

Practical Layered Rateless Codes for the Gaussian Channel: Power Allocation and Implementation

Jerome M. Shapiro and Richard J. Barron
 The MITRE Corporation
 Bedford, MA 01730
 Email: jshapiro@mitre.org, rbarron@mitre.org

Gregory W. Wornell
 Massachusetts Institute of Technology
 Cambridge, MA 02139
 Email: gww@mit.edu

Abstract—Practical rateless codes for the additive white Gaussian noise channel based on the layer-dither-repeat construction of Erez, et al. [5] are developed and analyzed. We use an iteratively decoded base code for each layer and a maximal-ratio combining receiver, which provides suitable log-likelihood ratios to the decoder. For this construction we derive an algorithm for computing the optimal power allocation among the layers that takes into account the base code gap to capacity. We further derive the initial and limiting gaps to capacity of the rateless code. Finally, we verify the predicted performance characteristics via simulations using a turbo base code.

I. INTRODUCTION

Rateless codes are good codes whose codeword prefixes themselves form good codes. Such codes have recently received increased attention from the research community, in large part because they are attractive for use over channels with unpredictable channel characteristics, e.g., in many wireless environments. Using rateless codes, a transmitter only needs to transmit as much of the codeword as is necessary to decode, in effect making the code rate itself variable in response to the variability in the channel. Rateless codes allow the system designer to leave out link margin requirements because the rateless codes themselves adapt to changing channel conditions. While some of the best known rateless codes are the so-called fountain codes, designed for the application layer and erasure channel models [1], rateless codes of the type we develop here are closer in spirit to hybrid-ARQ (H-ARQ) systems familiar in many standards.

This paper describes a reduction to practice of an approach to rateless coding for the additive white Gaussian noise (AWGN) channel proposed in [5]. In [5], Erez, et al. show that when used in conjunction with capacity-achieving binary base codes of sufficiently low rate, a particular layer-dither-repeat architecture is sufficient to come within any desired fraction of capacity on the realized channel. In this paper, we derive the optimal power allocation algorithm for use of this architecture with practical codes, develop important measures of performance for the resulting rateless code, and validate the associated analysis via simulations with turbo base codes.

The rateless block code of interest [5], [6] is depicted in Figs. 1 and 2. A packet of information bits is divided into L

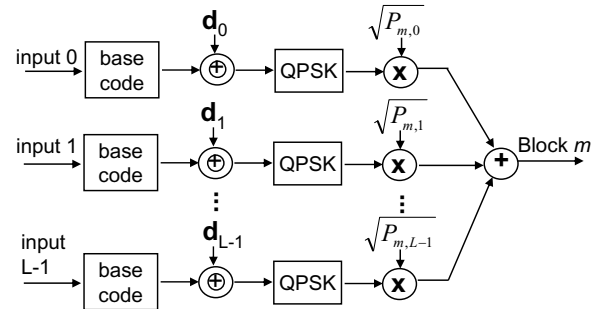


Fig. 1. Encoding with the rateless block code construction. $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{L-1}$ are binary dither sequences.

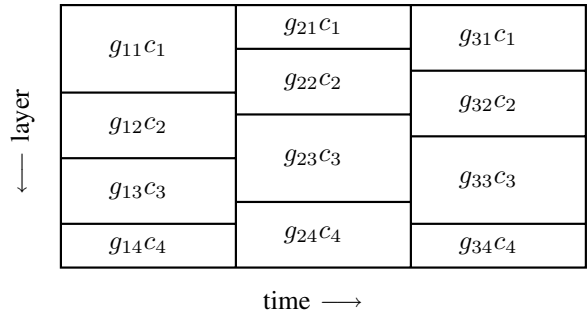


Fig. 2. A rateless code construction with 4 layers and 3 blocks of redundancy. Each block is a weighted linear combination of the base codewords c_1, c_2, \dots, c_4 , where g_{ml} denotes the weight for layer l of block m incorporating the effect of dithering and power allocation.

layers, each of which is encoded using a base code of rate R_0 (the same for all layers), dithered, modulated, and weighted by a derived amplitude. Then the layers are superimposed to form one redundancy block of the rateless code. The block is then sent over the channel. If the receiver can decode this block, it sends an acknowledgement (ACK). If not it either sends a no-acknowledgement (NACK) or there is time-out, and the transmitter sends the information again in a new redundancy block, encoding the same layers as before and using the same base codes, but with different power allocation and dithering in the superposition.

The receiver computes the channel log-likelihood ratios for soft-decision decoding by employing successive decoding with maximal ratio combining (MRC). Specifically, MRC is

applied to the first layer of the set of received blocks, treating the other layers as (white Gaussian) noise. If this first layer can be decoded, its contribution to all blocks is cancelled (removed), and decoding proceeds to the second layer. MRC and successive cancellation are now used on this second layer, and so on, until all layers are successfully decoded, at which time the receiver sends an ACK.

II. POWER ALLOCATION

In this section, we optimize the implementation of such rateless block codes when used in conjunction with practical base codes. In particular, we derive the appropriate power allocation to use when our base code of rate R_o is decodable only at a signal-to-noise ratio (SNR) ρ that is somewhat higher than what a capacity-achieving code would require, i.e.,

$$\rho = (2^{2R_o} - 1) \Delta, \quad (1)$$

for some SNR gap $\Delta > 1$ to capacity. Here we make the familiar simplifying assumption that the word-error rate curve is a perfect ‘‘cliff,’’ i.e., that the decoder always decodes a codeword perfectly for SNRs greater than or equal to ρ and fails to decode otherwise.

In our channel model, the received signal is, omitting the time index, $y = x+z$, where x is the channel input (transmitted signal) and z is the channel noise of variance N . We use N_m to denote the maximum noise threshold at which decoding is possible after receiving the m th block; the corresponding (minimum) SNR threshold is thus $\text{SNR}_m = P/N_m$.

The transmission is limited to average power P , so for an encoding of L layers, the power constraint takes the form

$$\sum_{l=0}^{L-1} P_{m,l} = P, \quad m = 1, 2, \dots \quad (2)$$

where $P_{m,l}$ is the power allocated to the l th layer of the m th block.

We now develop a simple block-by-block iteration for calculating the optimal power allocation, beginning with the first block.

A. Case: First Block

For block $m = 0$, we determine the noise threshold N_0 and power allocation $P_{0,l}$, $l = 0, 1, \dots, L-1$, such that each layer sees the same target SNR ρ that will allow decoding of the chosen base codes. Specifically,

$$\frac{P_{0,l}}{N_0 + \sum_{l'=l+1}^{L-1} P_{0,l'}} = \rho \quad (3)$$

for $l = 0, 1, \dots, L-1$.

Rewriting in matrix form

$$\begin{bmatrix} 1 & -\rho & -\rho & \cdots & -\rho \\ 0 & 1 & -\rho & \cdots & -\rho \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & -\rho \\ 1 & \cdots & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} P_{0,0} \\ P_{0,1} \\ \vdots \\ P_{0,L-1} \\ N_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ P \end{bmatrix} \quad (4)$$

and applying Gaussian elimination, we obtain¹

$$N_0 = \frac{P}{(1+\rho)^L - 1} \quad (5)$$

$$P_{0,l} = \frac{\rho(1+\rho)^{L-l-1}P}{(1+\rho)^L - 1}, \quad l = 0, 1, \dots, L-1. \quad (6)$$

Of course, the $P_{0,l}$ may be computed iteratively with l : we first calculate N_0 as in (5) and then proceed via

$$P_{0,L-1} = \rho N_0 \quad (7)$$

$$P_{0,l} = (1+\rho)P_{0,l+1}, \quad l = L-2, \dots, 1, 0. \quad (8)$$

Note that, rearranging (5), we obtain that the total rate in the transmission can be expressed in the form²

$$LR = L \frac{1}{2} \log(1+\rho) = \frac{1}{2} \log \left(1 + \frac{P}{N_0} \right), \quad (9)$$

which is the rate achievable by a capacity-achieving code over this channel when the noise level is N_0 .

Gap to Capacity: A layered rateless system using a capacity achieving base code approaches optimal efficiency in the limit of large L [5]. In this section, we show a counterbalancing effect — that using a practical base code the gap to capacity of Block 0 decoding increases monotonically with L . For this analysis, we fix the ceiling rate R for the code, which is the rate achieved if the transmission is decodable from a single received block.³ Then since

$$R = LR_o, \quad (10)$$

varying L requires varying R_o to keep R fixed. From (5) we obtain the SNR threshold for successful decoding after a single block is

$$\text{SNR}_L = (1+\rho)^L - 1, \quad (11)$$

which, substituting for ρ via (1) and using (10), yields

$$\text{SNR}_L = \left(1 + \Delta \left(2^{2R/L} - 1 \right) \right)^L - 1. \quad (12)$$

Now the corresponding SNR with a capacity-achieving code is obtained by setting $\Delta = 1$ in (12) and recognizing that the result is invariant to L , yielding $2^{2R} - 1$. Hence, the overall gap to capacity is simply a scaled version of (12):

$$\Delta_L = \frac{\left(1 + \Delta \left(2^{2R/L} - 1 \right) \right)^L - 1}{2^{2R} - 1}. \quad (13)$$

It is straightforward to verify using the convenient monotonic transformation

$$r_L = \ln(1 + \text{SNR}_L) \quad (14)$$

that SNR_L as defined in (12) converges to

$$\lim_{L \rightarrow \infty} \text{SNR}_L = 2^{\Delta 2R} - 1, \quad (15)$$

¹A similar power allocation formula is developed in [7] in the context of multi-user CDMA.

²Note that all logs are base 2 unless otherwise noted.

³In practice, a typical target R might be, for example, 4 b/s/Hz.

whence

$$\lim_{L \rightarrow \infty} \Delta_L = \frac{2^{\Delta 2R} - 1}{2^{2R} - 1}. \quad (16)$$

Moreover, Δ_L as given in (16) increases monotonically in L . To see this, it suffices to first note that 1) the second derivative of (14) is negative:

$$\frac{d^2 r_L}{dL^2} = -\frac{4(\ln 2)^2 \Delta (\Delta - 1) R^2 2^{2R/L}}{L^3 (1 + \Delta (2^{2R/L} - 1))^2} < 0, \quad (17)$$

which follows from simple calculus; and 2) r_L converges as $L \rightarrow \infty$ [cf. (15)]. Hence, r_L increases monotonically, and since r_L is a monotonic function of SNR_L , so must SNR_L .

B. Case: m th Block

We compute the optimal power allocation for the layers within block m , taking into account the use of MRC. Since we will be generating the power allocation iteratively in m , we assume the optimum power allocation for blocks $0, 1, \dots, m-1$ has been determined.

With MRC, SNRs add, so the total post-combining SNR seen by layer l after the cancelation of lower layers is the sum of SNRs for the individual redundancy blocks:

$$\text{SNR}_l(N_m) = \sum_{m'=0}^m \text{SNR}_{m',l}(N_m), \quad (18)$$

where

$$\text{SNR}_{m',l}(N_m) = \frac{P_{m',l}}{\sum_{l'=l+1}^{L-1} P_{m',l'} + N_m}. \quad (19)$$

When going from $m-1$ blocks to m blocks, the noise threshold increases from N_{m-1} to N_m , and thus the per-layer per-block SNRs all decrease. The new (m th) redundancy block must make up for this SNR shortfall to enable decoding. Since MRC is used, the SNR that must be contributed by this m th block is simply calculated as

$$\text{SNR}_{m,l}(N_m) = \rho - \sum_{m'=0}^{m-1} \text{SNR}_{m',l}(N_m), \quad (20)$$

for $l = 0, 1, \dots, L-1$, and thus we obtain L linear equations

$$P_{m,l} - \rho_{m,l} \left(\sum_{l'=l+1}^{L-1} P_{m,l'} + N_m \right) = 0. \quad (21)$$

where, for convenience, we have denoted the righthand side of (20) by $\rho_{m,l}$. The shortfall equations (21) and the power constraint (2) can together be written in matrix form

$$\begin{bmatrix} 1 & -\rho_{m,0} & -\rho_{m,0} & \cdots & -\rho_{m,0} \\ 0 & 1 & -\rho_{m,1} & \cdots & -\rho_{m,1} \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & -\rho_{m,L-1} \\ 1 & \cdots & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} P_{m,0} \\ P_{m,1} \\ \vdots \\ P_{m,L-1} \\ N_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ P \end{bmatrix} \quad (22)$$

Applying Gaussian elimination and (sub)iteration over l we can solve for N_m as

$$N_m = \frac{P}{\prod_{l'=0}^{L-1} (1 + \rho_{m,l'}) - 1}, \quad (23)$$

and then proceed from $l = L-1$ down to $l = 0$ via

$$P_{m,l} = \rho_{m,l} \left(\sum_{l'=l+1}^{L-1} P_{m,l'} + N_m \right). \quad (24)$$

Finally, note that rearranging (23) yields a total rate expression

$$\sum_{l=0}^{L-1} \frac{1}{2} \log(1 + \rho_{m,l}) = \frac{1}{2} \log \left(1 + \frac{P}{N_m} \right). \quad (25)$$

that generalizes (9). Substituting (19) and (20) into (25), we can solve for N_m from the nonlinear equation

$$\frac{1}{2} \log \left(1 + \frac{P}{N_m} \right) = \sum_{l=0}^{L-1} \frac{1}{2} \log \left[1 + \left(\rho - \sum_{m'=0}^{m-1} \frac{P_{m',l}}{\sum_{l'=l+1}^{L-1} P_{m',l'} + N_m} \right) \right]. \quad (26)$$

To verify that there is a unique solution, we find the following bounds on N_m :

$$N_{m-1} < N_m < \frac{P}{2^{2LR_o/(m+1)} - 1}. \quad (27)$$

Clearly $N_m > N_{m-1}$ because the right hand side becomes 0 for $N_m = N_{m-1}$. Otherwise, we would have achieved sufficient SNR by block $m-1$ to reach the threshold for decoding. Furthermore, the upper bound is the noise threshold for the case when we have a capacity-achieving code. Since our code is less efficient, we need a larger SNR to decode by block m .

Note that for the values of N_m in range of (27), the left hand side of (26) is a monotonically decreasing function of N_m and the right hand side is a monotonically increasing function of N_m . A search technique starting from the upper and lower bounds converges to the unique solution.

III. PERFORMANCE LOSS DUE TO IMPERFECT CODES

In this section, we examine the SNR gap to capacity as a function of the number of layers, the number of redundancy blocks, and the choice of power allocation.

In the sequel we use as an example base code a rate 1/6 CCSDS turbo code [2] using a block size of 16384 information bits, including a 32-bit CRC. The word error rate performance curve for the base code is shown in Fig. 3. We note an error floor at a word error rate of about 10^{-4} . The SNR threshold ρ (normalized by $2R_o$) is indicated by the dashed line at $E_b/N_0 = -0.3$ dB, which is 0.78 dB from capacity.

⁴ Although there have been more efficient codes reported in

⁴This value for ρ was selected by empirically optimizing over the block error performance for the rateless codes for multiple blocks.

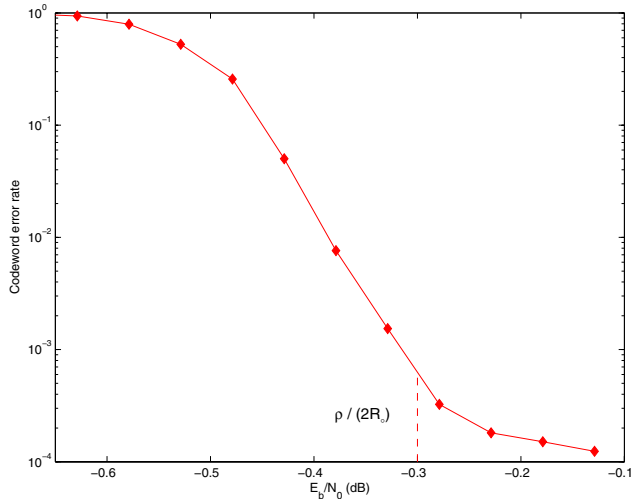


Fig. 3. Word error rate of base rate-1/6 CCSDS code. Blocklength=16384. SNR ρ used in power allocation computations shown as dashed line.

TABLE I
OPTIMAL POWER ALLOCATION FOR CAPACITY ACHIEVING CODE WITH QPSK, $R_c=1/6$, 12 LEVELS, 6 BLOCKS

Layer	Block						
	0	1	2	3	4	5	6
0	0.220	0.066	0.078	0.081	0.082	0.082	0.083
1	0.175	0.074	0.083	0.084	0.084	0.084	0.084
2	0.139	0.082	0.086	0.085	0.085	0.084	0.084
3	0.110	0.089	0.088	0.086	0.085	0.085	0.084
4	0.087	0.093	0.089	0.086	0.085	0.084	0.084
5	0.069	0.096	0.088	0.085	0.084	0.084	0.084
6	0.055	0.096	0.086	0.084	0.084	0.084	0.084
7	0.044	0.093	0.085	0.084	0.083	0.083	0.083
8	0.035	0.089	0.083	0.083	0.083	0.083	0.083
9	0.028	0.082	0.081	0.082	0.082	0.083	0.083
10	0.022	0.074	0.078	0.081	0.082	0.082	0.083
11	0.017	0.066	0.076	0.080	0.081	0.082	0.082

the literature, [3], [4], this performance is representative of the practically implementable codes currently used in practice.

A. Effect of Power Allocation on Gap to Capacity

The power allocation that results when assuming that the base code achieves capacity is given in Table I. The optimal power allocation taking into account that the code decodes with $E_b/N_0 = -0.3$ dB is given in Table II. In both cases, the allocation approaches equal power quite quickly. However, in the power allocation optimized for the particular code, initially, more power is allocated to Layer 0.

Fig. 4 shows three curves for the SNR gap to capacity as a function of the number of blocks under the following three conditions: 1) the power allocation is optimized for a decoding threshold of $E_b/N_0 = -0.3$ dB, 2) the power allocation is for the capacity achieving code, but we really have a decoding threshold of $E_b/N_0 = -0.3$ dB, and 3) equal power of $P/12$ at each layer and a decoding threshold of $E_b/N_0 = -0.3$ dB. Note that even in the absence of noise, using the power allocation of [5], Block 0 cannot decode. Similarly, the equal power case doesn't decode until Block 2 even under zero noise.

TABLE II
OPTIMAL POWER ALLOCATION FOR BASE CODE SNR THRESHOLD $E_b/N_0 = -0.3$ dB, 12 LEVELS, 6 BLOCKS.

Layer	Block						
	0	1	2	3	4	5	6
0	0.247	0.066	0.079	0.082	0.083	0.083	0.083
1	0.188	0.076	0.085	0.086	0.085	0.085	0.085
2	0.144	0.086	0.089	0.088	0.086	0.085	0.085
3	0.110	0.094	0.091	0.088	0.086	0.085	0.085
4	0.084	0.099	0.091	0.087	0.086	0.085	0.084
5	0.064	0.101	0.090	0.086	0.085	0.084	0.084
6	0.049	0.099	0.087	0.084	0.084	0.084	0.083
7	0.037	0.094	0.084	0.083	0.083	0.083	0.083
8	0.028	0.086	0.081	0.081	0.082	0.082	0.083
9	0.022	0.076	0.078	0.080	0.081	0.082	0.082
10	0.016	0.066	0.074	0.078	0.080	0.081	0.082
11	0.013	0.056	0.071	0.076	0.079	0.080	0.081

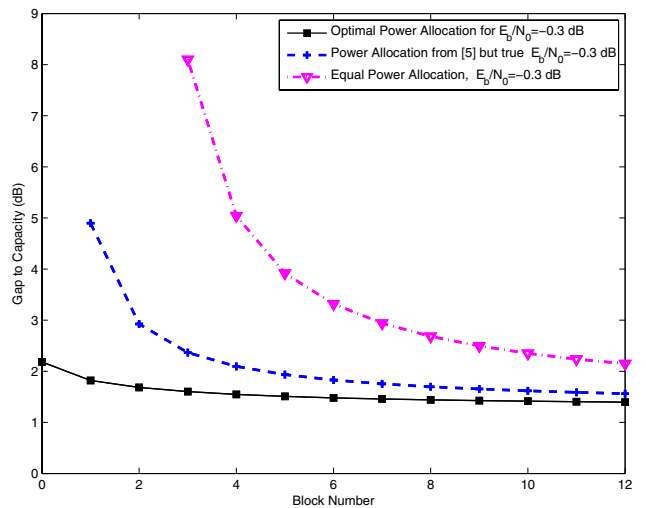


Fig. 4. SNR gap to capacity as a function of number m of redundancy blocks for different power allocations. The code is constructed from $L = 12$ layers. Note power allocation of [5] has no gap to capacity regardless of m .

With the optimal power allocation, we are 2.18 dB from capacity on Block 0, but improve to 1.4 dB from capacity by Block 12.

B. Effect of Number of Layers on Gap to Capacity

In Sec. II-A we determined that the gap to capacity of the initial block increases monotonically with the number of layers L . Here we examine the total effect of suboptimal coding and combining by considering an empirical example. Fig. 5 is a plot of the gap to capacity versus number of blocks for a layered rateless coding system with $\Delta = 0.78$ dB, corresponding to the rate 1/6 CCSDS code, and $R = 4$ b/s/Hz using a QPSK constellation. We use the iterative computation of the optimal power allocation developed in Sec. II-B. Each curve corresponds to a different number of layers L . We have run the same computations for several Δ and R , and we observe the same general trends as this representative case.

We observe that for $m = 0$ the gap increases with L , as expected. Notably for all $m > 0$, the behavior inverts and the gap decreases with L , reaching a limiting value for large

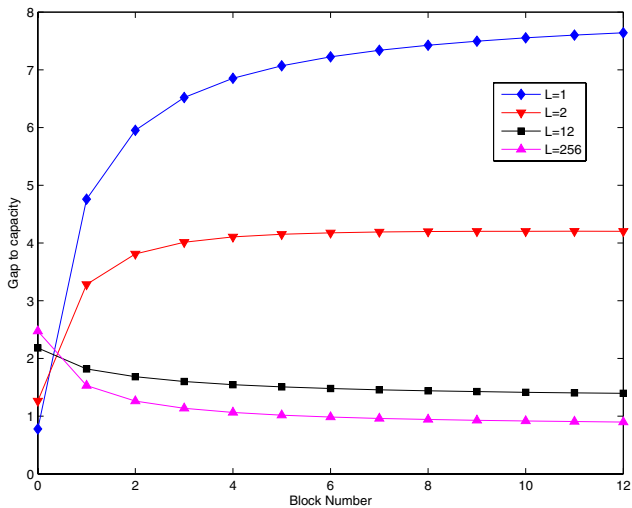


Fig. 5. Gap to capacity as a function of the number of redundancy blocks for $\Delta = 0.78$ dB and $R = 4$ b/s/Hz.

m . We also observe that for large block number, the gap to capacity reaches an asymptotic value in L , which we show below equals the base code gap to capacity Δ .

We derive the limiting value of the curves in Fig. 5 for large m . As $m \rightarrow \infty$ the channel noise will dominate the interference for all of the layers, and the effect of interference can be ignored when computing SNR. It follows that an equal power allocation for all of the blocks will provide optimal performance for large m , a fact illustrated in Fig. 4. MRC for equal power allocation will yield a combined single-layer SNR of $\rho = (m+1)P/(LN_m)$, which yields a channel SNR of $P/N_m = L\rho/(m+1)$. Taking the ratio of this channel SNR to that of a capacity-achieving code yields the gap to capacity

$$\Delta_{m,L} = \frac{L\rho}{(m+1)((1+\rho/\Delta)^{L/(m+1)} - 1)}.$$

In the limit of large m we have, using (10),

$$\lim_{m \rightarrow \infty} \Delta_{m,L} = \frac{\Delta(2^{2R/L} - 1)}{2R/L \ln 2}. \quad (28)$$

We see that the limit of (28) as $L \rightarrow \infty$ is Δ , the gap to capacity of the base code.

There is no clear objective function that one should select to determine an optimal L . A logical choice for an objective function is a weighted sum of gaps to capacity for several blocks $m = 1, 2, \dots, M$ for some $M \geq 1$. It is difficult to develop an analytical solution to such an optimization, because the determination of performance is based on a recursive formulation and solutions to transcendental equations. A numerical solution for an optimal L should be straightforward.

IV. SIMULATION RESULTS

Finally, we simulate an actual layered rateless coding system using a rate-1/6 CCSDS turbo code for a base code with QPSK modulation and $L = 12$ layers, resulting in a maximum throughput of 4 bits/sec/Hz for the Block 0 code.

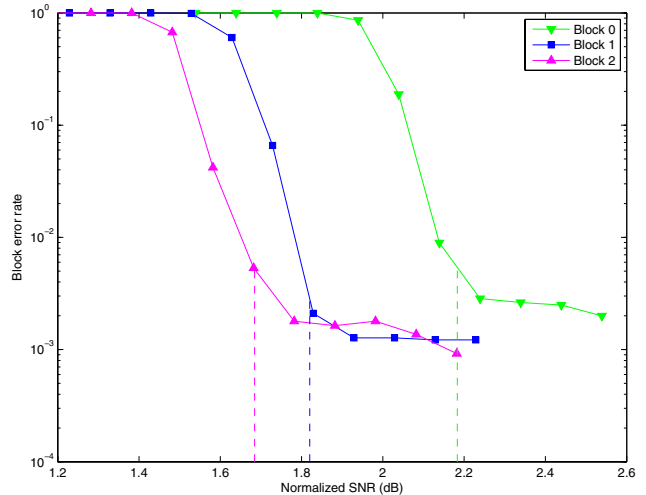


Fig. 6. Block error rate performance for the rateless code for Blocks 0, 1, and 2. Base rate-1/6 CCSDS code. Blocklength=16384. Normalized SNR = $\text{SNR}/(2^{2R/(m+1)} - 1)$. SNR threshold calculated from optimal power allocation shown as dashed line.

We simulate the AWGN channel around the SNR thresholds for Block 0, Block 1 and Block 2 using our optimal power allocation. In Fig. 6, we show the block error rate curves for Blocks 0, 1, and 2. The curves are plotted with respect to normalized SNR, $\text{SNR}_{\text{norm}} = \text{SNR}/(2^{2R/(m+1)} - 1)$. Thus capacity for each curve is at 0 dB. We show as dashed lines the SNR thresholds P/N_m for each block m as computed by the power allocation algorithm, and we see that they are in good agreement with the corresponding performance curve. Because capacity is at 0 dB, these SNR thresholds expressed as SNR_{norm} equal the gaps to capacity corresponding to the first three points of the $L = 12$ curve in Fig. 5.

We see that in all cases the error floor for the block error rate is about 10^{-3} , which is approximately $L = 12$ times higher than the error floor for the base code. This error floor is in line with predicted behavior, because a block must decode L consecutive base codewords in order to decode.

REFERENCES

- [1] A. Shokrollahi, "Raptor Codes," *IEEE Trans. Info. Th.*, vol. 52, pp. 2551-2567, June 2006.
- [2] Consultative Committee for Space Data Systems. "Recommendations for Space Data System Standards. Telemetry Channel Coding". CCSDS 101.0-B-4. Blue Book. May 1999.
- [3] S. Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the Design of Low-Density Parity-Check Codes Within 0.0045 dB of the Shannon Limit," *IEEE Commun. Lett.*, vol 5, Feb 2001, pp. 58-60.
- [4] D. Divsalar, S. Dolinar, and F. Pollara, "Iterative Turbo Decoder Analysis Based on Density Evolution", *IEEE J. Sel. Areas Commun.*, vol 19, May 2001, pp. 891-907.
- [5] U. Erez, G. Wornell and M. Trott, "Coding for Faster-than-Nyquist signaling: the merits of a regime change," 42nd Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, Sept. 29-Oct. 1, 2004.
- [6] U. Erez, G. W. Wornell, and M. D. Trott, "Rateless Coding and Perfect Rate-Compatible Codes for Gaussian Channels," in *Proc. Int. Symp. Inform. Theory (ISIT)*, (Seattle, WA), pp. 528-532, July 2006.
- [7] D. Warrier and U. Madhoo, "On the Capacity of Cellular CDMA with Successive Decoding and Controlled Power Disparities," in *Proc. 48th IEEE Vehicular Technology Conf.*, vol. 3, pp. 1873-1877, May 1998.