ITU-Telecommunication Standardization Sector Temporary Document RN-029

STUDY GROUP 15 Original: English

Red Bank, NJ 21-25 May 2001

Question: 4/15

SOURCE[1]: *AT&T*.

TITLE: G.gen: G.dmt.bis: G.lite.bis: : Interleaver Design for Multi-Tone Turbo Trellis Coded Modulation Scheme for G.dmt.bis and G.lite.bis

---

Abstract

This contribution proposes to use Multi-Tone Turbo Trellis Coded Modulation (MTTCM) technique as an option for the forward error correction in ADSL modems. The contribution suggests a procedure to design an interleaver that is suitable for Turbo codes.

---

* Contact: Hamid R. Sadjadpour, Tel:(973)236-6807, Email:sadjadpour@research.att.com;

I. Interleaver design

Our proposed Turbo code for ADSL modem requires two interleavers. It is well-known that for a gives block length, the design of the interleaver can have a significant effect on the decoder performance, due to the different values of the minimum effective free distance that each interleaver can have. Therefore, in this section, we describe two new interleaver designs. One is a semi-random interleaver which is very powerful since it produces large minimum effective free distance d_{min} for Turbo code. The second one is a deterministic interleaver suitable for the second interleaver utilized in Turbo encoder structure.

A. Problem Statements

An interleaver π is a permutation $i \mapsto \pi(i)$ that maps a data sequence of N input symbols $d_1, d_2, \ldots, d_N$ into the same sequence in a new order. 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 1 in each row and column, all other entries being zero. Every interleaver has a corresponding de-interleaver (π^{-1}) that acts on the interleaved data sequence and restores it to its original order. The de-interleaving 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 the interleaver block size decreases, the performance of the 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 [1-4].

An S-random interleaver (where $S = 1, 2, 3, \ldots$) is a "semi-random" 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 $S \leq \sqrt{\frac{N}{2}}$, then this process converges [5] in a reasonable time. This interleaver design assures that the 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 [6]. Turbo codes utilize an iterative decoding process based on the MAP or other algorithms that can provide 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. Figure 8 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 Figure 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 to 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}.

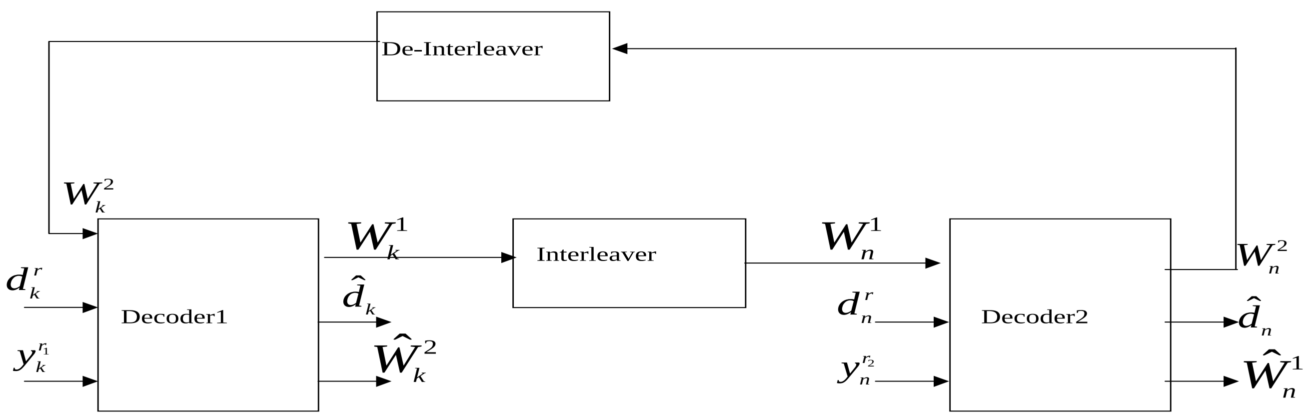

Fig. 1. Turbo Decoder.

It has been shown [1] 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, $\hat{r}^2_{\mathbf{W}^2,\mathbf{d}}$, is approximated by

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

where the two terms in the right hand side of (2) correspond to the correlation coefficients between $\mathbf{W}^2$ and the input data, i.e., $\mathbf{W}^1$ and $\mathbf{d}$ [1]. 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 de-interleaver. 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

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

where

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

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

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

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 [6].

B. 2-step S-random Interleaver Design

A new interleaver design, a 2-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. [1,6] introduced the IDS criterion to evaluate this correlation properties. The two vectors for the computation of IDS in (5) are very similar and for most interleavers 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

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

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

$$IDS_1 = \frac{1}{2N} \sum_{k_1=1}^{N} (V_{k_1} + V_{k_1}^{\prime(new)}) \tag{7}$$

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:

$$IDS_2 = \frac{1}{2N^2} \sum_{k_1=1}^{N} \sum_{k_2=1}^{N} ((\hat{r}^2_{W^2_{k_1},d_{k_2}})^2 + (\hat{r'}^2_{W^2_{k_1},d_{k_2}})^2) \tag{8}$$

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

$$IDS_{(new)} = \frac{1}{2}(IDS_1 + IDS_2) \tag{9}$$

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 $|i-j| \leq S$, then $|\pi(i)-\pi(j)| > S$. However, this does not exclude the possibility that $\pi(j) = j$, which can degrade the performance of the iterative decoding of the 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_k^i. 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 [6], 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 $IDS_{(new)}$ 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 $i - j \leq S_1$ 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 pre-determined weight w_{det} for input data sequences and the minimum permissible effective free distance code $d_{min,w_{det}}$. Find all input data sequences of length N and weight $w_l \leq w_{det}$ and their corresponding effective free distance d_{w_l} for the Turbo encoder with an interleaver design based on step 1 such that $d_{w_l} \leq d_{min,w_{det}}$. 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 non-zero elements in locations $(i_1, i_2, \ldots, i_{w_1})$ and $d_{min,w_1} \leq d_{min,w_{det}}$. Compute $IDS_{(new)}$ 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, $IDS'_{(new)}$, based on the new interleaver design. If $IDS'_{(new)} \leq IDS_{(new)}$, 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 $w_l \leq w_{det}$ and $d_{w_l} \leq d_{min,w_{det}}$. After completing this operation, return to step 2 and find all input data sequences of weight $w_l \leq w_{det}$ with $d_{w_l} \leq d_{min,w_{det}}$ for the new interleaver. Continue this step until it converges and there is no input data sequence of weight $w_l \leq w_{det}$ with $d_{w_l} \leq d_{min,w_{det}}$. Obviously if $d_{min,w_{det}}$ is too large, the second step may never converge, and in this case $d_{min,w_{det}}$ should be reduced.

An interleaver design proposed in [7] and [8] 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 2-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 pre-determined value ($d_{min,w_{det}}$), while attempting at least not to increase the correlation between the information input data and the soft output of each decoder corresponding to its parity bits. Obviously, if $d_{min,w_{det}}$ 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.

C. Deterministic Interleaver Design

The following theorem describes a deterministic interleaver based on the 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 $S_1 = \min\left\{\alpha, \left\lfloor \frac{N}{\alpha+1} \right\rfloor\right\}$, $S_2 = \left\lfloor \frac{\alpha-1}{2} \right\rfloor$. Then there is a permutation $\pi \in S_N$ such that (a) if $|(i-j) \mod N| \leq S_1$ and $i \neq j$ then $|(\pi(i) - \pi(j)) \mod N| \geq S_1$, and (b) for all i, $|(i - \pi(i)) \mod N| \geq S_2$.

Proof: Let $\beta = \lfloor(\alpha - 1)/2\rfloor$ and define $\pi : \{1, \ldots, N\} \to \{1, \ldots, N\}$ 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 $\alpha \mod N$, then $\pi^{-1}(j) = \alpha^{-1}(j - \beta)$ is the inverse permutation to π.

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

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

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

$$N - \frac{N}{\alpha} \leq N - \frac{N}{\alpha + 1} \leq N - S_1 \ ,$$

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

$$|(\pi(i) - \pi(j)) \mod N| = |\alpha(i - j) \mod N| = \min\{\alpha N - \alpha(i - j), \alpha(i - j) - (\alpha - 1)N\}.$$

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

(b) Let $i \in \{1, \ldots, N\}$. Then $|(i - \pi(i)) \mod N| = |(\alpha - 1)i + \beta \mod N|$. 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 $\geq \sqrt{N}$. Therefore the 'balls' with radius $\sqrt{N}/2$ cover the $\sqrt{N}\sqrt{N} = N$ numbers $\{1, \ldots, N\}$ completely. So Theorem 1 yields a solution where S_1 is already optimal.

D. Simulation Results

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

For all the examples, the number of iterations (using the logarithmic version of BCJR algorithm [9]) is 18. For the first two examples, the signal is BPSK with a code rate of $\frac{1}{3}$. In the first example, the interleaver block size is 192. The BER performance of the new interleaver design is compared to 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 Figure 2 it can be concluded that the new interleaver design performs better than other interleavers at low BER. It is also obvious that the error floor for Turbo codes is much lower with the new inteleaver 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 2-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. Figure 3 compares the BER performance of the 2-step S-random interleaver design to 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 S-random interleaver $S = 14$. The 2-step S-random interleaver has better BER performance than the S-random interleaver at low BER and it 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.

II. Summary

This contribution should be presented under the activity in G.gen.

We propose to include Multi-Tone Turbo Trellis Coded Modulation with fixed interleaver size to G.lite.bis and G.dmt.bis.

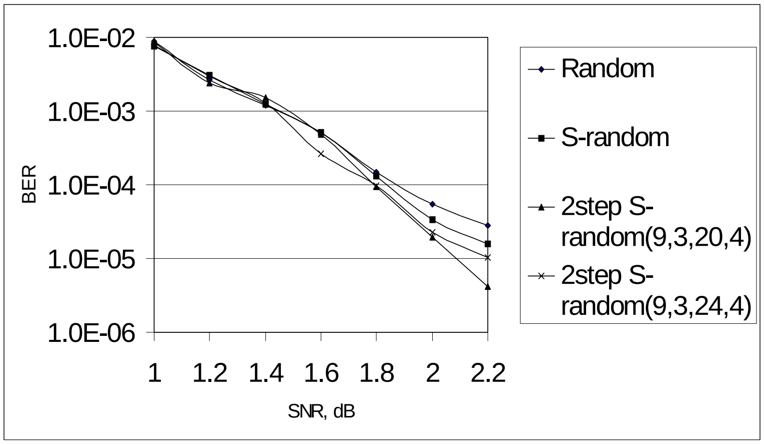

Fig. 2. Performance of Turbo code for different interleavers of size 192 bits and BPSK signal.

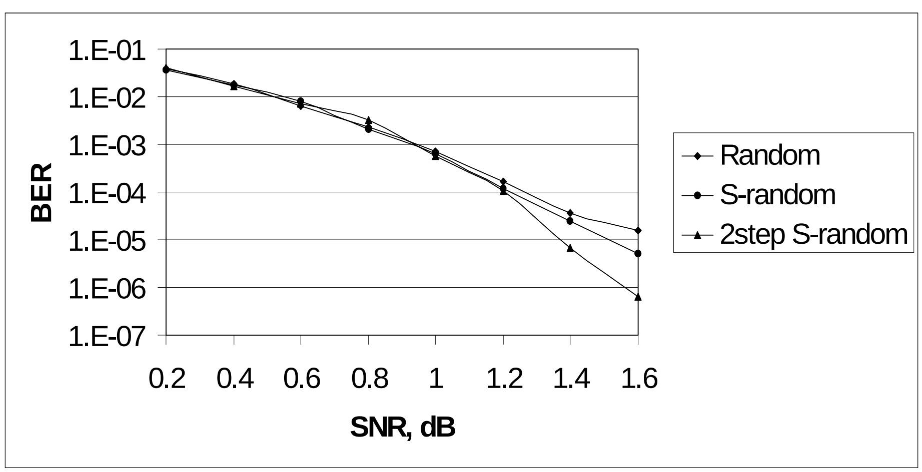

Fig. 3. Performance of Turbo code for different interleavers of size 400 bits and BPSK signal.

References

[1] J. Hokfelt, O. Edfors, and T. Maseng, *"Turbo Codes: Correlated Extrinsic Information and its Impact on Iterative Decoding Performance,"* Proceeding of IEEE VTC '99, Houston, Texas.

[2] A.K. Khandani, *"Group Structure of Turbo Codes with Applications to the Interleaver Design,"* International Symposium on Information Theory, pp. 421, August 1998.

[3] O.Y. Takeshita and D.J. Costello, Jr., *"New Classes of Algebraic Interleavers for Turbo Codes,"* International Symposium on Information Theory, pp. 419, August 1998.

[4] H. Herzberg, *"Multilevel Turbo Coding with Short Interleavers,"* IEEE Journal on selected areas in Communications, vol.16, no. 2, pp. 303-309, February 1998.

[5] S Dolinar and D. Divsalar, *"Weight Distribution for Turbo codes Using Random and Nonrandom Permutations,"* JPL Progress report 42-122, pp. 56-65, August 15, 1995.

[6] J. Hokfelt, O. Edfors, and T. Maseng, *"Interleaver Design for Turbo Codes Based on the Performance of Iterative Decoding,"* Proceeding of IEEE ICC '99, Vancouver, Canada.

[7] A.K. Khandani, *"Optimization of the interleaver structure for Turbo codes,"* Proceeding of the 1999 Canadian workshop on Information theory, pp. 25-28, June 1999.

[8] J. Yuan, B. Vucetic, and W. Feng, *"Combined Turbo Codes and Interleaver Design,"* IEEE Trans. on Comm., vol. 47, no. 4, pp. 484-487, April 1999.

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