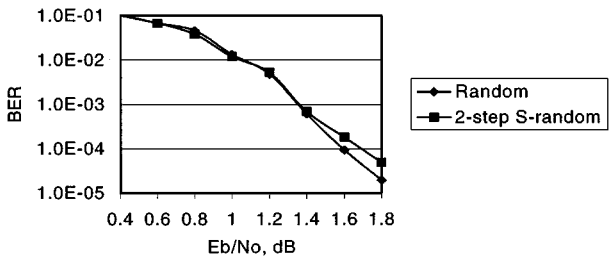

Fig. 4. Performance of Turbo code for different interleavers of size 1024 bits and QPSK signal.

For the last example, the signal is quaternary phase-shift keying (QPSK) with a code rate of 1/2. Equal number of parity bits are punctured from both encoders. The code block length is 1024. Fig. 4 compares the BER performance of a random interleaver with a deterministic interleaver described in Section IV with design parameters $(\alpha, S_1, S_2) = (33, 30, 16)$, with β the same as S_2. The performance of this deterministic interleaver is slightly worse than that of a random interleaver. However, the interleaving and deinterleaving operations can be carried out algebraically in the receiver and transmitter, thus reducing storage requirements.

Appendix I
Polynomials Divisible by a Primitive Polynomial

Let $R = GF(2)[X]$ be the ring of polynomials with binary coefficients, and let $p(X) \in R$ be a primitive irreducible polynomial of degree $m > 1$. We wish to determine all the polynomials $f(X) \in R$ which have low weight and are divisible by $p(X)$. (The weight of a polynomial is the number of nonzero terms).

Choose a zero α of $p(X)$. Then α generates $GF(2^m)$ as a field. Since $p(X)$ is primitive, by definition the minimal $n > 0$ with $\alpha^n = 1$ is $n = 2^m - 1$. Note that the nonzero elements of $GF(2^m)$ are precisely the n zeros of the polynomial $X^n - 1$.

Since $p(X)$ is irreducible, a polynomial $f(X) \in R$ is divisible by $p(X)$ if and only if $f(\alpha) = 0$. If $i, j \in \mathbf{N}$ satisfy $i \equiv j \pmod n$, then $\alpha^j = \alpha^i$, hence $X^i + X^j$ is divisible by $p(X)$. Let T_2 be the set of polynomials $X^i + X^j \in R$ with $0 \leq i < j$, $i \equiv j \pmod n$. More generally, let T_{2k} $(k = 2, 3, \ldots)$ be the sum of k disjoint (i.e., all monomials are distinct) terms from T_2.

Let H be the Hamming single-error-correcting code with generator polynomial $p(X)$, and let A_w be the set of codewords of H of weight w, written in the usual way as polynomials of degree $< n$ corresponding to residue classes in $R/(X^n - 1)$. Note that A_i is empty unless $i \equiv 3$ or $0 \pmod 4$, i.e., A_1, A_2, A_5, A_6, $\ldots$ are empty.

Theorem 2: Let $f(X) \in R$ have weight w and write

$$f(X) = g(X) + h(X)$$

where $g(X) \in T_{2i}$, $h(X) \in R$ has weight j, no two exponents of $h(X)$ are congruent modulo n, and the terms of $g(X)$ and $h(X)$ are disjoint (i.e., $w = 2i + j$). Then $f(X)$ is divisible by $p(X)$ if and only if $\phi(h(X)) \in A_j$ where ϕ means "read exponents $\bmod n$."

Proof:

"$\Leftarrow$" Let $f(X) = g(X) + h(X)$ be as in the theorem. Since $\phi(h(X))$ is divisible by $p(X)$, one has $\phi(h(\alpha)) = h(\alpha) = 0$. Therefore $g(X) \in T_{2i}$ and $h(X)$ are both divisible by $p(X)$ and so is $f(X)$. By construction the weight of $f(X)$ is $w = 2i + j$.

"$\Rightarrow$" Let $f(X) \in R$ be divisible by $p(X)$. By construction $g(X)$ and hence $h(X)$ is divisible by $p(X)$, where $\phi(h(X)) \in A_j$ for some j. Again by construction the weight of $h(X)$ is the weight of $\phi(h(X))$ and the weight of $f(X)$ is $2i + j = w$.

Q.E.D.

Note that the polynomials $g(X)$ and $h(X)$ are not necessarily unique. But one may define $g(X)$ by starting from the highest exponent of $f(X)$ and always taking the first term that fits to make the decomposition unique.

We discuss the first few values of w individually, and illustrate by taking $m = 3, n = 7$ and $p(X) = X^3 + X + 1$. Then H is a Hamming code of length seven, containing seven words of weight three, seven of weight four, and one word of weight seven.

Weight $w = 1$: No monomials are divisible by $p(X)$.

Weight $w = 2$: A weight two polynomial is divisible by $p(X)$ if and only if it is in T_2.

Examples: $1 + X^7$, $X^4 + X^{39}$.

General form: $f(X) = X^i + X^{i+7j}$, $i \geq 0$, $j \geq 1$.

Weight $w = 3$: A weight three polynomial is divisible by $p(X)$ if and only if it reduces to a weight three codeword in H when the exponents are read $\bmod n$.

Example: The seven words in A_3 are the cyclic shifts of $p(X)$ itself. So, for instance, $X^{32} + X^{16} + X^8$ is divisible by $p(X)$, since it reduces to $X^4 + X^2 + X = Xp(X) \in A_3$.

General form: $f(X) = X^{i+7j} + X^{i+1+7k} + X^{i+3+7l}$, $i, j, k, l \in \mathbf{Z}$, $i + 7j, i + 1 + 7k, i + 3 + 7l \geq 0$.

Weight $w = 4$: A polynomial of weight 4 is divisible by $p(X)$, if and only if it is in T_4, or it reduces to an element of A_4 when the exponents are read $\bmod n$.

Examples: $1 + X^7 + X^{10} + X^{17} \in T_4$, $1 + X^2 + X^3 + X^4 \in A_4$.

ACKNOWLEDGMENT

The authors would like to thank D. Rowitch, Editor, and the anonymous reviewers for their helpful comments to improve the paper. They would also like to thank J. Hokfelt for his comments on IDS criterion.

REFERENCES

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE Int. Conf. Communications*, vol. 2, Geneva, Switzerland, 1993, pp. 1064–1070.

[2] J. Hokfelt, O. Edfors, and T. Maseng, "Turbo codes: Correlated extrinsic information and its impact on iterative decoding performance," in *Proc. IEEE 49th Vehicular Technology Conf.*, vol. 3,, Houston, TX, May 1999, pp. 1871–1875.

[3] A. K. Khandani, "Group structure of turbo codes with applications to the interleaver design," in *Int. Symp. Information Theory*. Boston, MA: MIT, Aug. 1998, p. 421.

[4] O. Y. Takeshita and D. J. Costello Jr., "New classes of algebraic interleavers for turbo codes," in *Int. Symp. Information Theory*. Boston, MA: MIT, Aug. 1998, p. 419.

[5] H. Herzberg, "Multilevel turbo coding with short interleavers," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 303–309, Feb. 1998.

[6] L. C. Perez, J. Seghers, and D. J. Costello, "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1698–1709, Nov. 1996.

[7] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimum decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[8] W. Blackert, E. Hall, and S. Wilson, "Turbo code termination and interleaver conditions," *Electron. Lett.*, vol. 31, no. 24, pp. 2082–2084, Nov. 23, 1995.

[9] A. S. Barbulescu and S. S. Pietrobon, "Terminating the trellis of turbo codes in the same state," *Electron. Lett.*, vol. 31, pp. 22–23, Jan. 1995.

[10] M. C. Reed and S. S. Pietrobon, "Turbo code termination schemes and a novel alternative for short frames," in *7th IEEE Int. Symp. Personal, Indoor, Mobile Communications*, vol. 2, Taipai, Taiwan, Oct. 15–18, 1996, pp. 354–358.

[11] S. Dolinar and D. Divsalar, "Weight distribution for turbo codes using random and nonrandom permutations,", JPL Progress report 42-122, Aug. 15, 1995.

[12] J. Hokfelt, O. Edfors, and T. Maseng, "Interleaver design for turbo codes based on the performance of iterative decoding," in *Proc. IEEE ICC*, vol. 1, Vancouver, BC, Canada, June 1999, pp. 93–97.

[13] S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Trans. Commun.*, vol. 44, pp. 591–600, May 1996.

[14] A. K. Khandani, "Optimization of the interleaver structure for turbo codes," in *Proc. Canadian Workshop Information Theory*, June 1999, pp. 25–28.

[15] J. Yuan, B. Vucetic, and W. Feng, "Combined turbo codes and interleaver design," *IEEE Trans. Commun.*, vol. 47, pp. 484–487, Apr. 1999.

[16] D. Divsalar and R. J. McEliece, "Effective free distance of turbo codes," *Electron. Lett.*, vol. 32, no. 5, pp. 445–445, Feb. 29, 1996.

[17] E. K. Hall and S. G. Wilson, "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, Feb. 1998.

[18] Third Generation Partnership Project (3GPP), "Universal mobile telecommunications system (UMTS); Multiplexing and channel coding (FDD),", ETSI TS 125 212 V3.3.0, June 2000.

[19] Third Generation Partnership Project 2 (3GPP2), "Physical layer standard for CDMA2000 spread spectrum systems,", Release A, 3GPP2 C.S0002-A, June 2000.

Hamid R. Sadjadpour (S'90–M'95–SM'00) received the B.S. and M.S. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran in 1986 and 1988, respectively, and the Ph.D. degree in electrical engineering from University of Southern California (USC), Los Angeles, CA, in 1996.

During 1994–1995, he was also with LinCom Corporations, Los Angeles as a Member of Technical Staff. Since 1995, he has been with AT&T Research Laboratory, Florham Park, NJ, currently as a Principle Technical Staff Member. During 1999, he was also an Adjunct Professor at Lehigh University, Bethlehem, PA. His research interests include equalization techniques for wireless systems and DSL modems, error control coding and Turbo codes, communication theory, and signal processing. He holds three patents and eight patents pending.

Neil J. A. Sloane (S'62–M'66–SM'77–F'78) received the Ph.D. degree from Cornell University, Ithaca, NY, in 1967.

He was an Assistant Professor with Cornell University between 1967–1969. He joined AT&T Bell Laboratories, Murray Hill, NJ in 1969 as a Member of Technical Staff. He is currently a Technology Leader with AT&T. He is the author or coauthor of seven books: *A Handbook of Integer Sequences* (Academic Press, 1973), *A Short Course on Error-Correcting Codes* (Springer-Verlag, 1975), *Hadamard Transfrom Optics* (Academic Press, 1979), *Sphere-Packings, Lattices and Groups* (Springer-Verlag, 1998), *Claude Elwood Shannon: Collected Papers* (IEEE Press, 1993), *The Encyclopedia of Integer Sequences* (Academic Press, 1995), *Orthogonal Arrays* (Springer-Verlag, 1999), and *Rock Climbing Guide to New Jersey* (Globe Peqout Press, 2000).

Dr. Sloane is an AT&T Fellow and a Member of National Academy of Engineering, American Math. Society, Mathematic Association Amer., and American Stat. Association. He is also a recipient of the Chauvenet Price Award from the Mathematic Association of America in 1979. He was Editor-in-Chief of IEEE TRANSACTIONS ON INFORMATION THEOREY from 1978 to 1980. He also received 1987 and 1995 Information Theory Society Prize Paper Award, 1997 Shannon Lecturer of IEEE Information Theory Society, and 1984 Earle Raymond Hedrick Lecturer of the Mathematic Association America.

Masoud Salehi received the B.S. degree from Tehran University, Iran, and M.S. and Ph.D. degrees from Stanford University, Stanford, CA, all in electrical engineering.

Before joining Northeastern University, he was with the Electrical Engineering Department, Isfahan University of Technology and Tehran University. From February 1988 to May 1989, he was a Visiting Professor with the Information and Communication Theory Research Group, Eindhoven, The Netherlands, where he did research in network information theory and coding for storage media. In 1989, he joined the Department of Electrical Engineering and Computer Engineering, Northeastern University, Boston, MA, where he is currently an Associate Professor involved in teaching and research. He is the coauthor of two textbooks *Communication Systems Engineering* (Prentice-Hall, 1994) and *Communication Systems with MATLAB* (PWS-Kent, 2000). His main areas of research interest are coding, data compression, and information theory.

Gabriele Nebe received the Dipl. degree and the Ph.D. degree in mathematics from the RWTH, Aachen, Germany, in 1990 and 1995, respectively.

She was a Teaching Assistant during 1990 to 2000 with RWTH, Aachen. Since October 2000, she has been a Professor in Mathematics with the University of Ulm, Ulm, Germany.