AN OPTIMAL DEGREE DISTRIBUTION DESIGN AND A CONDITIONAL RANDOM INTEGER GENERATOR FOR THE SYSTEMATIC LUBY TRANSFORM CODED WIRELESS INTERNET

T. D. Nguyen, L. L. Yang, S. X. Ng and [1]*L. Hanzo*

School of ECS, University of Southampton, SO17 1BJ, UK.
Tel: +44-23-8059 3125, Fax: +44-23-8059 4508
Email: [1]lh@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract – **Recently, the authors of this paper proposed Systematic Luby Transform (SLT) codes and their soft decoding using the classic loglikelihood message passing algorithm for transmission over hostile Wireless Internet channels, where the transmitted data is affected by both packet loss events and random Gaussian noise. This scheme is further improved here with the aid of a new degree distribution and a novel random integer generator, which are termed as *the truncated degree distribution and the conditional random integer generator*. The SLT code using the new design is capable of achieving $BER \leq 10^{-5}$ at low E_b/N_0 values. For example, the SLT(1200,3600) code attains $BER \leq 10^{-5}$ in excess of an E_b/N_0 value of $1.5dB$ for transmission over the AWGN channel and above $3.5dB$ over the uncorrelated Rayleigh channel if additionally a packet erasure probability P_e of 0.1 is inflicted, an E_b/N_0 value above $2dB$ is required for transmission over the AWGN channel and in excess of $4dB$ over the uncorrelated Rayleigh channel, when using a maximum of $Iter = 20$ iterations and Quadrature Phase-Shift Keying.**

1. INTRODUCTION

When designing Luby Transform (LT) codes, there are three important factors, which determine the attainable performance of the LT code, namely the degree distribution, the integer random generator used for coining a particular packet's degree and the total number of source packets to be transmitted [1], [2], [3], [4]. However, as mentioned in [2], [3] and [4] the degree distribution design is the most influential factor. The original distribution proposed in [2], [3] and [4] for Systematic Luby Transform (SLT) codes is no longer optimal, when using soft bit decoding. Hence, the novel contribution of this paper is that of designing an optimal degree distribution for SLT codes. This distribution is referred to as the Truncated Degree Distribution (TDD) and will be outlined in Section 3. Furthermore, we also introduce a novel random integer generator for determining the specific degree of each packet during the LT encoding process. This random generator is termed as the conditional random integer generator, which will be detailed in Section 4. In Section 2 the soft bit based SLT is discussed, while in Section 5 its performance is analysed by using EXIT charts. Finally, in Section 6 the achievable BER performance is detailed and our conclusions are offered in Section 7.

THE FINANCIAL SUPPORT OF THE EPSRC, UK AS WELL AS THAT OF THE EUROPEAN UNION IS GRATEFULLY ACKNOWLEDGED.

2. SYSTEMATIC LUBY TRANSFORM ENCODING

The SLT encoder's operation may be followed with the aid of Fig 1, where the original m-bit information packets are encoded by the SLT encoder, which generates n number of m-bit packets from k number of m-bit packets. Owing to its systematic encoding regime, first the original k number of m-bit packets are directly copied into the encoder's output buffer and then $(n-k)$ number of m-bit parity packets are created according to the modulo-2 encoding rule of conventional LT codes. For the sake of a more compact graphical representation the n number of m-bit SLT encoded packets may be portrayed as a vertically stacked set of n number of m-bit packets, as seen in Fig 1. According to this representation the top k number of bits stacked vertically correspond to the bits belonging to the same bit positions in the k number of consecutive original m-bit packets, while the remaining $(n-k)$ bits at the bottom of each column represent the corresponding bits of the $(n-k)$ SLT-encoded parity packets. Hence, if we refer to the generator matrix and the parity check matrix of the SLT code as G and H, then we have the syndrome equation of $G \cdot H^T = 0$. The graphic representation of the SLT code seen in Figure 1 may be interpreted as a set of m codewords of an (n,k) block code, such as an LDPC code for example, where each of the vertical (n,k) codes is represented by the generator and parity check matrices of $(G_0; H_0), (G_1; H_1), \cdots, (G_m; H_m)$, which obey the structure of Figure 1.

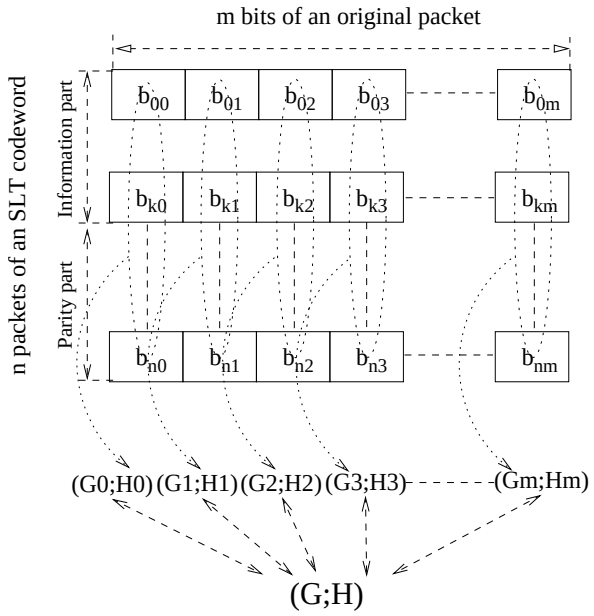

Fig. 1. The m number of (n, k) codewords of SLT codes and their generator as well as parity check matrices.

Proposition:

If G and H are the generator matrix as well as the parity check matrix of a SLT code and they satisfy the syndrome equation $G \cdot H^T = 0$, then all generator and parity check sub-matrices G_i and H_i also satisfy the equation $G_i \cdot H_i^T = 0$.

Proof:

As seen in Figure 1, the bits of the SLT codewords are generated by the resultant generator matrices G_i and all of these matrices satisfy $G_i = G, i = 1, \cdots, n$ and $H_i = H$. Since we have $G \cdot H^T = 0$, consequently $G_i \cdot H_i^T = 0$. From the above proposition we infer the statement that a codeword of the SLT code has a length of n bits. Hence, we infer the following Lemma as:

Lemma:

The decoding process of a single codeword of the SLT code may be decomposed into m decoding processes of the (n, k) constituent codewords.

3. THE TRUNCATED DEGREE DISTRIBUTION FOR SYSTEMATIC LUBY TRANSFORM CODES

For the sake of finding the optimal degree distribution, we commence our discourse by outlining the original Robust Soliton Degree Distribution (RSDD) $\mu(d)$ used in [2] and [5]. The RSDD $\mu(d)$ is composed of two parts, namely $\rho(d)$ and $\tau(d)$ formulated as follows [2]:

$$\rho(d) = \begin{cases} 1/K & \text{for } d = 1, \\ \frac{1}{d(d-1)} & \text{for } d = 2, 3, ..., K, \end{cases} \quad (1)$$

and

$$\tau(d) = \begin{cases} \frac{S}{K}\frac{1}{d} & \text{for d=1, 2, . . ., } \frac{K}{S} - 1, \\ \frac{S}{K}\log(\frac{S}{\delta}) & \text{for } d = \frac{K}{S}, \\ 0 & \text{for } d > \frac{K}{S}, \end{cases} \quad (2)$$

which are combined as follows:

$$\mu(d) = \frac{\rho(d) + \tau(d)}{Z'}, \quad (3)$$

where K, S, d are the number of input packets, the number of packets having a degree-one and the degree of packets, respectively. Finally, we have $Z' = \sum_d [\rho(d) + \tau(d)]$. The number of degree-one packets generated by this distribution is [2] $S \equiv c \cdot log_e\ (K/\delta)\ \sqrt{K}$, where c and δ constitute the tuneable parameters of the distribution. There are two conditions, which have to be satisfied by all input packets for them to be recovered from the received LT-encoded packets. Firstly, the number of LT-encoded packets has to satisfy [2] $N \geq K + 2 \cdot log_e(\frac{S}{\delta}) \cdot S$ and secondly, the number of packets having the highest degree must obey $d \geq \frac{K}{S}$ [2], [3]. These are the prerequisites for designing the original LT code. When we design the systematic Luby Transform code of [1], these are still sufficient conditions for decodability of SLT codes having a code rate higher than $R = \frac{1}{2+\epsilon}$, where $\epsilon = 2 \cdot log_e(\frac{S}{\delta})$ is the relative packet-level overhead of the original LT codes. However, these conditions are no longer sufficient for SLT codes having a code rate lower than $R = \frac{1}{2+\epsilon}$, because the density of both the parity and generator matrices of the SLT codes are low. For the sake of improving the density distribution of both the parity and generator matrices of the SLT codes, we re-define the degree distribution used for generating the parity check part of SLT codes as follows:

$$\Omega(d) = \begin{cases} \frac{1}{Z}\left[1 + \frac{S}{K} + \nu(d)\right] & \text{for } d = \gamma, \\ \frac{1}{Z}\left[\frac{\gamma}{d\cdot(\frac{\gamma}{d}+1)} + \frac{S}{K}\frac{\gamma}{d}\right] & \text{for } d = 2\gamma, 3\gamma, \cdots, \frac{K\cdot\gamma}{S} - 1, \\ \frac{S}{Z\cdot K}\log(\frac{S}{\delta}) & \text{for } d = \frac{\gamma\cdot K}{S}, \\ 0 & \text{for } d > \frac{\gamma\cdot K}{S} \text{ and } d = 1, \end{cases} \quad (4)$$

where K is the total number of original information source packets contributing to the SLT code constituted by m number of (n, k) codes, S is the number of packets having a specific degree γ, which satisfy the condition of [3] $S \equiv c \cdot log_e\cdot (K/\delta) \cdot \sqrt{K}$. Furthermore, $\nu(d)$ represents the extra fraction of redundant packets required for the Improved Robust Soliton Degree Distribution (IRSDD) of [5] to ensure decodability. Still referring to Equation 4, we have $Z = \sum_d [(\rho(d) + \tau(d)) + \nu(\gamma)]$, where γ is an integer number higher than unity. Maintaining the maximum packet degree of d_{max}=$\frac{\gamma\cdot K}{S}$

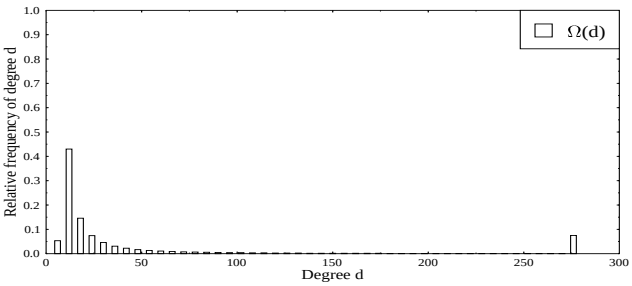

Fig. 2. An example of the truncated robust soliton degree distribution for the SLT-parity packets having parameters of $K = 10000, \delta = 0.5, c = 0.1, \gamma = 6$.

ensures that all original input packets will be represented in the set of SLT-encoded parity packets at least γ times [2], [3], [4]. Hence, the resultant packet degree distribution may be referred to as a truncated distribution having a maximum abscissa value of $d_{max} = \frac{\gamma \cdot K}{S}$ and an ordinate value step size of γ. Figure 2 shows the resultant Truncated Robust Soliton Degree Distribution (TRSDD), where we have $d_{max} = \frac{\gamma \times K}{S} = int\,(\frac{6 \times 100}{217}) = 276$.

The probability that upon generating a new SLT-encoded parity packet having a degree of γ, $K - \gamma$ SLT input information packets still remain unpresented in this SLT-encoded parity packet, encoder's output buffer was termed as the Degree Release Probability (DRP) [2], which was formulated as:

- $p(\gamma, K - \gamma) = 1$;
- $p(d, L) = \frac{d(d-1) \cdot L \cdot \Pi_{j=0}^{d-1}(K-(L+1)-j)}{\Pi_{j=0}^{d-1}(K-j)}$ for $d = 2\gamma, 3\gamma, \cdots, \frac{\gamma \cdot K}{S}$ and $L = K - d + 1, \cdots, \gamma$;
- for all other d and L values we have $p(d, L) = 0$,

where L is the number of still unrepresented input information packets, when a new SLT-encoded parity packet having a degree d is generated. Hence, given the degree distribution $\Omega(d)$ of Equation 4, the probability that there are still L unrepresented input packets, when an SLT-encoded parity packet is generated, is given by $P(d, L) = \Omega(d) \cdot p(d, L)$, where $p(d, L)$ is the DRP of an SLT-encoded parity packet having a degree of d.

Finally, the average degree of the SLT-encoded packets is calculated as follows:

$$D = \sum_{d} \frac{d \cdot (\rho(d) + \tau(d) + \nu(d))}{Z} + 1. \tag{5}$$

It can be readily shown that this average SLT-encoded packet degree gives the average row weight in the parity check matrix of the SLT codes designed in [1]. Having derived a novel degree distribution for SLT codes, in the next section we advocate a novel conditional integer random generator, which is used for determining the specific degree of each SLT-packet during the SLT encoding process.

4. CONDITIONAL RANDOM INTEGER GENERATOR DESIGNED FOR SYSTEMATIC LUBY TRANSFORM CODES

Two types of random integer generators have been proposed in [6], which are the Linear Congruential Random Integer Generator (LCRIG) and the Bit Swapping Random Integer Generator (BSRIG). Given the current integer I_j and the parameters C and M, the next integer generated by the LCRIG is given as follows [6]:

$$I_{j+1} = (aI_j + C) \bmod M, \tag{6}$$

where M is the basis of the modulo function, while a and C are positive integers referred to as the multiplier and the increment, respectively.

There are two types of the BSRIGs. The first type suggests that the bits representing the integer $(I_{n-j} + I_{n-k}) \bmod 2^b$ are rotated by r bit-positions according to:

$$I_n = [(I_{n-j} + I_{n-k}) \bmod 2^b] \text{ rot } r, \tag{7}$$

where rot denotes for the rotation function. By contrast, the second type proposes that the bits representing the integer I_{n-j} and I_{n-k} are rotated by r_1 and r_2 positions, respectively before their modulo 2 addition

$$I_n = [(I_{n-j} \text{ rot } r_1) + (I_{n-k} \text{ rot } r_2)] \bmod 2^b, \tag{8}$$

where I_n is an integer presented by b bits and the notation I_{n-j} rot r_1 means that the bits of I_{n-j} are shifted to the right by r_1 positions as exemplified by 00001111_2 (rot) $3 = 11100001_2$. Also as detailed in [6], the employment of a LCRIG has a limited benefit, because the probability of having different packet degrees is different. More explicitly, Figure 3 portrays the packet degree distribution produced by the LCRIG of [6] for the SLT code(1200,1800,$c = 0.1$,$\delta = 0.5$), where c and δ are two parameters of the SLT code of [1]. Similar to Figure 2, the horizontal axis quantifies the specific degrees of original information packets, while the vertical axis represents the relative frequency of the specific degrees having corresponding values along the horizontal axis. As we can see in Figure 3, the relative frequency of the degrees of the original information packets represented by the parity packets spans the range from degree-1 to degree-12. By contrast, the relative frequency of the degrees created by the bit swapping random integer generator of [6] characterized in Figure 4 is concentrated to a slightly narrower degree distribution range and most of the packet degrees range from $d = 2$ to $d = 6$. When designing the random integer generator used for encoding SLT codes, we aim for making this range as narrow as

possible, since this allow us to maintain a near-constant probability of the original information packets being represented by the parity packets. Consequently, the resultant probability that an original information packet is recovered from the parity packets also becomes near-constant. Hence, we may conclude from the above interpretation of Figure 3 and Figure 4 that the performance of SLT codes using the BSRIG expected to be better than that of the LCRIG of [6].

Nonetheless, the degree of SLT-encoded packets still varies

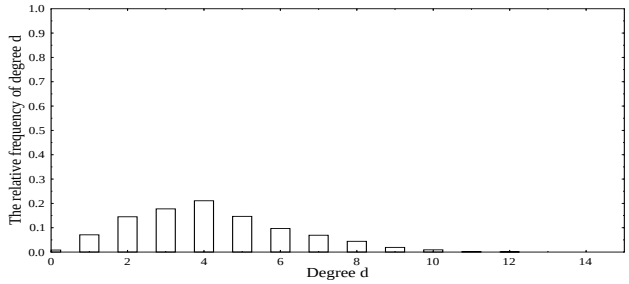

Fig. 3. The degree distribution of the congruential random integer generator of [6].

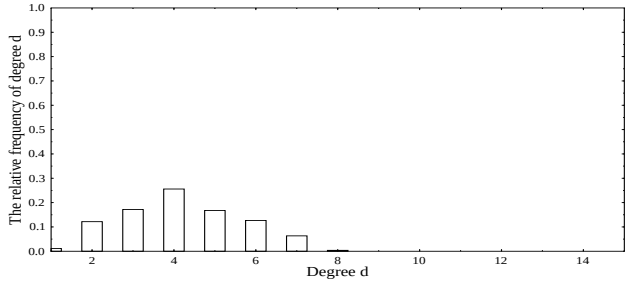

Fig. 4. The degree distribution of the swapping bit random integer generator of [6].

over quite a wide range even for the BSRIG as seen in Figure 4, which degrades the performance of the resultant SLT codes. Hence, for the sake of generating the desired Dirac-delta-like degree distribution associated with a single degree value, we specifically design a novel random integer generator, which we refer to as the conditional random integer generator, in order to improve the BSRIG of [6]. Here we invoke the bit BSRIG used of [6], while satisfying the condition of $d \leq \overline{D}$ for the output degree of the message packets, where $\overline{D}$ is the mean degree of the message packets, which is calculated as follows:

$$\overline{D} = \lceil \frac{1-R}{R} \cdot (\overline{d}_c - 1) \rceil, \tag{9}$$

and $\overline{d}_c$ is the average degree of the parity packets, which obey the Improved Robust Soliton Distribution of [5], while R is the code rate of the SLT code. When using this random integer generator, we arrive at the degree distribution seen in Figure 5. Observing that most of the output degree values are concentrated around $d = 4$. This is expected to improve the attainable performance of the resultant SLT, as we will demonstrate in Section 6.

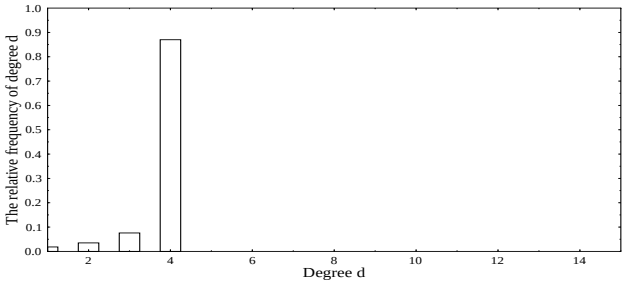

Fig. 5. The degree distribution of the conditional random integer generator of [6].

5. EXIT CHARTS OF SYSTEMATIC LUBY TRANSFORM CODES

As argued in [1] the EXIT chart of SLT codes is constituted by a message passing function between the information nodes and parity check nodes specified by the Parity Check Matrix (PCM). Let us now consider the EXIT chart of our novel SLT code, which employs the degree distributions mentioned in Section 3 for transmission over the AWGN channel. At the beginning of the decoding process, the Log Likelihood Ratios (LLRs) of the SLT-encoded bits are calculated from the output of the matched filter in the demodulator according to:

$$\log\left(\frac{Pr_{(bi=0/r)}}{Pr_{(bi=1/r)}}\right) = \log\left(\frac{\sum_{s_i \in S_0} e^{(-\frac{|r-s_i|^2}{N_0})}}{\sum_{s_i \in S_0} e^{(-\frac{|r-s_i|^2}{N_0})}}\right), \tag{10}$$

where $s_i \in S_0$ denotes a modulated symbol with the ith bit being equal to zero. Initially, the *a priori* LLRs of the parity check nodes are set to the above-mentioned matched filter output LLRs. The output LLRs of check nodes are set to zero at this stage. At the first iteration, the LLR messages $m_c(j)$ passed from the jth check nodes to the information nodes are updated as follows:

$$m_c(j) = 2 \cdot \text{atanh}\left(\prod_{i}^{d_c} m_v(i)\right), \tag{11}$$

where $m_v(i)$ represents the associated LLR message passed from the ith message node to the check nodes and d_c is the degree of the parity check node, which obeys our truncated degree distribution proposed in Section 3. The extrinsic LLRs $m_v(i)$ passed from the ith information node to the jth check node are updated according to:

$$m_v(i) = \sum_{j}^{d_v} m_c(j), \tag{12}$$

where $m_c(j)$ is the associated LLR message received from the jth check node and d_v is the degree of the information node.

From above extrinsic LLRs we generated the EXIT curves

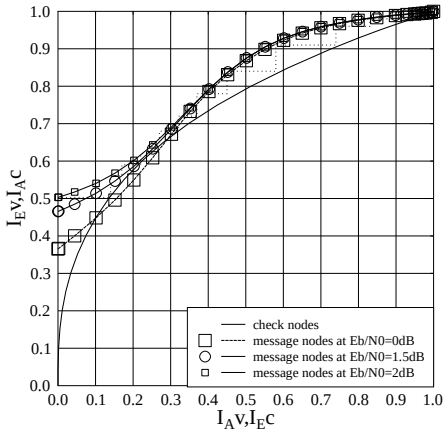

Fig. 6. EXIT chart of the SLT(1200,3600) code.

for the SLT code (1200,3600), which is seen in Figure 6. The message-node-related EXIT curve of the SLT(1200,3600) code recorded at $E_b/N_0 = 0dB$ intersects the check-node-related curve. Consequently, at $E_b/N_0 = 0dB$ the EXIT curves of the SLT(1200,3600) code cannot achieve an infinitesimally low BER. At $E_b/N_0 = 1.5dB$ the EXIT curves of the SLT code no longer intersect and after about $Iter = 20$ decoding iterations an infinitesimally low BER may be attained. By contrast, at $E_b/N_0 = 2dB$ only $Iter = 12$ decoding iterations are necessitated. Finally, our simulation-based BER results are provided in the following section.

6. SIMULATION RESULTS

An AWGN-contaminated Binary Erasure Channel (BEC) typically encountered in line-of-sight wireless Internet scenarios was assumed, where the statistical multiplexing-included random uniformly distributed packet dropping events had a probability of $P_e = 0.1$. The parameters of the system investigated are summarised in Table 1.

The BER performance of the SLT code using the truncated degree distribution of Equation 4 and the conditional random integer generator of Section 4 is shown in Figure 7 for transmission over the AWGN channel using a QPSK modulator. Figure 8 characterizes the BER performance of SLT codes for transmission over the BEC-AWGN channel having an erasure probability of $P_e = 0.1$. The BER performance of SLT codes is shown in Figure 9 for transmission over the uncorrelated Rayleigh fading channel. Finally, Figure 10 shows the

Erasure probability	P_e= 0.0 and P_e=0.1
LT code parameters in Equation 4	δ= 0.5 c= 0.1
The number of source packets	1200
SLT packet size	165 bits
SLT code rates R	1/3,1/2,2/3
SLT code rate R =1/3 and 1/2	$\gamma = 2$
SLT code rate R =2/3	$\gamma = 3$
Modulation	QPSK

Table 1. System parameters.

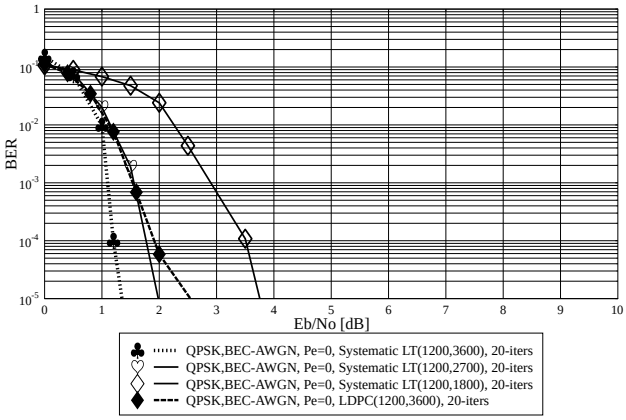

Fig. 7. BER performance of SLT codes over the AWGN channel when using a QPSK modulator and the parameters of Table 1.

attainable BER performance of SLT codes for transmission over the wireless Internet channel, where the data is contaminated by both uncorrelated, non-dispersive Rayleigh fading and erased owing to statistical multiplexing induced collissions.

7. CONCLUSIONS AND FUTURE WORKS

SLT codes using both the proposed truncated degree distribution and the conditional random integer generator achieved a high performance for transmission over various combinations of the BEC, AWGN and Rayleigh fading channels. The system using the SLT(1200,3600) code requires low E_b/N_0 values for attaining output BERs lower than 10^{-5}, when communicating over the different channels. Quantitatively, this system is capable of achieving BER $\leq 10^{-5}$ for an E_b/N_0 value around 1.5 dB over the AWGN channel and at 3.5 dB over the Rayleigh fading channel. By contrast, the corresponding benchmark system using the Low Density Parity Check code LDPC (1200,3600) requires an E_b/N_0 value upto 2.5 dB to achieve BER $\leq 10^{-5}$ over the AWGN channel.

8. REFERENCES

[1] T.D. Nguyen, L-L Yang, L. Hanzo, "Systematic Luby Transform Codes and Their Soft Decoding," *To appear in IEEE Workshop on Signal Processing Systems in Shanghai, October, 2007, at http://www.sips2007.org/AcceptedPapers.asp*.

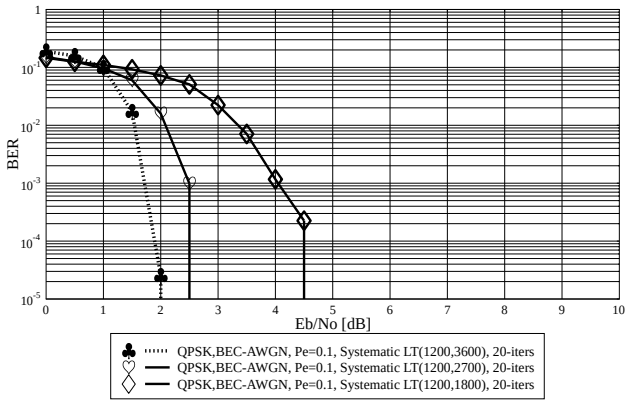

Fig. 8. BER performance of SLT codes over the BEC-AWGN channel associated with $P_e = 0.1$ when using a QPSK modulator and the parameters of Table 1.

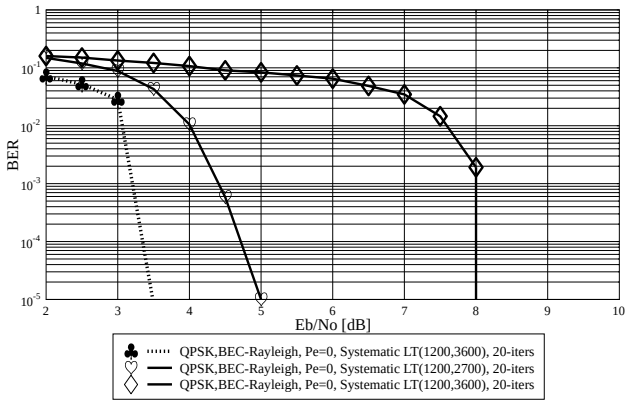

Fig. 9. BER performance of SLT codes over the uncorrelated non-dispersive Rayleigh channel when using a QPSK modulator and the parameters of Table 1.

[2] M. Luby, "LT codes," in *Proceeding of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002, pp. 271–282.

[3] D. J. C. MacKay, "Fountain Codes," *Cavendish Laboratory, University of Cambridge*, Oct 2004, at http://www.inference.phy.cam.ac.uk/mackay.

[4] A. Shokrollahi, "Raptor Codes," *Information Theory, IEEE Transactions*, pp. 2551–2567, 2006.

[5] R. Tee, T.D. Nguyen, L-L. Yang and L. Hanzo, "Serially Concatenated Luby Transform Coding and Bit-Interleaved Coded Modulation Using Iterative Decoding for the Wireless Internet," *Proceedings of VTC 2006 Spring, Melbourne,CD ROM,*, vol. 138, no. 4, pp. 177–182, May 2006.

[6] R. Tee, T.D. Nguyen, L-L. Yang and L. Hanzo, "Luby Transform Coding Aided Bit-Interleaved Coded Modulation for the Wireless Internet," *in Proceeding of VTC Fall 2007, Baltimore, September*, 2007.

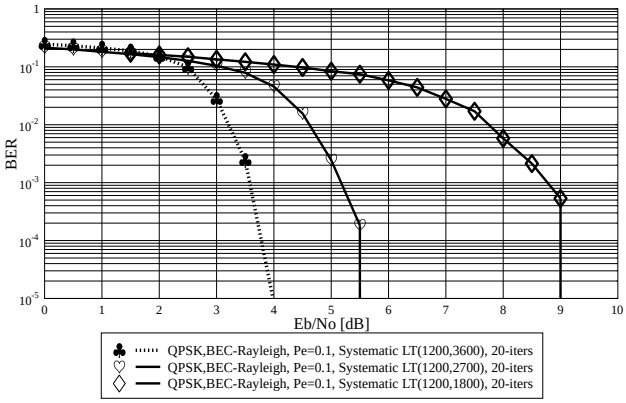

Fig. 10. BER performance of SLT codes over the Wireless Internet channel inflicting uncorrelated, non-dispersive Rayleigh fading and an erasure probability of $P_e = 0.1$ when using a QPSK modulator and the parameters of Table 1.