MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Rateless Codes on Noisy Channels

Ravi Palanki and Jonathan S. Yedidia

TR2004-037 April 2004

Abstract

We study the performance of the newly invented rateless codes (LT- and Raptor codes) on noisy channels such as the BSC and the AWGN channel. We find that Raptor codes outperform LT codes, and have good performance on a wide variety of noisy channels.

Proceedings of the 38th Annual Conference on Information Sciences and Systems

1. Published in the Proceedings of the Conference on Information Sciences and Systems, Princeton, March 2004. 2. April 2004, Included small corrections. 3. See also TR2003-124 for a longer version.

Rateless codes on noisy channels

Ravi Palanki
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, USA
email: ravi@systems.caltech.edu

Jonathan S. Yedidia
Mitsubishi Electric Research Laboratories
Cambridge, MA 02139, USA
email: yedidia@merl.com

Abstract

We study the performance of the newly invented rateless codes (LT and Raptor codes) on noisy channels such as the BSC and the AWGNC. We find that Raptor codes outperform LT codes, and have good performance on a wide variety of noisy channels.

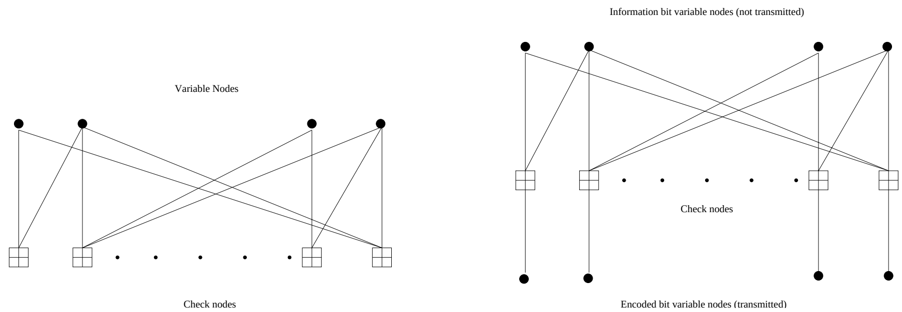
1 Introduction

In a recent landmark paper, Luby [1] designed a class of rateless codes called Luby Transform (LT) codes. These are low density generator matrix codes which are decoded using the same message passing decoding algorithm (belief propagation) that is used to decode LDPC codes. Also, just like LDPC codes, LT codes achieve capacity on every BEC. Unfortunately, LT codes also share the error floor problem endemic to capacity achieving LDPC codes. Shokrollahi [2] showed that this problem can be solved using Raptor codes, which are LT codes combined with outer LDPC codes.

The aim of this paper is to study the performance of LT and Raptor codes on channels other than the BEC. Since LDPC codes designed for the BEC perform fairly well on other channels, one might conjecture that such a result holds for LT codes as well. We test this conjecture using simulation studies.

2 LT codes

The operation of an LT encoder is very easy to describe. From k given information bits, it generates an infinite stream of encoded bits, with each such encoded bit generated as



(a) (b)

Figure 1: Tanner graph of (a) LDPC code (b) LT code.

follows:

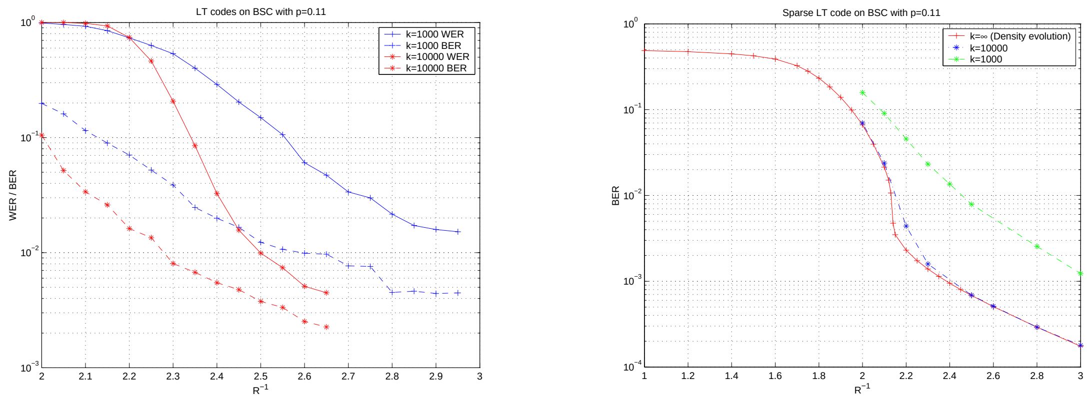
1. Pick a degree d at random according to a distribution $\mu(d)$.
2. Choose uniformly at random d distinct input bits.
3. The encoded bit's value is the XOR-sum of these d bit values.

The encoded bit is then transmitted over a noisy channel, and the decoder receives a corrupted version of this bit. Here we make the non-trivial assumption that the encoder and decoder are completely synchronized and share a common random number generator i.e., the decoder knows which d bits are used to generate any given encoded bit, but not their values. In other words, the decoder can reconstruct the LT code's Tanner graph without error. Having done that, the decoder runs a belief propagation algorithm on this Tanner graph. The message passing rules are straightforward and resemble those of an LDPC decoder. Clearly, for large block lengths, the performance of such a system depends mostly on the degree distribution μ. Luby uses the Robust Soliton (RS) distribution which is described in [1].

Luby's analysis and simulation studies show that this distribution performs very well on the erasure channel. The only disadvantage is the decoding complexity grows as $O(k \ln k)$, but it turns out that such a growth in complexity is in fact necessary to achieve capacity [2]. However, slightly sub-optimal codes called Raptor codes, can be designed with decoding complexity $O(k)$ [2].

3 LT codes on noisy channels

When the receiver tries decoding after picking up a finite number n of symbols from the infinite stream sent out by the transmitter, it is in effect trying to decode an (n, k) code, with a non-zero rate $R = k/n$. As R decreases, the decoding complexity goes up and the



(a) (b)

Figure 2: Performance on BSC with $p = 0.11$ of LT codes generated using (a) the RS(k,0.01,0.5) distribution (b) the distribution in Eqn. 1.

probability of decoding error goes down. In this paper, we have studied the variation of bit error rate (BER) and word error rate (WER) with the rate of the code on a given channel.

In Figure 2(a), we show some results for LT codes on a BSC with 11% bit flip probability. Note that the results are similar in nature on other BSCs and other AWGNCs as well. In this figure, we plot R^{-1} on the x-axis and BER/WER on the y-axis. The receiver buffers up kR^{-1} bits before it starts decoding the LT code using belief propagation. On a BSC with 11% bit flip probability, the Shannon limit is $R^{-1} = 2$ i.e., a little over $2k$ bits should suffice for reliable decoding in the large k limit. We see from the figure that an LT code with $k = 10000$ drawn using the RS(10000,0.1,0.5) distribution can achieve a WER of 10^{-2} at $R^{-1} = 2.5$ (or $n = 25000$). While this may suffice for certain applications, neither a 25% overhead nor a WER of 10^{-2} is particularly impressive. Moreover, the WER and BER curves bottom out into an error floor, and achieving very small WERs without huge overheads in nearly impossible. Going to higher block lengths is also not practical because of the $O(k \ln k)$ complexity.

The error floor problem is not confined to LT codes generated using a robust soliton distribution. Codes generated using distributions optimized by Shokrollahi for the BEC [2] also exhibit similar behaviour (Figure 2(b)). In this paper, we discuss the performance of one such distribution[1] from [2]:

$$\begin{aligned}\mu(x) &= 0.007969x + 0.493570x^2 + 0.166220x^3 \\ &\quad +0.072646x^4 + 0.082558x^5 + 0.056058x^8 + 0.037229x^9 \\ &\quad +0.055590x^{19} + 0.025023x^{65} + 0.0003135x^{66} \qquad (1)\end{aligned}$$

The main advantage of such distributions is that the average number of edges per node remains constant with increasing k, which means the decoding complexity grows only as $O(k)$ (assuming the number of iterations is fixed). On the minus side, there will be a small fraction of information bit nodes that are not connected to any check node. This means that even as k goes to infinity, the bit error rate does not go to zero and consequently, the word error rate is always one. The density evolution [3] analysis shown in Figure 2(b) supports this observation.

4 Raptor Codes

The error floors exhibited by LT codes suggest the use of an outer code. Indeed this is what Shokrollahi does in the case of the BEC [2] where he introduces[2] the idea of Raptor codes, which are LT codes combined with outer codes. Typically these outer codes are high rate LDPC codes. In this paper, we use the distribution in Eqn 1 for the inner LT code. For the outer LDPC code, we follow Shokrollahi [2] and use a left regular distribution (node degree 4 for all nodes) and right Poisson (check nodes chosen randomly with a uniform distribution).

Simulation studies, such as the one shown in Figure 3 clearly indicate the superiority of Raptor codes. Figure 3 shows a comparison between LT codes and Raptor codes on a BSC with bit flip probability 0.11. The LT code has $k = 10000$ and is generated using the distribution in Eqn. 1. The Raptor code has $k = 9500$ and uses an outer LDPC code of rate 0.95 to get $k' = 10000$ encoded bits. These bits are then encoded using an inner LT code, again generated using the distribution in Eqn 1. Figure 3 clearly shows the advantage of using the outer high-rate code.

Raptor codes not only beat LT codes comprehensively, but also have near-optimal performance on a wide variety of channels as shown in Figure 4, which shows the performance of the aforementioned Raptor code on four different channels. On each of these channels, the Raptor code has a waterfall region close to the Shannon capacity, with no noticeable error floors. Of course, this does not rule out error floors at lower WERs.

[1]Note that these distributions were not designed to be used directly in LT codes.

[2]We must mention here that Maymounkov [4] independently proposed the idea of using an outer code.

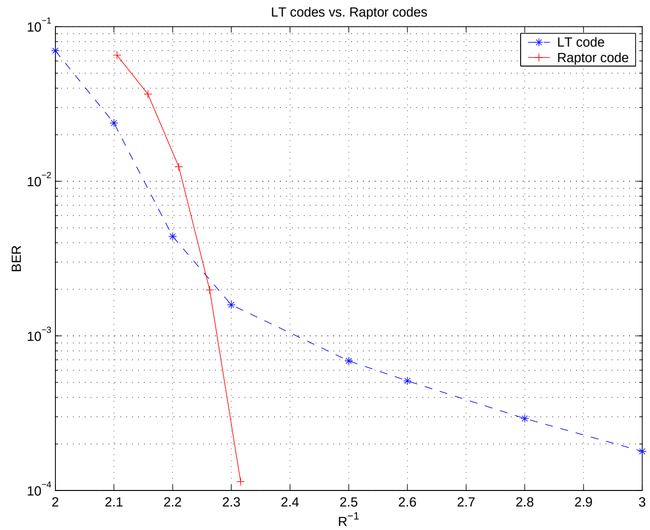


Figure 3: Comparing LT codes with Raptor codes on a BSC with p=0.11. The LT code has $k = 10000$ and is generated using the distribution in Eqn. 1. The Raptor code has $k = 9500$ and has two components: an outer rate-0.95 LDPC code and an inner rateless LT code generated using the distribution in Eqn. 1.

Figure 5 shows a histogram of the number of noisy bits needed for decoding the previously described Raptor code with $k = 9500$. We observe that the expected number of noisy bits required for successful decoding (20737) is fairly close to the Shannon limit (19000).

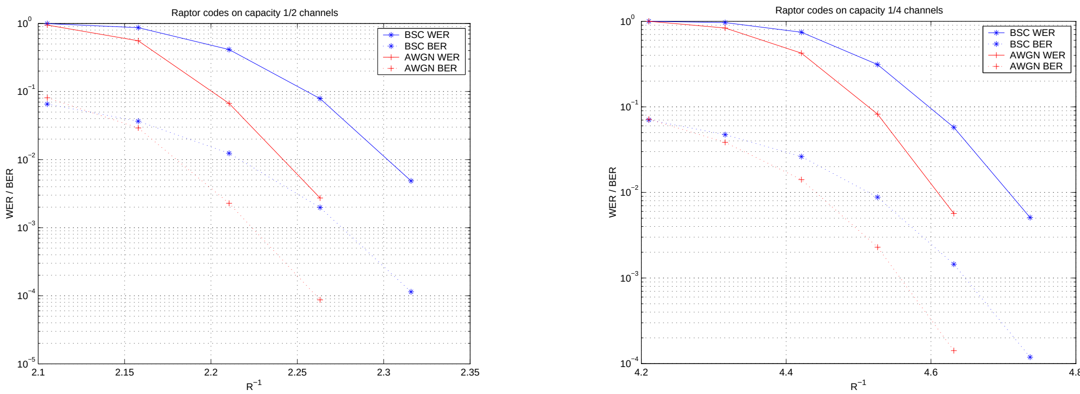
References

[1] M. Luby, "LT- codes," *Proceedings of the 43rd Annual IEEE Symposium on the Foundations of Computer Science*, pp. 271-280, 2002.

[2] A. Shokrollahi, "Raptor codes," preprint 2002. Available online at `www.inference.phy.cam.ac.uk/mackay/DFountain.html`

[3] T. Richardson and R. Urbanke, "The capacity of LDPC codes under message passing decoding," *IEEE Trans. Inform. Theory,* vol. 47, pp. 595-618, Feb 2001.

[4] P. Maymounkov, "Online codes", NYU Technical Report TR2003-883, Nov 2002. Available online at `http://www.scs.cs.nyu.edu/~petar`



(a) (b)

Figure 4: Performance of Raptor code with k=9500 and k'=10000 on different channels: (a) BSC with $p = 0.11$ and AWGNC with $E_s/N_0 = -2.83dB$. Both channels have capacity 0.5 (b) on BSC with $p = 0.2145$ and AWGNC with $E_s/N_0 = -6.81dB$. Both channels have capacity 0.25.

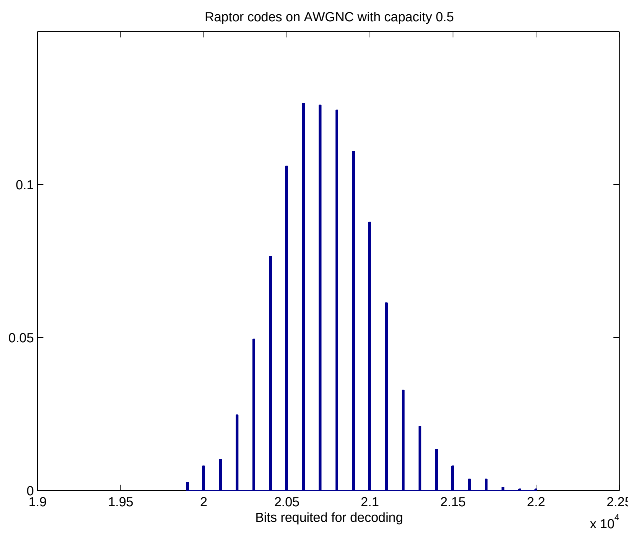


Figure 5: Histogram of number of bits required for successful decoding of Raptor code with $k = 9500$ on an AWGNC with $E_s/N_0 = -2.83dB$. The capacity of this channel is 0.5. The receiver first attempts decoding after receiving 19000 noisy bits (Shannon limit). Whenever decoding fails, the receiver waits for another 100 bits before attempting to decode again.