NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Shirvanimoghaddam, Mahyar; Johnson, Sarah "Raptor codes in the low SNR regime".
Published in IEEE Transactions on Communications Vol. 64, Issue 11, p. 4449-4460
(2016)

Available from: <http://dx.doi.org/10.1109/TCOMM.2016.2606410>

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting /republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Accessed from: <http://hdl.handle.net/1959.13/1344351>

Raptor Codes in the Low SNR Regime

Mahyar Shirvanimoghaddam, *Member, IEEE*, and Sarah Johnson, *Member, IEEE*

Abstract—In this paper, we revisit the design of Raptor codes for binary input additive white Gaussian noise (BI-AWGN) channels, where we are interested in very low signal to noise ratios (SNRs). A linear programming degree distribution optimization problem is defined for Raptor codes in the low SNR regime through several approximations. We also provide an exact expression for the polynomial representation of the degree distribution with infinite maximum output node degree in the low SNR regime, which enables us to calculate the exact value of the fractions of output nodes of small degrees. A more practical degree distribution design is also proposed for Raptor codes in the low SNR regime, where we include the rate efficiency and the decoding complexity in the optimization problem, and an upper bound on the maximum rate efficiency is derived for given design parameters. Simulation results show that the Raptor code with the designed degree distributions can approach rate efficiencies larger than 0.95 in the low SNR regime.

Index Terms—AWGN, Rateless codes, Raptor code, LT code, Low SNR regime.

I. INTRODUCTION

INSPIRED by Shannon’s revolutionary work and his seminal paper “A Mathematical Theory of Communication” published in 1948, a fundamental question has been posed in communications: *how can we efficiently and reliably transmit information?* Shannon also gave a basic answer: *coding can do it*. Since then, how to find the practical coding scheme that can approach the fundamental limits established by Shannon has been at the heart of coding theory and communications. This paper provides a practical framework to design a near-capacity code for an additive white Gaussian noise (AWGN) channel in the low signal-to-noise ratio (SNR) regime.

A. Fountain Codes

Fountain codes [1] are a class of codes on graphs designed mainly for reliable transmission over an erasure channel with unknown erasure probability, which can generate a potentially limitless number of coded symbols for a given set of information symbols [2]. The Luby Transform (LT) code [3] was the first practical realization of Fountain codes, where the degree of each coded symbol was obtained from a single probability mass function, referred to as the degree distribution. In LT codes, the Soliton distribution or robust Soliton distribution [4] was originally used as the degree distribution; however, a capacity approaching LT code was shown to have an average degree increasing with the message size with a logarithmic

slope, which makes it very hard to design a linear time encoder/decoder [2]. A more practical extension of LT codes, namely Raptor codes, were proposed by Shokrollahi in [5], which is a simple concatenation of a high rate low density parity check code (LDPC) and an LT code. The encoding and decoding of these codes are linear in terms of the message length; thus practical for several applications with large data transmission. Raptor codes were shown to be an excellent candidate for forward error correction in the application layer, where the goal is to recover from symbol erasures (rather than errors) [6–8].

LT and Raptor codes can also be used over other binary input channels [9]. Authors in [10] proposed protograph based Raptor like (PBRL) codes as a class of rate-compatible LDPC codes. These codes share some similarities with Raptor codes as new parity bits can always be generated by exclusive-OR operation on precoded bits. The authors in [2] studied Raptor code design for binary memoryless symmetric channels, where a Gaussian approximation was adopted to derive a simple linear programming optimization problem for the degree distribution design. A lower bound on the fraction of output symbols of degree 2 for a capacity approaching code was also proposed, which was shown to be channel dependent for binary input channels, unlike the erasure channel which allows channel independent capacity approaching codes. Accordingly, the authors in [2] proved that there is no universal Raptor code for binary input AWGN (BI-AWGN) channels and binary symmetric channels (BSC). An adaptive degree distribution for Raptor codes over BI-AWGN channels was designed in [11, 12], where the transmitter changes the degree distribution after sending a certain number of coded symbols. This approach has been shown to achieve rate efficiencies close to 90% for SNRs larger than -10 dB, where the optimization failed to provide good degree distributions for lower SNRs. In this paper, we study Raptor codes in the low SNR regime, where we design the degree distribution function and also formulate the asymptotic degree distribution function when SNR goes to zero and the maximum output node degree goes to infinity.

B. Motivations and Main Contributions

Many communication systems such as wireless ad-hoc networks, sensor networks, and ultra wide band networks, work under low-SNR conditions per node in the network, although the available degrees of freedom is large [13]. As shown in [14] in the slow-fading scenario, once a channel is in deep fade, channel coding may not help to increase the reliability of the transmission. If available, cooperative transmission can dramatically improve the performance by creating diversity using the antennas available at the other nodes of the network [15]. In many applications such as cellular-based machine-to-machine communications, dedicating resources, such as bandwidth or even relays, to a large number of users/devices is not

Manuscript received October 27, 2015; revised July 5, 2016 and August 22, 2016. The editor coordinating the review of this paper and approving it for publication was M. Ardakani.

The material in this paper was presented in part at the IEEE International Communications Conference (ICC), Kuala Lumpur, Malaysia, May 2016.

The authors are with the School of Electrical Engineering and Computer Science, The University of Newcastle, NSW, Australia (e-mail: {mahyar.shirvanimoghaddam, sarah.johnson}@newcastle.edu.au).

This work was supported by the Australian Research Council (ARC) under Grants LP130100783 and FT110100195.

efficient due to the random bursty nature of the transmission [16]. These necessitate the study of efficient transmission strategies for low-power communication systems in the low SNR regime, where increasing the transmission power or using cooperative techniques are not possible and channel coding is the only choice.

Very low rate error control codes are also required for the error reconciliation stage of quantum key distribution (QKD) systems. QKD is the first commercial application of quantum physics, which aims to provide a completely secure random key between two disjoint parties. More specifically, in continuous variable (CV) QKD systems, low SNRs are essential for the security of the system. Repeat-accumulate (RA) codes have been investigated to use in CV-QKD systems in [17], where the codeword length is 64800 and the lowest code rate is 1/20. Although, RA and LDPC codes have been proved in [18] to achieve the Shannon limit in AWGN channels when the coding rate approaches zero, in the low-rate region, both RA codes and LDPC codes suffer from performance loss and extremely slow convergence speed with iterative decoding. Multi-edge type LDPC (MET-LDPC) codes were shown in [19] to theoretically approach high efficiencies in low SNRs. Parallel-concatenated Zigzag-Hadamard (PCZH) codes are also proposed in [20] which offer the bit-error-rate (BER) of 10^{-5} at $E_b/N_0 = -1.2$ dB, which is only around 0.4 dB away from the Shannon limit. In [21], a new class of low-rate codes called repeat-zigzag-Hadamard (RZH) codes, were proposed which are essentially serially concatenated turbo-like codes where the outer code is a repetition code and the inner code is a punctured ZH code. Low-rate LDPC codes with simple protograph structure was also proposed in [22], which can be viewed as hybrid concatenations of simple modules such as accumulators, repetition codes, differentiators, and punctured single parity check codes. Existing very low rate codes, however have a very poor word error rate (WER) performance and so will only produce good efficiencies for theoretical threshold results or for finite-length results where the allowed WER is very high.

In this paper, we study Raptor codes over BI-AWGN channels in the low SNR regime. We use the general framework proposed in [2] for the degree distribution design of the LT code in the low SNR regime. We show that a simple linear program can be designed for the degree distribution optimization problem in the low SNR regime. An exact expression for the degree distribution function in the low SNR regime in the asymptotic case, where the maximum output node degree goes to infinity, is derived and shown to be a function of the elementary EXIT function, previously defined in [23]. We also characterize the fractions of low degrees through approximations of the proposed linear program. Using this approach, we characterize the rate efficiency, defined as the ratio of the maximum code rate and the channel capacity. A simple upper bound is also proposed for the rate efficiency. Similarly, the average output node degree is characterized and a simple lower bound is proposed. A more practical degree distribution design is proposed by taking into account both the rate efficiency and decoding complexity, where a set of degree distribution functions are also derived.

Fig. 1. Factor graph representation of a Raptor code truncated at codeword length $n + 1$.

C. Organization

The rest of the paper is organized as follows. Section II presents an overview of Raptor codes. In Section III, we investigate the Raptor code design in the low SNR regime. Section IV provides practical degree distribution design for Raptor codes in the low SNR regime and some bounds on the maximum rate efficiency of the code and the decoding complexity. Simulation results are shown in Section VI. Finally, Section VII concludes the paper.

II. PRELIMINARIES

A. Raptor Code Structure and Definitions

A binary Raptor code is constructed as the serial concatenation of an outer code V and an inner LT code C . The outer code is usually a high-rate LDPC code, which encodes a k -bit information vector into a k' -bit codeword of V , where the codeword bits are referred to as *input symbols*. Then, via an LT code a potentially limitless number of Raptor coded symbols, referred to as *output symbols*, are generated and transmitted over the channel, where each output symbol is the XOR (modulo-2 sum) of a randomly selected fraction of input symbols. More specifically, to generate each output symbol, first an integer d , called degree, is obtained from a predefined probability mass function, called the degree distribution function. An output symbol is then generated via the XOR operation on d distinct input symbols, uniformly selected at random. A Raptor code can then be characterized by $(k, V, \Omega(x))$, where $\Omega(x) := \sum_d \Omega_d x^d$ denotes the degree distribution polynomial and Ω_d denote the fraction of output nodes of degree d . A Raptor code can be represented by a factor graph, representing both LDPC and LT component codes (see Fig. 1).

The degree distribution can also be defined with respect to edges by $\omega(x) := \sum_d \omega_d x^{d-1}$, where ω_d is the fraction of edges connected to an output symbol of degree d . As shown in [5], ω_d and Ω_d can be calculated from each other as follows:

$$\omega_d = \frac{d\Omega_d}{\beta}, \quad \Omega_d = \frac{\omega_d/d}{\sum_d \omega_d/d}, \quad (1)$$

where $\beta = \sum_d d\Omega_d$ is the average output node degree. We can also compute the input node and edge degree distributions $\Lambda(x) := \sum_d \Lambda_d x^d$ and $\lambda(x) := \sum_d \lambda_d x^d$, respectively. It has been shown in [2] that in the asymptotic case, when k and n go to infinity, the input node and edge degree distribution

follow a Poisson distribution with parameter α , i.e.,

$$\Lambda(x) = \lambda(x) = e^{\alpha(x-1)}, \quad (2)$$

where $\alpha := \beta n/k'$ is the average input node degree and n is the number of output symbols. The realized rate of a Raptor code, denoted by R , is then defined as the ratio of k and the average number of coded symbols required to successfully decode the original k -bit message, i.e., $R = \frac{k}{\mathbb{E}[n]}$, where $\mathbb{E}$ is the expectation operator.

B. Decoding of Raptor Codes

Let y_i denote the channel output for output symbol c_j , where $y_j = (-1)^{c_j} + w_j$ and w_j is the zero mean Gaussian noise with variance σ_w^2 for a BI-AWGN channel. The channel SNR, γ , is then defined as $\gamma = 1/\sigma_w^2$. The channel log-likelihood ratio (LLR) is also defined as follows:

$$m_j^{(0)} := \ln \frac{P(c_j = 0|y_j)}{P(c_j = 1|y_j)} = \frac{2y_j}{\sigma_w^2}. \quad (3)$$

It is then easy to show that $m_j^{(0)}$ follows a symmetric Gaussian distribution with mean 2γ , when an all zero codeword is transmitted. Here, a probability density function (pdf) $f(x)$ is called symmetric if $f(x) = e^x f(-x)$ [24].

We consider the sum-product algorithm (SPA) [23] for the decoding of the Raptor code, where the LLR messages are passed along each edge in the factor graph between neighboring input symbols and output symbols in each iteration of the SPA. Let $m_{o \rightarrow i}^{(t)}$ and $m_{i \rightarrow o}^{(t)}$ denote messages passing from output node o to input node i and that from input node i to output node o , respectively, in the t^{th} iteration of the SPA. The message updating equations are formulated as follows [11]:

$$\begin{aligned} m_{o \rightarrow i}^{(t)} &= 2 \tanh^{-1} \left(\tanh \left(\frac{m_o^{(0)}}{2} \right) \prod_{i' \neq i} \tanh \left(\frac{m_{i' \rightarrow o}^{(t-1)}}{2} \right) \right), \\ m_{i \rightarrow o}^{(t)} &= \sum_{o' \neq o} m_{o' \rightarrow i}^{(t)}. \end{aligned} \quad (4)$$

After a predetermined number of iterations, T , the decoded LLR, m_i , for input node i is calculated as $m_i = \sum_o m_{o \rightarrow i}^{(T)}$. The decoded LLRs of the input nodes are passed to code V to recover the original information bits. The entire decoding process is repeated by increasing the number of output symbols, and accordingly their channel outputs, until the k -bit information vector is correctly decoded. Note that a decoder verifies the correctness of the decoding by the check equations of the outer code V, and the receiver sends an acknowledgment to the sender via a feedback channel upon the correct verification of the decoding.

C. Degree Distribution Optimization of Raptor Codes

In [2], a framework to construct a Raptor code for a given BI-AWGN channel has been proposed. This framework, referred to as the mean-LLR-exit chart approach, was proposed based on two assumptions. Firstly, the cycle-free assumption which means that the factor graph representation of Raptor codes is locally cycle-free, which is well justified when the graph is large and sparsely connected. This assumption enables

us to assume that all incoming SPA messages arriving at a given node are statistically independent. The second assumption is that the probability density of a message passed from an input symbol to an output node along a randomly chosen edge in the graph is a mixture of symmetric Gaussian distributions. This can be satisfied when the degree of each input node is relatively large, then according to the Central Limit Theorem the sum of incoming messages to an input node follows a Gaussian distribution. The mixture of Gaussian model is then due to the irregular degrees of input nodes [25].

Since we assume that the message from an input node to a neighboring output node follows a symmetric Gaussian distribution, its pdf can be characterized by its mean. Let $\mu^{(t)}$ denote the mean of message $m_{i \rightarrow o}^{(t)}$ along a randomly chosen edge in the t^{th} iteration of SPA. Similar to [25], we define a function $f_d(\mu)$, representing the mean of the message from an output node to an input node, as follows:

$$f_d(\mu) = 2 \mathbb{E} \left[\tanh^{-1} \left(\tanh \left(\frac{Z}{2} \right) \prod_{q=1}^{d-1} \tanh \left(\frac{X_q}{2} \right) \right) \right], \quad (5)$$

where X_q , which correspond to the LLR message from input node to output node, is the symmetric Gaussian random variable with mean μ and Z is a symmetric Gaussian random variable with mean 2γ representing the channel output LLR. As shown in [2], under the above assumptions and that the all-zero codeword is transmitted, we have:

$$\mu^{(t+1)} = \alpha \sum_d \omega_d f_d(\mu^{(t)}). \quad (6)$$

The goal of Raptor code design is then to find the output node or edge degree distribution to maximize the design rate of the LT code, defined as $R_{\text{design}} := \beta/\alpha$, such that $\mu^{(t+1)} > \mu^{(t)}$ for $t \geq 0$, so that the bit error rate (BER) of input symbols decreases through SPA iterations. For this aim, [2] proposed the following linear program to minimize $\alpha \sum_{d=1}^D \omega_d/d$, which is equivalent to maximizing R_{design} , for given α , D , and the target maximal message mean μ_o :

$$\text{minimize } \alpha \sum_{d=1}^D \omega_d/d \quad (7)$$

$$\text{s.t. (i). } \alpha \sum_{d=1}^D \omega_d f_d(\mu_j) > \mu_j, \quad \forall j = 1, \dots, N,$$

$$(ii). \sum_{d=1}^D \omega_d = 1,$$

$$(iii). \omega_d \geq 0, \quad \forall d = 1, \dots, D,$$

where $\{\mu_j | j = 1, \dots, N\}$ is a set of equally spaced values in range $(0, \mu_o]$ and $\mu_N = \mu_o$. It is important to note that μ_o is chosen to be large enough to ensure that the decoding of outer code V is successful. A joint optimization of the degree distribution and α was further proposed in [25] to maximize the design rate of the Raptor code at a given SNR, where higher rate efficiencies can be achieved compared to the codes obtained in [2].

III. RAPTOR CODES IN THE LOW SNR REGIME

In this section, we consider the asymptotic case where the channel SNR goes to zero and the maximum output node degree goes to infinity. We redefine the degree distribution optimization problem and provide the exact expression for the degree distribution polynomial in this asymptotic case.

A. Degree Distribution Optimization in the Low SNR Regime

Authors in [25] provided a detailed study of the feasibility of the linear program (7) based on parameters μ_o and α . More specifically, they showed that for SNRs higher than $\text{SNR}_{\text{low}} = \mu_o/2\alpha$ the linear program is feasible. The following lemma shows that the linear program (7) is always feasible for all SNRs larger than zero in the asymptotic case, where the average input node degree, α , goes to infinity.

Lemma 1: The linear program (7) is feasible for every SNR $\gamma > 0$, when α , the average input node degree, goes to infinity. The proof of this lemma is provided in Appendix A. It is important to note that in this paper, we are interested in the design of the degree distribution function for the Raptor code, when the channel SNR goes to zero. In this setting, the channel capacity also goes to zero. This means that for a given set of information symbols of size k , one should generate a potentially infinite number of coded symbols to be able to decode the message when the SNR goes to zero. As the degree of each coded symbol is at least one, the average input node degree, α , goes to infinity. Therefore, the linear program is feasible for every $\gamma > 0$, assuming that the average input node degree α goes to infinity.

Let us have a closer look at function $f_d(\mu)$ defined in (5), when the channel SNR, γ , goes to zero. We first define the function $h_d(Z, \mathbf{X}_d)$ as follows:

$$h_d(Z, \mathbf{X}_d) = 2 \tanh^{-1} \left(\tanh \left(\frac{Z}{2} \right) \prod_{q=1}^{d-1} \tanh \left(\frac{X_q}{2} \right) \right),$$

where $\mathbf{X}_d \triangleq (X_1, \dots, X_{d-1})$. Then, it is clear that $f_d(\mu) = \mathbb{E}[h(Z, \mathbf{X}_d)]$. We expand this function in a Taylor series about the mean value of Z , i.e., 2γ , that is:

$$\begin{aligned} h_d(Z, \mathbf{X}_d) &= h_d(2\gamma, \mathbf{X}_d) + (Z - 2\gamma) \frac{\partial h_d}{\partial Z} + \frac{(Z - 2\gamma)^2}{2} \frac{\partial^2 h_d}{\partial Z^2} + \dots, \end{aligned}$$

where the derivatives are evaluated at 2γ . By truncating the above equation at the second order terms and taking the expectation with respect to Z from both sides, we have:

$$\mathbb{E}_Z [h_d(Z, \mathbf{X}_d)] \approx h_d(2\gamma, \mathbf{X}_d) + \frac{1}{2} \frac{\partial^2 h_d}{\partial Z^2} \sigma_Z^2, \quad (8)$$

where $\mathbb{E}[Z] = 2\gamma$ and $\sigma_Z^2 = 4\gamma$, as Z follows a symmetric Gaussian distribution with mean 2γ . The second derivative of $h_d(Z, \mathbf{X}_d)$ with respect to Z at $Z = 2\gamma$ is given by:

$$\left. \frac{\partial^2 h_d}{\partial Z^2} \right|_{Z=2\gamma} = \frac{P_d(1 - P_d^2) \tanh(\gamma)(1 - \tanh^2(\gamma))}{1 - P_d^2 \tanh^2(\gamma)}, \quad (9)$$

where $P_d \triangleq \prod_{q=1}^{d-1} \tanh(X_q/2)$. Therefore, (8) can be rewritten as follows:

$$\begin{aligned} \mathbb{E}_Z [h_d(Z, \mathbf{X}_d)] &\approx 2 \tanh^{-1} (\tanh(\gamma) P_d) \\ &\quad + 2\gamma \frac{P_d(1 - P_d^2) \tanh(\gamma)(1 - \tanh^2(\gamma))}{1 - P_d^2 \tanh^2(\gamma)}. \end{aligned} \quad (10)$$

Since γ goes to zero, (10) can be further expanded in a Taylor series about zero as follows:

$$\mathbb{E}_Z [h_d(Z, \mathbf{X}_d)] \approx 2\gamma P_d + \mathcal{O}(\gamma^2),$$

which follows from $h_d(0, \mathbf{X}_d) = 0$, $\left. \frac{\partial^2 h_d}{\partial Z^2} \right|_{Z=0} = 0$, and $\left. \frac{\partial h_d}{\partial Z} \right|_{Z=0} = P_d$. Using only the linear term in the above approximation, we can approximate $f_d(\mu)$ as follows:

$$\begin{aligned} f_d(\mu) &= \mathbb{E}_{\mathbf{X}} [\mathbb{E}_Z [h_d(Z, \mathbf{X})]] \approx \mathbb{E}_{\mathbf{X}} \left[2\gamma \prod_{q=1}^{d-1} \tanh \left(\frac{X_q}{2} \right) \right] \\ &= 2\gamma \prod_{q=1}^{d-1} \mathbb{E} \left[\tanh \left(\frac{X_q}{2} \right) \right], \end{aligned} \quad (11)$$

which follows from the fact that Z and X_q are mutually independent random variables for $q = 1, \dots, d$. As X_q is assumed to follow a symmetric Gaussian distribution of mean μ , we can define function $\varphi(\mu)$ for $\mu \geq 0$ as follows:

$$\varphi(\mu) = \mathbb{E} \left[\tanh \left(\frac{X}{2} \right) \right] = \frac{\int_{-\infty}^{\infty} \tanh \left(\frac{u}{2} \right) e^{-\frac{(u-\mu)^2}{4\mu}} du}{\sqrt{4\pi\mu}}. \quad (12)$$

Thus, $f_d(\mu)$ can be further simplified as follows:

$$f_d(\mu) \approx 2\gamma [\varphi(\mu)]^{d-1}. \quad (13)$$

When SNR goes to zero, the left hand side of condition (i) in linear program (7) can be rewritten as follows:

$$\alpha \sum_d \omega_d f_d(\mu) \approx 2\gamma \alpha \sum_d \omega_d [\varphi(\mu)]^{d-1}. \quad (14)$$

It is important to note that when γ goes to zero, by using the well-know Maclaurin series of the function $\ln(1 + \gamma)$, the AWGN channel capacity, denoted by $C_b(\gamma)$, can be approximated as follows:

$$C_b(\gamma) = \frac{1}{2} \log_2(1 + \gamma) = \frac{1}{2 \ln(2)} (\gamma + \mathcal{O}(\gamma^2)). \quad (15)$$

Therefore, for a near capacity code (i.e., the code rate R is assumed to be very close to the channel capacity) and for a very small γ , we have:

$$\beta = \alpha R \approx \frac{\eta \alpha}{2 \ln(2)} (\gamma + \mathcal{O}(\gamma^2)), \quad (16)$$

where $\eta := \frac{R}{C_b(\gamma)}$ is defined as the rate efficiency. Thus by using (1) and (14), we have:

$$\begin{aligned} \alpha \sum_d \omega_d f_d(\mu) &\approx \frac{4 \ln(2) \beta}{\eta} \sum_d \omega_d [\varphi(\mu)]^{d-1} \\ &= \frac{4 \ln(2)}{\eta} \sum_d d \Omega_d [\varphi(\mu)]^{d-1}. \end{aligned} \quad (17)$$

Here, we choose to maximize the error correction capability of the code, which is equivalent to maximizing μ_o in such a way that conditions (i) – (iii) in the following optimization problem are satisfied:

$$\begin{aligned} &\text{maximize } \mu_o \\ &\text{s.t. (i). } \sum_{d=1}^D d \Omega_d [\varphi(\mu_j)]^{d-1} > \frac{\eta \mu_j}{4 \ln(2)}, \quad \forall j = 1, \dots, N, \\ &\quad \text{(ii). } \sum_{d=1}^D \Omega_d = 1, \\ &\quad \text{(iii). } \Omega_d \geq 0, \quad \forall d = 1, \dots, D, \end{aligned} \quad (18)$$

where $\mu_N = \mu_o$ and condition (i) was obtained by replacing (17) in the original linear program (7). We choose to maximize μ_o for a given D and η , to maximize the error correction

capability of the code as condition (i) in the above problem guarantees that the average message mean LLR of μ_o can be achieved by using the degree distribution obtained from this optimization problem for a Raptor code in the low SNR regime ($\gamma > 0$). Table I shows the maximum μ_o for different values of D when $\eta = 1$ (i.e., near capacity code). As can be seen in this table, by increasing the maximum degree, higher values of μ_o , and accordingly higher error correction capability for the Raptor code, is obtained as the linear program will have a higher degrees of freedom for larger D .

It is important to note that we have chosen discretized values for μ_j in the above optimization problem, as satisfying condition (i) for $0 \leq \mu \leq \mu_o$ is numerically impractical. Therefore, the results of the above linear program is not optimal, and it only gives suboptimal solutions. Moreover, the linear program (18) is an approximations of the original linear program (7) when the channel SNR goes to zero and the average input node degree goes to infinity. One could also solve the original linear program (7) in the low SNR regime by setting α to a very large value. However, this requires to set the maximum output node degree D to a very large value. Therefore, one need to calculate $f_d(\mu)$ for a large range of d , which is computationally complex. In (18), we approximated function $f_d(\mu)$ in the low SNR regime and simplifies the original linear program, so the optimizations can be easily performed in linear time.

B. An Exact Expression for the Degree Distribution Polynomial in the Low SNR Regime

We first summarize some of the properties of the function $\varphi(x)$ in the following proposition, which are used in the remainder of the paper.

Proposition 1: The function $\varphi(x)$, defined in (12), has the following properties:

- It is continuous, concave, and monotonically increasing with x .
- The inverse function $\varphi^{-1}(x)$ exists in $[0, 1]$, where $\varphi(\varphi^{-1}(x)) = x$. The inverse function is continuous, convex, and monotonically increasing with x .
- $\varphi(0) = \varphi^{-1}(0) = 0$.
- $\lim_{x \rightarrow \infty} \varphi(x) = 1$.
- $\lim_{x \rightarrow 1} \varphi^{-1}(x) = \infty$.
- $\lim_{\delta \rightarrow 0} \int_0^{1-\delta} \varphi^{-1}(t) dt = 4 \ln(2)$.

The proof of this proposition is provided in Appendix B. Now, we define $x_j \triangleq \varphi(\mu_j)$ for $j = 1, \dots, N$, and substitute it into constraint (i) in (18), which is then rewritten as follows:

$$\sum_{d=1}^D d\Omega_d x_j^{d-1} > \frac{1}{4 \ln(2)} \varphi^{-1}(x_j), \quad \forall j = 1, \dots, N, \quad (19)$$

where $x_N = \varphi(\mu_o)$. Eq. (19) can then be written in a more general form as follows:

$$\sum_{d=1}^D d\Omega_d x^{d-1} > \frac{1}{4 \ln(2)} \varphi^{-1}(x), \quad 0 \leq x \leq 1 - \delta_0, \quad (20)$$

where $\delta_0 = 1 - \varphi(\mu_o)$. This can be also rewritten as

$$\Omega'(x) > \frac{1}{4 \ln(2)} \varphi^{-1}(x), \quad 0 \leq x \leq 1 - \delta_0, \quad (21)$$

TABLE I
OPTIMIZED VALUE OF μ_o FOR DIFFERENT VALUES OF THE MAXIMUM DEGREE D WHEN $\eta = 1$.

D	μ_o	δ_0
50	16.22	0.0068
100	18.75	0.0034
200	21.31	0.0017
300	22.81	0.0011
500	24.71	0.0007
1000	27.35	0.0003

where $\Omega'(x)$ is the derivative of $\Omega(x)$ with respect to x . We can then reformulate the optimization problem to find the degree distribution which minimizes δ_0 (which is equivalent to maximizing μ_o) for a given maximum degree D such that (21) and constraints (ii)-(iii) in (18) are satisfied. By taking the integral of both sides of (21) and knowing that $\Omega(0) = 0$ and $\varphi^{-1}(0) = 0$, we have:

$$\Omega(x) > \frac{1}{4 \ln(2)} \int_0^x \varphi^{-1}(t) dt, \quad 0 \leq x \leq 1 - \delta_0. \quad (22)$$

The following lemma shows how the value of δ_0 changes with the maximum degree D .

Lemma 2: Let δ_D denote the minimum value of δ_0 such that (20) holds for $0 \leq x \leq 1 - \delta_0$ for a given maximum degree D and $\Omega^{(D)}(x)$ denote the respective degree distribution function. Then the following arguments hold:

- $\delta_D > 0$ for a finite D .
- δ_D is decreasing in D .

The proof of this Lemma is provided in Appendix C. As can be seen in Table I, δ_D decreases with D and it is very close to zero, which is in close agreement with the results of this lemma. The following theorem then gives the degree distribution function when $D \rightarrow \infty$ (see Appendix D for the proof).

Theorem 1: The degree distribution polynomial of a Raptor code, which satisfies (21), when $D \rightarrow \infty$ and SNR goes to zero is bounded by:

$$\left| \Omega^{(\infty)}(x) - \frac{1}{4 \ln(2)} \int_0^x \varphi^{-1}(t) dt \right| \leq \epsilon_\infty, \quad 0 \leq x < 1, \quad (23)$$

where $\Omega^{(\infty)}(x) \triangleq \lim_{D \rightarrow \infty} \Omega^{(D)}(x)$,

$$\epsilon_\infty = 1 - \frac{1}{4 \ln(2)} \int_0^{1-\delta_\infty} \varphi^{-1}(t) dt, \quad (24)$$

and $\delta_\infty \triangleq \lim_{D \rightarrow \infty} \delta_D$.

It is important to note that according to Lemma 1, there exists a degree distribution function with a potentially infinite maximum degree for every $\delta_0 > 0$, thus, Theorem 1 gives the exact expression for the degree distribution function when D goes to infinity. More specifically, as δ_D is decreasing with D and there exists at least one degree distribution function for each $\delta_0 > 0$, which satisfies condition (21), we can conclude that $\lim_{D \rightarrow \infty} \delta_D = 0$. Therefore, by using Theorem 1 and Lemma 2, we have $\epsilon_\infty = 0$ and accordingly the asymptotic degree distribution polynomial, $\Omega^{(\infty)}(x)$ is given by:

$$\Omega^{(\infty)}(x) = \frac{1}{4 \ln(2)} \int_0^x \varphi^{-1}(t) dt, \quad x \in [0, 1]. \quad (25)$$

C. Fraction of Output Symbols with Small Degrees

When x approaches zero, the left hand side of (25) can be realized as the Maclaurin series of $\frac{1}{4\ln(2)} \int_0^x \varphi^{-1}(t)dt$, which is given by:

$$\frac{1}{4\ln(2)} \int_0^x \varphi^{-1}(t)dt = \frac{1}{4\ln(2)} \sum_{i=1}^{\infty} \frac{x^i}{i!} \left[\frac{d^{i-1}}{dt^{i-1}} \varphi^{-1}(t) \right]_{t=0}.$$

The derivatives of an inverse function can be found as follows [26]:

$$\frac{d^i}{dt^i} \varphi^{-1}(t) = \frac{1}{(\varphi'(t))^{2i-1}} \det(\mathbf{A}^{(i)}), \quad (26)$$

where the elements of the matrix $\mathbf{A}^{(i)}$ are given by:

$$A_{k,\ell}^{(i)} = \begin{cases} \frac{(k-\ell+1)i-k}{i!} \varphi^{(k-\ell+1)}(t), & 1 \leq \ell \leq \min\{k, i-1\} \\ 1, & k=1 \text{ \& } \ell=i, \\ 0, & \text{otherwise.} \end{cases}$$

The optimized degrees can then be found as follows for $j > 1$:

$$\Omega_j = \frac{\eta}{4\ln(2)j!} \left[\frac{d^{j-1}}{dt^{j-1}} \varphi^{-1}(t) \right]_{t=0}. \quad (27)$$

As shown in (26), to find the asymptotic degree distribution, we need the derivatives of $\varphi(x)$ at $x=0$. They can be found by using the Maclaurin series of $\tanh(x)$ [27], as follows:

$$\begin{aligned} \varphi(x) &= \frac{1}{\sqrt{4\pi x}} \int_{-\infty}^{\infty} \tanh\left(\frac{u}{2}\right) e^{-\frac{(u-x)^2}{4x}} du \\ &= \frac{1}{\sqrt{4\pi x}} \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{2(2^{2n}-1)B_{2n}}{(2n)!} u^{2n-1} e^{-\frac{(u-x)^2}{4x}} du \\ &= \sum_{n=1}^{\infty} \frac{2(2^{2n}-1)B_{2n}}{(2n)!} M_{2n-1}(x, 2x), \end{aligned} \quad (28)$$

where B_{2n} is the $(2n)^{th}$ Bernoulli number [27] and $M_{2n-1}(x, 2x)$ is the $(2n-1)^{th}$ moment of a Gaussian distribution with mean x and variance $2x$, which can be found as follows:

$$M_{2n-1}(x, 2x) = \sum_{j=0}^{n-1} \binom{2n-1}{2j} (2j-1)!! (2x)^j x^{2n-2j-1}. \quad (29)$$

The double factorial for odd numbers is defined as $(2j-1)!! = (2j-1)(2j-3)\dots 5.3.1$ and can be also represented as follows:

$$(2j-1)!! = \frac{(2j)!}{2^j j!}. \quad (30)$$

By substituting (29) and (30) in (28), we have

$$\begin{aligned} \varphi(x) &= \sum_{n=1}^{\infty} \sum_{j=0}^{n-1} \frac{2^{j+1}(2^{2n}-1)B_{2n}}{(2n)!} \binom{2n-1}{2j} (2j-1)!! x^{2n-1-j} \\ &= \sum_{n=1}^{\infty} \sum_{j=0}^{n-1} \frac{2^{2n}-1}{n j! (2n-2j-1)!} B_{2n} x^{2n-1-j} \\ &= \sum_{\ell=1}^{\infty} \sum_{n=\lceil \frac{\ell+1}{2} \rceil}^{\ell} \frac{2^{2n}-1}{n(2n-\ell-1)!(2\ell-2n+1)!} B_{2n} x^{\ell}, \end{aligned} \quad (31)$$

where the last equality is derived by replacing $2n-j-1$ with ℓ and changing the bounds of the summations accordingly. By calculating the first few terms of the right hand side of (31),

we have:

$$\varphi(x) = \frac{1}{2}x - \frac{1}{4}x^2 + \frac{5}{24}x^3 - \frac{13}{48}x^4 + \frac{227}{480}x^5 + \dots$$

The i^{th} derivatives of $\varphi(x)$ at $x=0$ can then be calculated as follows for $i \geq 1$:

$$\varphi^{(i)}(0) = \sum_{n=\lceil \frac{i+1}{2} \rceil}^i \frac{(2^{2n}-1) i!}{n(2n-i-1)!(2i-2n+1)!} B_{2n}. \quad (32)$$

By substituting (32) in (26) and (27), and due to the fact that $\varphi(0)=0$, the value of the first few terms of the asymptotic degree distribution function can be found as follows:

$$\begin{aligned} \Omega_2 &= \frac{1}{4\ln(2)} = 0.3607, \quad \Omega_3 = \frac{1}{6\ln(2)} = 0.2404, \\ \Omega_4 &= \frac{1}{24\ln(2)} = 0.0601, \quad \Omega_5 = \frac{1}{10\ln(2)} = 0.1443. \end{aligned}$$

Higher degrees cannot be found through this approach as the approximation returns negative values. However, a numerical optimization can be performed for a relatively large maximum degree to obtain the complete degree distribution if required. In Appendix E, an alternative approach is proposed to find the asymptotic degree distribution function which is based on the fact that the left hand side of (25) can be realized as a pdf, thus its moments can be easily found. The degree distribution function can then be found through a system of linear equations.

D. A Lower Bound on the Average Output Node Degree

As the degree distribution function is a polynomial with non-negative coefficients, all its derivatives are non-negative in $[0, 1]$. It is then clear that for a Raptor code with the degree distribution function obtained from (18), we have:

$$\Omega^{(D)'}(x) \leq \Omega^{(D)'}(1), \quad (33)$$

and due to (21), we have:

$$\Omega^{(D)'}(1) > \frac{1}{4\ln(2)} \varphi^{-1}(1 - \delta_D). \quad (34)$$

Therefore, for degree distribution functions with maximum degree D , which satisfy (21) for $x \in [0, 1 - \delta_D]$, the average output node degree, denoted by $\beta^{(D)}$ is lower bounded by:

$$\beta^{(D)} > \frac{1}{4\ln(2)} \varphi^{-1}(1 - \delta_D). \quad (35)$$

IV. A MORE PRACTICAL DEGREE DISTRIBUTION DESIGN

The asymptotic degree distribution formulated in the previous section provides near capacity performance only in the theoretical sense, but in practice the guaranteed performance requires an infinite number of decoding iterations. It is well known that practical codes require a gap between curves in the EXIT chart to reduce the number of decoding iterations. In this section, we introduce the parameter ϵ to force this gap between the optimized degree distribution function and the asymptotic degree distribution function. For the decoder to proceed in iterations, ϵ should be larger than zero and the higher the value of ϵ the less number of iterations for the decoder to converge. It also has another effect on the decoders,

as larger ϵ provides higher tolerance to random changes in the convergence of the decoder due to noise. We formulate the optimization problem by taking ϵ and rate efficiency η into account, then provide bounds on the average output node degree, the maximal achievable rate efficiency, and the maximum number of decoding iterations, as a function of ϵ and η .

More specifically in the practical design problem, one should fix μ_o , D , and ϵ , and find the degree distribution which provides the desired level of error correction (i.e., μ_o) and maximize the rate efficiency when the SNR goes to zero. We then formulate the degree distribution optimization problem as follows, when γ goes to zero:

$$\begin{aligned} & \text{maximize } \eta \\ & \text{s.t. (i). } \sum_{d=1}^D d\Omega_d x_j^{d-1} \geq \frac{\eta(\varphi^{-1}(x_j) + \epsilon)}{4 \ln(2)}, \quad \forall j = 1, \dots, N, \\ & \text{(ii). } \sum_{d=1}^D \Omega_d = 1, \\ & \text{(iii). } \Omega_d \geq 0, \quad \forall d = 1, \dots, D. \end{aligned} \quad (36)$$

We choose to maximize the rate efficiency η , for a given D , μ_o and ϵ . The above optimization is linear for a given η , ϵ and D . To solve this problem, we gradually increase η and find the degree distribution which satisfies conditions (i) – (iii). Table II shows some degree distributions obtained by this approach for different values of D and their corresponding maximum efficiencies, when $\epsilon = 0.05$. It is important to note that there are several degree distributions which provide the same rate efficiency and satisfy constraints (i)–(iii) at the same time. In Table II, we showed the degree distribution with the maximum average output node degree, β . This is important for the practical implementation of Raptor codes in very low SNRs, where the code is operating over a given channel with a very low SNR, e.g., -30 dB, and according to (16) the rate is proportional to β assuming that the average input node degree is very large.

It is important to note that condition (i) in the above optimization guarantees that the fraction of degree 1 output nodes is at least $\eta\epsilon/(4 \ln(2))$, which is easily obtained by replacing x_j with 0 in condition (i). This will significantly reduce the number of decoding iterations while it slightly reduces the rate efficiency. We will provide more details on these in the following. The following lemma gives a lower bound on the average output node degree as a function of μ_o , ϵ , and η .

Lemma 3: The average output node degree for the near-capacity Raptor code in low SNRs is lower bounded by:

$$\sum_{d=1}^D d\Omega_d \geq \frac{\eta(\mu_o + \epsilon)}{4 \ln(2)}. \quad (37)$$

The proof of this lemma is provided in Appendix F. For a sufficiently large value of μ_o , $\varphi(\mu_o)$ is very close to 1, and the lower bound (37) would be very tight. For example, when $\mu_o = 40$ and $\epsilon = 0.05$, the average output node degree is lower bounded by 14.445η , which is very close to the respective value in Table II, for different maximum degrees. Fig. 2 shows the average output node degree of Raptor codes with degree distributions obtained from the above linear program versus

TABLE II
OPTIMIZED DEGREE DISTRIBUTIONS FOR DIFFERENT D AND THEIR CORRESPONDING MAXIMUM EFFICIENCIES, WHEN $\mu_o = 40$ AND $\epsilon = 0.05$.

D	50	100	200	300	500	1000
Ω_1	0.0155	0.0167	0.0173	0.0174	0.0175	0.0176
Ω_2	0.3140	0.3351	0.3451	0.3488	0.3517	0.3536
Ω_3	0.1764	0.2074	0.2300	0.2309	0.2335	0.2353
Ω_4	0.1508	0.1218	0.0601	0.0695	0.0667	0.0636
Ω_5	-	-	0.1075	0.0873	0.0948	0.1038
Ω_6	-	0.0067	-	0.0002	-	-
Ω_7	-	0.1380	-	0.0805	0.0664	0.0311
Ω_8	0.0599	-	0.0959	0.0004	0.0106	0.0577
Ω_9	0.0970	-	-	-	-	0.0023
Ω_{11}	-	-	-	0.0191	0.0539	-
Ω_{12}	-	-	-	0.0518	0.0087	0.0088
Ω_{13}	-	-	-	-	-	0.0463
Ω_{15}	-	-	0.0341	-	-	-
Ω_{16}	-	-	0.0307	-	-	-
Ω_{17}	-	0.0474	-	-	-	-
Ω_{18}	-	0.0356	-	-	-	-
Ω_{20}	-	-	-	-	0.0427	-
Ω_{22}	-	-	-	-	-	0.0211
Ω_{23}	-	-	-	0.0123	-	0.0133
Ω_{24}	-	-	-	0.0310	-	-
Ω_{38}	-	-	0.0291	-	-	-
Ω_{39}	-	-	0.0078	-	-	-
Ω_{41}	-	-	-	-	0.0251	0.0056
Ω_{42}	-	-	-	-	-	0.0151
Ω_{50}	0.1864	-	-	-	-	-
Ω_{59}	-	-	-	0.0220	-	-
Ω_{60}	-	-	-	0.0020	-	-
Ω_{85}	-	-	-	-	-	0.0021
Ω_{86}	-	-	-	-	-	0.0097
Ω_{100}	-	0.0913	-	-	-	-
Ω_{101}	-	-	-	-	0.0028	-
Ω_{102}	-	-	-	-	0.0109	-
Ω_{200}	-	-	0.0424	-	-	-
Ω_{211}	-	-	-	-	-	0.0026
Ω_{212}	-	-	-	-	-	0.0038
Ω_{300}	-	-	-	0.0268	-	-
Ω_{497}	-	-	-	-	0.0022	-
Ω_{498}	-	-	-	-	0.0125	-
Ω_{995}	-	-	-	-	-	0.0033
Ω_{996}	-	-	-	-	-	0.0031
β	12.4457	13.3772	13.8436	13.9957	14.1176	14.2248
η	0.8612	0.9253	0.9569	0.9668	0.9741	0.9790

the maximum degree, which are very close to the proposed lower bound, especially in the lower degrees. It is important to note that for a fixed μ_o , the gap between the average output node degree and its lower bound increases with the maximum degree. The reason is that with increasing D , condition (i) in (36) can be satisfied for a larger μ_o . One could jointly optimize η , μ_o , and $\Omega(x)$, so the average output node degree in this case would be very close to the lower bound with the suboptimal μ_o for each D .

The following lemma provides an upper bound on the maximal rate efficiency of a Raptor code with parameter ϵ when the SNR goes to zero.

Lemma 4: The maximum achievable rate efficiency η , is upper bounded by:

$$\eta < \frac{4 \ln(2)}{4 \ln(2) + \epsilon}. \quad (38)$$

The proof of this lemma is provided in Appendix G. For a nonzero ϵ , there is always a gap to the capacity, which can be characterized by (38) in terms of rate efficiency, η . Fig. 3 shows the maximum rate efficiency of a Raptor code versus the maximum degree. As can be seen in this figure, with increasing the maximum degree, the maximum efficiency get closer to the upper bound provided in (38).

Fig. 2. The average output node degree of the degree distributions obtained from (36) versus the maximum degree for different values of ϵ .

Fig. 3. The maximum rate efficiency of the degree distribution functions obtained from (36) versus the maximum degree for different values of ϵ .

A. A Discussion on Determining ϵ

The parameter ϵ in (36) determines the average number of decoding iterations in the SPA decoding. The maximum number of decoding iterations for a Raptor code with the maximum degree D , denoted by $n_D(\epsilon)$, can be upper bounded by:

$$n_D(\epsilon) \leq \frac{\mu_0}{\epsilon}, \quad (39)$$

which arises from condition (i) in (36), where the minimum gap between the left hand side and the right hand side of the inequality is considered to be ϵ . As can be seen in (39), the maximum number of decoding iterations is proportional to the inverse of ϵ , which can also be seen in Fig. 4. This shows that for practical implementation of Raptor codes, a small ϵ leads to a huge decoding complexity. Moreover, ϵ can be chosen such that the rate efficiency is high enough with a reasonable decoding complexity. For example as shown in Fig. 4, when the number of decoding iterations is set to 1000, ϵ must be set to greater than 0.01. Although, $\epsilon = 0.01$ gives a higher rate efficiency in theory, as will be shown in the next section, it leads to a worse efficiency in actual simulations when the SNR is not zero.

V. SIMULATION RESULTS AND DISCUSSION

In this section, we carry out simulations on Raptor codes in the low SNR regime. Fig. 5 shows the rate efficiency, η , of a Raptor code with the optimized degree distribution function

Fig. 4. The average number of decoding iterations and rate efficiency vs ϵ for a Raptor code with degree distributions obtained from (36) and different maximum degrees.

versus the SNR when the maximum number of iterations is 1000. We assume a rate 0.98 LDPC precoder which is constructed the same way as in [25]. As can be seen in this figure, the code achieves higher efficiencies in lower SNRs, which is due to the fact that the degree distributions were designed for the low SNR regime. Moreover, by increasing the message length, the rate efficiency improves which is due to the fact that the tree assumption and independence of the messages in the SPA decoding are more likely to be met as the message length increases. It is important to note that the efficiency at SNRs larger than -10 dB using the degree distributions obtained from Table II is not high enough which is due to the fact that these degree distributions were designed for very low SNRs. For SNR larger than -10 dB, one could solve the original linear program (7) to find the degree distribution, which has been shown that can achieve efficiencies larger than 92% at those SNRs [11]. The rate efficiencies for these codes have been also shown in Fig. 5

It is also important to note that the choice of ϵ plays and important role in the efficiency of the Raptor code. As can be seen in Fig. 5, when $\epsilon = 0.05$, the code achieves higher efficiencies compared to the code with $\epsilon = 0.01$. Although for infinite length codes, higher efficiencies can be achieved for a lower ϵ , in real applications where the message length is finite, ϵ cannot be very small, since otherwise the decoder does not converge due to the random nature of noise and the code graph.

We have also compared rate 1/10 and 1/50 MET-LDPC codes designed in [28], which were used in [19] for quantum key distribution systems, and several RA codes in [17] in Fig. 5. As shown, Raptor codes with an appropriately designed degree distribution can achieve higher efficiencies in comparison with the fixed rate codes in the entire SNR range. Moreover, the rate efficiency of fixed rate codes decreases with decreasing the word error rate, which is also clear from Fig. 5. For the sake of comparison, the rate efficiency of a few existing repeat ZigZag Hadamard codes [21] have been also plotted in Fig. 5.

Fig. 5. Rate efficiency versus the SNR for a Raptor code with the optimized degree distribution and different message lengths, different values of D and ϵ , when $\mu_o = 40$. The code length for RA codes [17] is 64800 and the code length for MET-LDPC is 10^6 [19]. The WER for rate 1/50 and 1/10 MET-LDPC codes are respectively $\{0.54, 0.16, 0.038, 0.021, 0.014, 0.012, 0.009\}$ and $\{1.8 \times 10^{-3}, 5.7 \times 10^{-4}, 1.2 \times 10^{-4}, 3 \times 10^{-5}, 1.5 \times 10^{-5}\}$ from top to the bottom.

VI. CONCLUSIONS

This paper provided a detailed study on Raptor codes in the low SNR regimes, where the degree distribution design was formulated and an exact expression for degree distribution function in the asymptotic case, when the maximum degree goes to infinity, was provided. Accordingly, the fraction of output nodes of small degrees were found, which are in a close agreement with the results obtained from the linear programming solution of the suboptimal degree distribution function. A more practical degree distribution design was also proposed, where the code efficiency and decoding complexity were taken into account when optimizing the degree distribution. An upper bound on the maximum rate efficiency of the designed Raptor code was also proposed. Simulation results show that Raptor code with the designed degree distribution functions closely approach the capacity of the AWGN channel in very low SNRs.

APPENDIX A PROOF OF LEMMA 1

The proof of this lemma follows directly from the proof of Lemma 2 and Theorem 1 in [25]. More specifically, $f_d(\mu)$ is shown to be increasing with μ for $d \geq 1$, and it also decreases with d for any μ and γ . Therefore, we have:

$$f_d(\mu) \leq f_1(\mu) = 2\gamma.$$

Thus, constraint (i) in (7) can be rewritten as follows:

$$\mu_i < \alpha \sum_{d=1}^D \omega_d f_d(\mu_i) \leq \alpha \sum_{d=1}^D \omega_d f_1(\mu_i) = \alpha f_1(\mu_i) = 2\alpha\gamma.$$

For the maximum value of μ_i , we also have $\mu_o < 2\alpha\gamma$. Now suppose that $\gamma > \mu_o/2\alpha$, then it is clear that $\omega(x) = 1$ satisfies all constraints in (7), thus the optimization problem is feasible. In an asymptotic case, when the average input node degree, α goes to infinity, the optimization problem is feasible for $\gamma > \lim_{\alpha \rightarrow \infty} \mu_o/2\alpha = 0$. This completes the proof.

APPENDIX B PROOF OF PROPOSITION 1

Statements (a)-(e) were previously proven in [25, 29], therefore we only provide the proof for (f). It is clear that for the inverse function we can write:

$$\int_0^{1-\delta} \varphi^{-1}(t) dt = \varphi^{-1}(1-\delta) - \int_0^{\varphi^{-1}(1-\delta)} \varphi(t) dt. \quad (40)$$

As shown in [29], $\varphi(x)$ can be represented by:

$$\varphi(x) = 1 - 4 \sum_{k=0}^{\infty} (-1)^k e^{x(k^2+k)} Q\left(\sqrt{\frac{x}{2}}(1+2k)\right), \quad (41)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$, which can also be calculated using the following integral [30]:

$$Q(x) = \frac{1}{\pi} \int_0^{\pi/2} e^{-\frac{x^2}{2 \sin^2 \theta}} d\theta. \quad (42)$$

By substituting (41) and (42) in (40), and since $\varphi^{-1}(1) = \infty$, we have

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \int_0^{1-\delta} \varphi^{-1}(t) dt \\ &= 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \int_0^{\pi/2} \int_0^{\infty} e^{x\left(k^2+k-\frac{(2k+1)^2}{4 \sin^2 \theta}\right)} dx d\theta \\ &= 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \int_0^{\pi/2} \left. \frac{e^{x\left(k^2+k-\frac{(2k+1)^2}{4 \sin^2 \theta}\right)}}{k^2+k-\frac{(2k+1)^2}{4 \sin^2 \theta}} \right|_{x=0}^{\infty} d\theta \\ &\stackrel{(i)}{=} 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \int_0^{\pi/2} \frac{4 \sin^2 \theta}{4(k^2+k) \cos^2 \theta + 1} d\theta \\ &\stackrel{(ii)}{=} 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \int_0^{\infty} \frac{4u^2}{(2k+1)^2 + u^2} \frac{du}{1+u^2} \\ &= 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \int_0^{\infty} \left(\frac{\frac{(2k+1)^2}{k^2+k}}{(2k+1)^2 + u^2} + \frac{\frac{-1}{k^2+k}}{1+u^2} \right) du \\ &= 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{\pi} \left[\frac{2k+1}{k^2+k} \tan^{-1} \left(\frac{k^2+k}{(2k+1)^2} u \right) - \frac{\tan^{-1}(u)}{k^2+k} \right]_{u=0}^{\infty} \\ &\stackrel{(iii)}{=} 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \stackrel{(iv)}{=} 4 \ln(2), \end{aligned}$$

where step (i) follows from the fact that $k^2+k-\frac{(2k+1)^2}{4 \sin^2 \theta} < 0$, in step (ii) we performed variable change $u = \tan \theta$, step (iii) follows from $\tan^{-1}(\infty) = \pi/2$, and finally step (iv) follows from the Maclaurin series of $\ln(1+x)$ for $x = 2$. This completes the proof.

APPENDIX C PROOF OF LEMMA 2

For the first condition, let us assume that there is a degree distribution function of maximum degree D , which satisfies the following inequality for $0 \leq x \leq 1$, i.e., $\delta = 0$:

$$\Omega'(x) \geq \frac{1}{4 \ln(2)} \varphi^{-1}(x),$$

then it is clear that $\Omega'(x)$ must go to infinity when $x \rightarrow 1$, as $\lim_{x \rightarrow 1} \varphi^{-1}(x) = \infty$. This contradicts the fact that the

maximum degree D is finite. This proves the first condition. The second condition is directly proven from the fact that by increasing the maximum degree the optimization problem will have a larger degree of freedom.

APPENDIX D PROOF OF THEOREM 1

For the degree distribution $\Omega^{(D)}(x)$ with maximum degree D , we have:

$$\Omega^{(D)'}(x) \geq \frac{1}{4\ln(2)} \varphi^{-1}(x) dt, \quad 0 \leq x \leq 1 - \delta_D, \quad (43)$$

and accordingly

$$\Omega^{(D)}(x) \geq \frac{1}{4\ln(2)} \int_0^x \varphi^{-1}(t) dt, \quad 0 \leq x \leq 1 - \delta_D.$$

We then define function $\Delta^{(D)}(x)$ as follows:

$$\Delta^{(D)}(x) = \Omega^{(D)}(x) - \frac{1}{4\ln(2)} \int_0^x \varphi^{-1}(t) dt,$$

which is obviously an increasing function of x due to (43) for $0 \leq x \leq 1 - \delta_D$. Therefore, we have:

$$0 \leq \Delta^{(D)}(x) \leq \Delta^{(D)}(1 - \delta), \quad 0 \leq x \leq 1 - \delta_D.$$

According to Lemma 2, δ_D is a decreasing function of D , therefore, we have:

$$\begin{aligned} & \lim_{D \rightarrow \infty} \Delta^{(D)}(1 - \delta_D) \\ &= \lim_{D \rightarrow \infty} \Omega^{(D)}(1 - \delta_D) - \lim_{D \rightarrow \infty} \frac{1}{4\ln(2)} \int_0^{1 - \delta_D} \varphi^{-1}(t) dt = 0 \\ &\leq 1 - \frac{1}{4\ln(2)} \int_0^{1 - \delta_D} \varphi^{-1}(t) dt, \end{aligned}$$

which arises from the fact that $\Omega(x) \leq 1$ for $x \in [0, 1]$. This completes the proof.

APPENDIX E

AN ALTERNATIVE APPROACH TO FIND THE ASYMPTOTIC DEGREE DISTRIBUTION OF THE RAPTOR CODE

By using Lemma 2, it can be verified that $g(x) \triangleq \frac{1}{4\ln(2)} \varphi^{-1}(x)$ is a probability distribution function defined in $[0, 1]$, as it is always positive and its integral over $[0, 1]$ is 1. Therefore, we can find the n^{th} moments of $g(x)$, referred to as $\mathcal{G}_n$, as follows:

$$\mathcal{G}_n = \int_0^1 t^n g(t) dt, \quad n = 0, 1, \dots \quad (44)$$

Also, from Theorem 1 and assuming that the optimization problem (7) is always feasible for D being sufficiently large and when γ goes to zero, we have:

$$\Omega^{(\infty)}(x) = \frac{1}{4\ln(2)} \int_0^x \varphi^{-1}(t) dt, \quad 0 \leq x \leq 1,$$

and equivalently,

$$\Omega^{(\infty)'}(x) = g(x), \quad 0 \leq x < 1. \quad (45)$$

By substituting (45) into (44), we have:

$$\mathcal{G}_n = \int_0^1 t^n \sum_{d=1}^{\infty} d\Omega_d t^{d-1} dt = \sum_{d=1}^{\infty} \frac{d}{d+n} \Omega_d,$$

which can be also represented as follows:

$$\frac{1 - \mathcal{G}_n}{n} = \sum_{d=1}^{\infty} \frac{1}{d+n} \Omega_d.$$

The above equation can be represented in matrix form as follows:

$$\begin{bmatrix} \frac{1-\mathcal{G}_1}{1} \\ \frac{1-\mathcal{G}_2}{2} \\ \frac{1-\mathcal{G}_3}{3} \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\ \cdots & \cdots & \cdots & \ddots \end{bmatrix} \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \vdots \end{bmatrix}$$

where the square matrix in the above equation is the Hilbert matrix, where the i^{th} entry in the j^{th} column, denoted by $h_{i,j}$, is $h_{i,j} = 1/(i+j-1)$ [31]. The entries of the inverse Hilbert matrix, denoted by H^{-1} , can be calculated as follows:

$$(H^{-1})_{i,j} = (-1)^{i+j} (i+j-1) \binom{n+i-1}{n-j} \binom{n+j-1}{n-i} \binom{i+j-2}{i-1}^2,$$

where n is the order of the Hilbert matrix [31]. Therefore, the asymptotic degree distribution can be calculated as follows:

$$\Omega_d = \sum_{j=1}^{\infty} (H^{-1})_{d+1,j} \frac{1 - \mathcal{G}_j}{j}.$$

APPENDIX F PROOF OF LEMMA 3

Condition (i) in (36) must hold for $x_0 = \varphi(\mu_0)$ for a near-capacity code. As $\Omega(x)$ and therefore $\Omega'(x)$ are continuous and monotonically increasing with x , we have:

$$\sum_{d=1}^D d\Omega_d \geq \sum_{d=1}^D d\Omega_d x_0^{d-1} \geq \frac{\eta(\varphi^{-1}(x_0) + \epsilon)}{4\ln(2)}$$

which directly results in (37).

APPENDIX G PROOF OF LEMMA 4

From (36), we have:

$$\frac{4\ln(2)}{\eta} \int_0^{\varphi(\mu_0)} \sum_d d\Omega_d x^{d-1} dx > \int_0^{\varphi(\mu_0)} (\varphi^{-1}(x) + \epsilon) dx,$$

which can be further simplified as follows:

$$\frac{4\ln(2)}{\eta} \sum_d \Omega_d [\varphi(\mu_0)]^d dx > \mu_0 - \int_0^{\mu_0} \varphi(x) dx + \epsilon \varphi(\mu_0).$$

Since $\lim_{x \rightarrow \infty} \varphi(x) = 1$ and $\sum_{d=1}^D \Omega_d = 1$, we can then rewrite the above inequality as follows:

$$\frac{4\ln(2)}{\eta} > \lim_{\mu_0 \rightarrow \infty} \left(\mu_0 - \int_0^{\mu_0} \varphi(x) dx \right) + \epsilon. \quad (46)$$

According to Lemma 2, we have:

$$\lim_{\mu_0 \rightarrow \infty} \left(\mu_0 - \int_0^{\mu_0} \varphi(x) dx \right) = \int_0^1 \varphi^{-1}(t) dt = 4\ln(2).$$

By replacing this in (46), (38) is easily derived.

REFERENCES

- [1] D. MacKay, "Fountain codes," *IEEE Proc. Commun.*, vol. 152, no. 6, pp. 1062 – 1068, Dec. 2005.
- [2] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2033 – 2051, May. 2006.
- [3] M. Luby, "LT codes," in *Proc. 43rd Annual IEEE Symp. Foundations Comput. Science*, Nov. 2002, pp. 271 – 280.
- [4] K.-K. Yen, Y.-C. Liao, C.-L. Chen, and H.-C. Chang, "Modified robust soliton distribution (MRSD) with improved ripple size for LT codes," *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 976–979, May 2013.
- [5] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551 – 2567, Jun. 2006.
- [6] I. De Fez, F. Fraile, R. Belda, and J. C. Guerri, "Analysis and evaluation of adaptive LDPC AL-FEC codes for content download services," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 641–650, 2012.
- [7] "Universal mobile telecommunications system (UMTS); LTE; multimedia broadcast/multicast service (MBMS); protocols and codecs (3GPP TS 26.346 v. 12.6.0)," ETSI, Technical Specification, Jul. 2015.
- [8] "Digital video broadcasting (DVB); upper layer FEC for DVB systems (ETSI TR 102 993 v. 1.1.1)," ETSI, Technical Specification, Feb. 2011.
- [9] R. Palanki and J. Yedidia, "Rateless codes on noisy channels," in *Proc. Intl. Symp. Inform. Theory (ISIT)*, Jul. 2004, p. 37.
- [10] T. Y. Chen, K. Vakilinia, D. Divsalar, and R. D. Wesel, "Protograph-based Raptor-like LDPC codes," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1522–1532, May 2015.
- [11] S.-H. Kuo, Y. L. Guan, S.-K. Lee, and M.-C. Lin, "A design of physical-layer Raptor codes for wide SNR ranges," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 491–494, Mar. 2014.
- [12] R. Barron, C. Lo, and J. Shapiro, "Global design methods for Raptor codes using binary and higher-order modulations," in *Proc. IEEE Military Comm. Conf.*, Nov. 2009, pp. 1 – 7.
- [13] T. Quek and H. Shin, "Bursty relay networks in low-SNR regimes," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 694–705, Feb. 2010.
- [14] A. Avestimehr and D. Tse, "Outage capacity of the fading relay channel in the low-SNR regime," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1401–1415, Apr. 2007.
- [15] B. Wang, J. Zhang, and L. Zheng, "Achievable rates and scaling laws of power-constrained wireless sensory relay networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4084–4104, Sep. 2006.
- [16] M. Shirvanimoghaddam, Y. Li, M. Dohler, B. Vucetic, and S. Feng, "Probabilistic rateless multiple access for machine-to-machine communication," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2015.
- [17] S. J. Johnson, V. A. Chandrasekty, and A. M. Lance, "Repeat-accumulate codes for reconciliation in continuous variable quantum key distribution," in *Australian Commun. Theory Workshop (AusCTW)*, Jan. 2016, pp. 18–23.
- [18] H. Jin and R. McEliece, "RA codes achieve AWGN channel capacity," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, ser. Lecture Notes in Computer Science, M. Fossorier, H. Imai, S. Lin, and A. Poli, Eds. Springer Berlin Heidelberg, 1999, vol. 1719, pp. 10–18.
- [19] P. Jouguet, S. Kunz-Jacques, and A. Leverrier, "Long-distance continuous-variable quantum key distribution with a Gaussian modulation," *Phys. Rev. A*, vol. 84, p. 062317, Dec. 2011.
- [20] W. Leung, G. Yue, L. Ping, and X. Wang, "Concatenated Zigzag Hadamard codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1711–1723, Apr. 2006.
- [21] K. Li, G. Yue, X. Wang, and L. Ping, "Low-rate repeat-Zigzag-Hadamard codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 531–543, Feb. 2008.
- [22] D. Divsalar, S. Dolinar, and C. Jones, "Low-rate LDPC codes with simple protograph structure," in *Proc. Intl. Symp. Inf. Theory*, Sep. 2005, pp. 1622–1626.
- [23] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [24] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
- [25] Z. Cheng, J. Castura, and Y. Mao, "On the design of Raptor codes for binary-input Gaussian channels," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3269–3277, Nov. 2009.
- [26] P. G. Todorov, "New explicit formulas for the n th derivative of composite functions," *Pacific J. Math.*, vol. 92, no. 1, pp. 217–236, 1981.
- [27] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Corporation, 1964, no. 55.
- [28] T. Richardson, R. Urbanke *et al.*, "Multi-edge type LDPC codes," in *Workshop honoring Prof. Bob McEliece on his 60th birthday, California Institute of Technology, Pasadena, California*, 2002, pp. 24–25.
- [29] S.-Y. Chung, T. Richardson, and R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657–670, Feb. 2001.
- [30] J. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Nov. 1991, pp. 571–575 vol.2.
- [31] M.-D. Choi, "Tricks or treats with the Hilbert matrix," *The American Mathematical Monthly*, vol. 90, no. 5, pp. 301–312, 1983.

Mahyar Shirvanimoghaddam Mahyar Shirvanimoghaddam received the B. Sc. degree with 1st Class Honours from University of Tehran, Iran, in September 2008, the M. Sc. Degree with 1st Class Honours from Sharif University of Technology, Iran, in October 2010, and the Ph.D. degree from The University of Sydney, Australia, in January 2015, all in Electrical Engineering. He then held a research assistant position at the Centre of Excellence in Telecommunications, School of Electrical and Information Engineering, The University of Sydney, before coming to the University of Newcastle, Australia, where he is now a Postdoctoral Research Associate at the School of Electrical Engineering and Computer Science. His general research interests include channel coding techniques, cooperative communications, compressed sensing, machine-to-machine communications, and wireless sensor networks.

Sarah J. Johnson received the B.E. degree in electrical engineering, and the Ph.D. degree from the University of Newcastle, Australia, in 2000 and 2004, respectively. She held a postdoctoral position at NICTA, Australia's Information and Communications Technology Research Centre of Excellence, before returning to the University of Newcastle, where she is now an Associate Professor and Australian Research Council Future Fellow. Her research interests include error correction codes and information theory for multiple-user communication networks.

She has authored a book entitled *Iterative Error Correction* (Cambridge University Press).