Research Article

An Efficient Design of Systematic Rateless Codes

Jingke Dai , Fenggan Zhang, Wenke Yang, and Kai Kang

School of Military Operational Support, Rocket Force University of Engineering, Xi'an 710025, China

Correspondence should be addressed to Jingke Dai; djk029@163.com

Received 20 October 2019; Revised 20 February 2020; Accepted 5 May 2020; Published 19 May 2020

Academic Editor: L. J. García Villalba

Copyright © 2020 Jingke Dai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The asymptotic performance of systematic rateless codes is first analysed using Gaussian approximation (GA) based on mutual information, which provides more accurate decoding thresholds than the method based on message mean. A modified linear program (LP) algorithm is proposed, where two key parameters are precomputed to efficiently search degree distributions with low overhead and appropriate average degree. The asymptotic analysis and simulation results show that the optimized code outperforms the codes obtained by conventional LP. Furthermore, the effects of outer code on overall code rate and decoding complexity are discussed.

1. Introduction

Rateless codes keep generating symbols until successful decoding; thus, they are more robust than traditional fixed-rate codes on noisy channels. The first practical rateless code is Luby Transform (LT) code [1], the extension of which is Raptor code constructed by an inner LT code serially concatenated with an outer high-rate low-density parity-check (LDPC) code [2]. They are originally proposed on a binary erasure channel (BEC) and then extended to noisy channels [3, 4], where the well-known sum-product algorithm (SPA) is employed to decode.

Many asymptotic analysis tools have been implemented to analyse and design LT code and Raptor code on binary input additive white Gaussian noise (BIAWGN) channels, such as Gaussian approximation based on message mean (GA-mean) [5, 6], extrinsic information transfer (EXIT) chart [7–9], and discretized density evolution (DDE) [10]. Recently, the systematic version of rateless codes have received extensive attention due to their several advantages over the nonsystematic codes. For example, the original information bits may be directly visible in the codewords of systematic codes; nonsystematic rateless codes usually need at least one symbol with degree-1 to start decoding [5] while systematic codes have a lot; systematic codes may have smaller generating/decoding matrix thanks to the systematic part and then lower encoding/decoding complexity may be achieved.

Systematic LT (SLT) and systematic Raptor (SR) codes on BEC are proposed in [2, 11], respectively. However, on noisy channels, SLT codes are generated in a straightforward manner [12], i.e., the original symbols are first transmitted followed by LT-coded symbols. If an inner SLT code is serially concatenated with a high-rate outer code, the resulting code can also be referred as an SR code [13]. The SLT codes are analysed and designed based on the GA-mean method in [14, 15], but this method may introduce some nonignorable errors of message density resulting in poor performance of designed codes [5]. DDE provides an accurate asymptotic result, and SR codes have been analysed in [13], but this method has pretty high complexity because it is essentially an infinite-dimensional analysis.

The EXIT chart has the same calculation complexity as GA-mean since both are one-dimensional analysis, but the former may provide asymptotic results closer to DDE, which have been shown in the analysis of LDPC codes [16] and SR codes [17]. As the EXIT chart is also based on Gaussian approximation and the tracking variable in decoding iteration is mutual information, we refer it as Gaussian approximation based on mutual information (GA-MI) in the following. The contributions of this paper are summarized as follows. First, we provide the asymptotic curves of SLT codes and SR codes on BIAWGN channels using GA-MI, which is very close to DDE. Second, we provide a modified linear program (LP) algorithm to design SR codes, where

two important optimization parameters are precomputed according to the requirement of outer code and some lower bound analysis. The optimized degree distribution has lower overhead and appropriate average degree than those obtained by conventional LP. Moreover, the balance of code-rate and decoding complexity is discussed.

2. Encoding and Decoding of Systematic Rateless Codes

Like conventional Raptor codes, the SR codes include two stages of encoding and decoding. Firstly, the source symbols $\mathbf{s} = [s_1, s_2, \dots, s_K]$ are precoded using a high-rate LDPC code to produce the intermediate symbols $\mathbf{e} = [e_1, e_2, \dots, e_{K'}]$. Note the outer code rate $R_o = K/K'$. Then, the intermediate symbols are encoded using an inner SLT code to generate output symbols as given by $\mathbf{c} = [c_1, c_2, \dots, c_N] = \mathbf{e} \times \mathbf{G}_{\text{SLT}}$, where the generation matrix $\mathbf{G}_{\text{SLT}} = [\mathbf{I}_{K'}, \mathbf{G}_{\text{LT}}]$, $\mathbf{I}_{K'}$ is an identity matrix, and $\mathbf{G}_{\text{LT}}$ is the generation matrix of size $K' \times (N - K')$ like conventional LT code [15]. The rates of inner SLT code and SR code are $R_i = K'/N$ and $R = K/N$, respectively.

The output symbols are transmitted on BIAWGN channels with binary phase shift keying modulation modelled as $y_i = x_i + n_i$, where $x_i = 1 - 2c_i$ and n_i is a zero-mean Gaussian noise with variance σ^2 . The channel log-likelihood ratio (LLR) is given by $z_i = 2y_i/\sigma^2$. Two-step sequential decoding is considered [13], where the decoding of inner SLT codes is followed by decoding of outer LDPC codes. The SPA is implemented in the inner decoder according to the bipartite graph related to $\mathbf{G}_{\text{LT}}$, which is constructed using intermediate symbols and output symbols as variable nodes (VNs) and check nodes (CNs). The message update rule between VNs and CNs is given by

$$u_{m,n}^{(l)} = 2 \tanh^{-1} \left[\tanh \left(\frac{z_{K'+n}}{2} \right) \prod_{k \in \{S_n\} \setminus m} \tanh \left(\frac{v_{n,k}^{(l)}}{2} \right) \right], \quad (1)$$

$$v_{n,m}^{(l+1)} = z_m + \sum_{k \in \{S_m\} \setminus n} u_{m,k}^{(l)}, \quad (2)$$

where $u_{m,n}^{(l)}$ and $v_{n,m}^{(l)}$ denote the message passed from the n th check node to the m th variable node and the message passed in the opposite direction at the l th decoding iteration, and $\{S_n\} \setminus m$ represents the set of all neighbours of n th node except for the m th node. After the maximum iteration $l_{\max}$, the decision $v_m^{(l_{\max})} = z_m + \sum_{k \in \{S_m\}} u_{m,k}^{(l_{\max}-1)}$ is provided as the initial LLR to the outer decoder.

3. Asymptotic Analysis Based on Mutual Information

3.1. GA-MI for Inner Code. Using polynomial notations, the CN degree distribution of the inner SLT code is written compactly as $\Omega(x) = \sum_{j=1}^{d_c} \Omega_j x^j$, where d_c is the maximum degree. The corresponding edge degree distribution is calculated by $\omega(x) = \sum_{j=1}^{d_c} \omega_j x^{j-1} = \Omega'(x)/\Omega'(1)$, and the average degree of

CNs is given by $\beta = \Omega'(1)$. The average degree of VNs is denoted by α , and the variable node and edge degrees are both assumed to be Poisson distributed as $\Lambda(x) = e^{\alpha(x-1)}$ and $\lambda(x) = e^{\alpha(x-1)}$, respectively [7]. They can also be truncated to obtain polynomials as $\Lambda(x) = \sum_{i=1}^{d_v} \Lambda_i(\alpha) x^i$ and $\lambda(x) = \sum_{i=1}^{d_v} \lambda_i(\alpha) x^{i-1}$, where d_v is large enough to ensure $\Lambda(1) \approx 1$. The overhead of the inner SLT code is defined by $\varepsilon = (N - K')/K'$ which is related to the code rates as $\varepsilon = 1/R_i - 1 = R_o/R - 1$.

Assuming the all-zero codeword is transmitted, the initial LLR has the symmetric Gaussian density as $z \sim N(2/\sigma^2, 4/\sigma^2)$. The mutual information of z is computed by [8]

$$J(\sigma) = 1 - \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp \left[-\frac{(\xi - \sigma^2/2)^2}{2\sigma^2} \right] \log_2(1 + e^{-\xi}) d\xi, \quad (3)$$

where $J(\sigma)$ is a monotonically increasing function and then has an inverse function $\sigma = J^{-1}(I)$. Both $J(\cdot)$ and $J^{-1}(\cdot)$ do not have close-form expressions, but they can be closely approximated as suggested in [18].

From (2), we know that the output of VNs is the sum of message of channel LLR and CNs; thus, the mutual information passed from variable nodes at the l th iteration is calculated by

$$I_{E,V}^{(l)}(I_{E,C}^{(l-1)}) = \sum_{i=1}^{d_v} \lambda_i(\alpha) J \left(\sqrt{\sigma_{ch1}^2 + (i-1)\sigma_V^2} \right), \quad (4)$$

where the $\sigma_{ch1}^2 = 4/\sigma^2$ is the variance of initial LLR and $\sigma_V^2 = [J^{-1}(I_{E,C}^{(l-1)})]^2$ is the variance of message output by check nodes. The mutual information passed by CNs at the l th iteration is like conventional LT codes as given by

$$I_{E,C}^{(l)}(I_{E,V}^{(l)}) = \sum_{j=1}^{d_c} \omega_j \left[1 - J \left(\sqrt{\sigma_{ch2}^2 + (j-1)\sigma_C^2} \right) \right], \quad (5)$$

where $\sigma_{ch2} = J^{-1}[1 - J(\sqrt{4/\sigma^2})]$ and $\sigma_C = J^{-1}(1 - I_{E,V}^{(l)})$. Combining (4) and (5), we can get the iteration expression of $I_{E,C}^{(l)}$ as the function of $I_{E,C}^{(l-1)}$, σ , and α :

$$\begin{aligned} I_{E,C}^{(l)} &= F(I_{E,C}^{(l-1)}, \sigma, \alpha) \\ &= \sum_{j=1}^{d_c} \omega_j \left[1 - J \left(\sqrt{\sigma_{ch2}^2 + (j-1) \left(J^{-1} \left(1 - I_{E,V}^{(l-1)}(I_{E,C}^{(l-1)}) \right) \right)^2} \right) \right]. \end{aligned} \quad (6)$$

At the end of iteration $l_{\max}$, the inner decoder outputs the extrinsic information provided to the outer decoder as

$$I_{\text{ext}} = \sum_{i=1}^{d_v} \Lambda_i J \left(\sqrt{\frac{4}{\sigma^2} + i \left(J^{-1} \left(I_{E,C}^{(l_{\max})} \right) \right)^2} \right), \quad (7)$$

and the corresponding error probability is calculated as

$$P_e^{\text{SLT}} = \sum_{i=1}^{d_v} \Lambda_i(\alpha) Q \left[\sqrt{\frac{1}{\sigma^2} + \frac{i \left(J^{-1} \left(I_{E,C}^{(l_{\max})} \right) \right)^2}{4}} \right], \quad (8)$$

where the function $Q(x) = 1/\sqrt{2\pi} \int_x^{+\infty} e^{-t^2/2} dt$.

To compare the accuracy of different methods, we plot the asymptotic bit-error-rate (BER) of the SLT code using the above GA-MI method and GA-mean method in Figure 1, where the degree distribution takes the optimized result of $\sigma = 0.977$ in [15]. Also presented are the results by DDE and simulated BER with finite symbol length K' . It is shown that the curve of GA-MI is much closer to the curve of DDE than GA-mean, and the simulation BERs of finite-length codes converge to the results of GA-MI with the increase of K' . Note that the asymptotic BER curve by GA-mean is lower than the ones by DDE and GA-MI in the water-fall region of overhead. It is not to say that the GA-mean method has better performance than the other two methods but the approximated message density represented by its mean has a relatively larger deviation from the actual density given by DDE. Therefore, we say that the GA-MI provides a more accurate asymptotic performance than the GA-mean method. In addition, the line of critical BER in Figure 1 is explained in the next subsection.

3.2. GA-MI for Outer Code and Critical BER. The extrinsic information I_{ext} output by an inner decoder is taken as a priori information of outer decoder, and the asymptotic performance of outer codes can also be obtained by GA-MI. Assuming the message output by an inner decoder is Gaussian distrib-

FIGURE 1: Comparison of asymptotic and simulated BER of SLT codes, $\sigma = 0.977$.

uted [10], the variance of initial LLR of an outer decoder is calculated by

$$\sigma_{o, ch} = J^{-1}(I_{\text{ext}}). \quad (9)$$

We denote the edge degree distributions of variable nodes and check nodes of outer LDPC codes as $\gamma(x) = \sum_{i=1}^{d_{o,v}} \gamma_i x^{i-1}$ and $\rho(x) = \sum_{j=1}^{d_{o,c}} \rho_j x^{j-1}$. The exchange of mutual information is analogous to the SLT code, and then, the iteration expression is given by

$$I_{o,E,C}^{(l)} = \sum_{j=1}^{d_{o,c}} \rho_j \left[1 - J \left(\sqrt{(j-1) \left[J^{-1} \left(1 - \sum_{i=1}^{d_{o,v}} \gamma_i J \left(\sqrt{\sigma_{o, ch}^2 + (i-1) \left(J^{-1} \left(I_{o,E,C}^{(l-1)} \right) \right)^2} \right) \right]^2} \right) \right) \right]. \quad (10)$$

The decoder outputs the error probability after the last iteration as

$$P_e^{\text{SR}} = \sum_{i=1}^{d_{o,v}} \Gamma_i Q \left[\sqrt{\frac{\left(\sigma_{o, ch}^2 + i \left[J^{-1} \left(I_{o,E,C}^{(l_{\max})} \right) \right]^2 \right)}{4}} \right], \quad (11)$$

where $\Gamma(x) = \sum_{i=1}^{d_{o,v}} \Gamma_i x^{i-1}$ is the variable node degree distribution.

As shown in [10], the outer code would be successfully decoded if the BER of inner code lowers to a certain level, which is called as “critical BER.” Combining (7)–(11), we can calculate the critical BER which is denoted by P_e^c , i.e., $P_e^{\text{SR}} \rightarrow 0$ if $P_e^{\text{SLT}} < P_e^c$. The critical BER for a 50/51-rate regular (4,204)-LDPC code is $P_e^c = 3.78 \times 10^{-3}$, which is very close to the results of 3.83×10^{-3} obtained by DDE in [10]. This critical BER is also depicted in Figure 1, and the corresponding overheads obtained by GA-MI and DDE are between 1.3 and 1.4, which could be taken as the decoding thresholds of SR codes. Furthermore, we see that the

employed code takes a wide range of overhead (from 0.9 to 1.4) to reduce to the critical BER; therefore, the performance may be further improved if we can narrow this region.

4. Efficient Design of Rateless Codes

The outer code is fixed as a high-rate LDPC code, and we concentrate the optimization on inner SLT codes. The classic linear program algorithm is combined with GA-MI to design SLT codes, which is described as follows.

$$\begin{aligned} & \text{Minimizing } \frac{\sum_{j=1}^{d_c} \omega_j}{j} \\ & \text{subject to } F(I_{E,C}^{(l-1)}, \sigma, \alpha) > I_{E,C}^{(l-1)}, \quad I_{E,C}^{(l-1)} \in [0, I_{E,C}^{\max}], l = 1, 2, \dots, L \\ & \sum_{j=1}^{d_c} \omega_j = 1 \\ & \omega_j \geq 0, j = 1, \dots, d_c \end{aligned} \quad (12)$$

where the function $F(I_{E,C}^{(l-1)}, \sigma, \alpha)$ is defined in (6) and L is an integer indicating the maximum number of sampling points of $I_{E,C}^{(l-1)}$. The objective of the above design model indicates the minimized overhead as $\varepsilon = \alpha/\beta = \alpha \sum_{j=1}^{d_c} \omega_j/j$. The second and third constraints ensure the distribution validity. The first constraint ensures that the tracking variable, i.e., mutual information from CNs, converges to a fixed value $I_{E,C}^{\max}$ in decoding iterations. Note that $I_{E,C}^{\max}$ is the maximum extrinsic information output by CNs, different from the minimum I_{ext} (see Equation (7)) which is the extrinsic information required by an outer decoder for successful decoding.

The selection of parameters $I_{E,C}^{\max}$ and α is very important because they affect the performance of optimization results in a great measure. If $I_{E,C}^{\max}$ is too small, the results cannot achieve the critical BER quickly, whereas the average degree β , which indicates decoding complexity to some extent, will be increased if $I_{E,C}^{\max}$ is too large. If α is too small, the optimization process may not succeed, while the decoding complexity will be increased if α is too large considering $\beta = \alpha/\varepsilon$. Furthermore, $I_{E,C}^{\max}$ and α influence each other (see Equation (8)) so that it is easy to fall into the circle of “chicken and egg, which came first” problem.

The conventional LP method chooses these two parameters separately by experience or many attempts [5, 6, 9], which may not find the best result or even not converge successfully if the selected parameters are inappropriate. A rough range of these two parameters is given in [7] but no exact result is provided. In [15], a new constraint condition about the lower bound of error probability is added, but the results are not suitable for SR codes (see the analysis for Figure 1 in former section). Hence, we propose a new approach to efficiently locate the values of these two key parameters in the following, using which the LP search is definitely convergent and the resulting codes have excellent performance.

Step 1. Assume σ is fixed and rewrite the error probability of Equation (8) to a binary function as given by

$$f(\alpha, I_{E,C}) = \sum_{i=1}^{d_r} \Lambda_i(\alpha) Q \left[\sqrt{\frac{1}{\sigma^2} + \frac{i(J^{-1}(I_{E,C}))^2}{4}} \right], \quad (13)$$

where the coefficients $\Lambda_i(\alpha) = \alpha^i e^{-\alpha/i!}$ [7]. Note that $f(\alpha, I_{E,C})$ is monotonically decreasing relative to either of two inputs.

Step 2. Calculate the bounds of two parameters, namely, $I_{E,C}^{ub}$ and α^{lb} . As in [8], the LLR output by VNs may be perfect ($\tanh v_{n,k}^{(l)} \rightarrow 1$) in the final iteration; from (1), we know the output of CNs is equal to the channel LLR, which is a zero-mean Gaussian variable with variance $4/\sigma^2$. Therefore, the upper bound of mutual information output by CNs is

$$I_{E,C}^{ub} = J \left(\sqrt{\frac{4}{\sigma^2}} \right). \quad (14)$$

Assuming the critical BER as P_e^c , the required average degree is

$$\alpha^{lb} = f_{\alpha}^{-1} \left(P_e^c, I_{E,C}^{ub} \right), \quad (15)$$

where $f_{\alpha}^{-1}(P_e^c, I_{E,C}^{ub})$ is the reverse function of (13) related to α . In consider of the monotonicity of (13), α^{lb} is the lowest possible value of α satisfying P_e^c . However, $I_{E,C}^{ub}$ is not achieved for small overheads, thus the selected α must be larger than α^{lb} .

Step 3. Set the sampled average degree as $\alpha^* = \alpha^{lb} + t \cdot \Delta\alpha$, where $t = 1, 2, 3, \dots, t_{\max}$, $\Delta\alpha$ is the sampling interval, and $t_{\max}$ controls the maximum α^* . Then, get the maximum value of mutual information corresponding to α^* as given by

$$I_{E,C}^{\max} = f_I^{-1}(P_e^c, \alpha^*), \quad (16)$$

where $f_I^{-1}(P_e^c, \alpha^*)$ is the reverse function of (13) related to $I_{E,C}$.

Step 4. Implement the LP algorithm of Equation (12) with the above α^* and corresponding $I_{E,C}^{\max}$ to search degree distributions.

This method is applied with channel parameters $\sigma = 0.977$, $L = 400$, and $P_e^c = 0.0039$. From (14) and (15), we have $I_{E,C}^{ub} = 0.5013$ and $\alpha^{lb} = 7.7$; thus, we start with $\alpha^* = 8$ but the searched degree distribution has a large overhead as $\varepsilon = 2.65$. For different α^* , the corresponding $I_{E,C}^{\max}$ and the resulting overheads are both plotted in Figure 2. It is seen that the overhead decreases very quickly to a minimum and then increases slowly with the increase of α^* . The corresponding β (not listed) is monotonically increasing as α^* and so does the decoding complexity. Therefore, the best distribution

FIGURE 2: The required $I_{E,C}^{max}$ and resulting overheads of optimized codes for different α^* .

TABLE 1: Parameters of optimized SR codes with different outer code rates.

R_o	P_e^c	ϵ	R	β
0.90	0.0202	0.96414	0.45822	6.742
0.91	0.0182	0.97248	0.46135	6.684
0.92	0.0159	0.98256	0.46405	6.615
0.93	0.0138	0.99879	0.46528	7.009
0.94	0.0118	1.00996	0.46767	7.921
0.95	0.0097	1.01953	0.47041	8.337
0.96	0.0077	1.02608	0.47382	8.771
0.97	0.0057	1.03431	0.47682	9.185
0.98	0.0039	1.04023	0.48034	9.613
0.99	0.0020	1.05176	0.48251	10.459

could be chosen as the one with minimum overhead, and here is $\epsilon = 1.04$ with $\alpha^* = 10$. If changing to another channel parameter σ and P_e^c , we still observe the similar phenomenon; thus, the last two steps of the above method could be combined into one step: the algorithm stops if the resulting overhead starts to increase and then the previous one is taken as the optimized results.

The proposed method can also be extended to the case of variable outer code rate. Setting the outer codes as regular LDPC codes with a rate from 0.90 to 0.99, we present the critical BERs and the feature parameters of resulting codes in Table 1. In general, with the increase of outer code rate R_o , the critical BER decreases and then the overhead of inner codes ϵ increases. The overall rate $R = R_o/(1 + \epsilon)$ increases along with the rising of R_o , but the decoding complexity, indicated by β , also increases. In practice, one can choose the codes balancing the rate and decoding complexity based on system requirements.

5. Simulation Results

We choose a regular (4,200)-LDPC code with a rate of $R_o = 0.98$ as the outer code, and the distributions of the inner code optimized by the above method are shown in Table 2, the bottom line of which gives the corresponding overhead threshold. The code of $\Omega_{opt1}(x)$ is optimized under $\sigma = 0.977$, which is also the channel condition for code optimization in [15]. We denote the distribution of the corresponding code in [15] as $\Omega_{s1}(x)$. For comparison, we also denote the classic distribution optimized with $K = 65536$ in [2] by $\Omega_{n1}(x)$. The asymptotic performance of SR codes with these three distributions is shown in Figure 3.

It is seen that the optimized distribution has the best performance in overhead threshold as $\epsilon = 1.05$, which is smaller than $\epsilon = 1.36$ of $\Omega_{s1}(x)$. The corresponding SR code-rate thresholds of $\Omega_{opt1}(x)$ and $\Omega_{s1}(x)$ are around 0.478 and 0.415, respectively, equivalent to about 15% increase in code rate. Furthermore, the average degree of $\Omega_{opt1}(x)$ is $\beta = 9.613$, which is smaller than $\beta = 12.507$ of $\Omega_{s1}(x)$, equivalent to about 23% decrease in average degree. The code of $\Omega_{n1}(x)$ has the smallest average degree as $\beta = 5.87$, but the overhead threshold of $\epsilon = 1.77$ is the highest. The reason of the optimized code having advantages in both overhead and average degree is that the proposed method has found the optimal design parameters for LP, using which the searched result can achieve the critical BER as soon as possible with an appropriate average degree. We can see that the curves of $\Omega_{opt1}(x)$ descends sharply and “exactly” to the intersection point with the line of critical BER, while $\Omega_{s1}(x)$ have a wide transitional region and the start point of error floor is far below critical BER (would increase average degree).

To further verify the advantage of optimized code, we depict the simulated BER performance of SR codes with those distributions in Figure 4, where information length is

TABLE 2: Degree distributions of the optimized SR codes.

$\sigma = 0.977, \alpha = 10$ $\Omega_{\text{opt1}}(x)$		$\sigma = 0.945, \alpha = 9.6$ $\Omega_{\text{opt2}}(x)$		$\sigma = 0.5, \alpha = 3.5$ $\Omega_{\text{opt3}}(x)$	
Ω_6	0.84682	Ω_6	0.54882	Ω_{34}	0.58683
Ω_7	0.04504	Ω_7	0.35176	Ω_{35}	0.41317
Ω_{38}	0.00061	Ω_{43}	0.06956		
Ω_{39}	0.10753	Ω_{44}	0.02986		
$\varepsilon = 1.050$		$\varepsilon = 0.960$		$\varepsilon = 0.102$

FIGURE 3: Asymptotic performance of SR codes with $\sigma = 0.977$. The dotted lines are the BERs obtained after inner decoding, and the solid lines are those after outer decoding,

fixed to $K = 10000$. Due to the degree distributions of $\Omega_{n1}(x)$ which is originally designed for nonsystematic Raptor (NR) code, we also provide the simulation result of the NR code with $\Omega_{n1}(x)$. It is seen that the optimized SR code indeed outperforms the existing SR codes, which validates the asymptotic results and verifies the effectiveness of our design method. However, the performance of the NR code with $\Omega_{n1}(x)$ and the optimized SR code is very close to each other, the reason of which is analysed as follows. Under the channel with $\sigma = 0.977$, those two codes achieving a BER below 10^{-5} have a rate close to $1/2$, and from [10], we know that the performance of rate- $1/2$ NR code with $\Omega_{n1}(x)$ is actually approaching the channel capacity, i.e., the space for performance improvement is limited. In addition, the obvious performance difference is observed between the NR code and the SR code with the same degree distribution ($\Omega_{n1}(x)$).

The code of $\Omega_{\text{opt2}}(x)$ in Table 2 is optimized under $\sigma = 0.945$ which has an overhead of $\varepsilon = 0.96$, corresponding to

the overall rate $R = R_o/(1 + \varepsilon) = 1/2$ and the bit signal-to-noise ratio $E_b/N_0 = 0.491$ dB. In Figure 5, the SR code of $\Omega_{\text{opt2}}(x)$ is compared with the rate- $1/2$ SR code and NR code optimized in [10, 13], the distributions of which are denoted by $\Omega_{s2}(x)$ and $\Omega_{n2}(x)$, respectively. It can be seen that our optimized SR code outperforms those two codes, though the improvement is not very big when compared with the NR code. The code of $\Omega_{\text{opt3}}(x)$ in Table 2 is optimized under $\sigma = 0.5$, corresponding to the high symbol signal-to-noise ratio (SNR) $E_s/N_0 = 3.01$ dB. It is also compared with the SR code of $\Omega_{s2}(x)$ and the NR code of $\Omega_{n2}(x)$ in Figure 5. The optimized SR code achieves a BER of 10^{-5} when the overhead $\varepsilon = 0.15$, indicating that only a few LT-coded symbols of the inner code are needed for successful decoding. The overheads of the SR code with $\Omega_{s2}(x)$ and the NR code with $\Omega_{n2}(x)$ achieving a BER of 10^{-5} are around 0.36 and 0.30, respectively. This result implies that the optimized SR code may have better performance than the NR code under high SNR (or high rate).

FIGURE 4: Simulated BER performance of Raptor codes with $K = 10000$ and $\sigma = 0.977$. FIGURE 5: Simulated BER performance of Raptor codes with $K = 10000$ under different σ .

6. Conclusions

In this paper, we have investigated the asymptotic performance and distribution design of systematic rateless codes using the GA-MI method, which gives a more accurate result than the GA-mean method. Based on critical BER analysis, the average degree of VNs and the maximum mutual infor-

mation in linear program are analysed and located efficiently, using which we have obtained the codes with lower overhead and appropriate decoding complexity than existing codes. The influence of outer code rate on the overall code rate and decoding complexity is also discussed. In addition, the optimized SR codes are compared with NR codes under different SNR.

Data Availability

The numerical and simulation data used to support the findings of this study are included within the supplementary information file.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61501469) and supported by the China Postdoctoral Science Foundation (Grant No. 2015M572782).

Supplementary Materials

(Supplementary Materials)

References

- [1] M. Luby, "LT codes," in *Proceedings of 43rd Symposium on Foundations of Computer Science, ser. FOCS '02*, pp. 271–280, Washington, DC, USA, 2002.
- [2] A. Shokrollahi, "Raptor codes," *IEEE Transactions on Communications*, vol. 52, no. 6, pp. 2551–2567, 2006.
- [3] R. Palanki and J. S. Yedidia, "Rateless codes on noisy channels," in *International Symposium on Information Theory, 2004. ISIT 2004. Proceedings*, pp. 1–6, Chicago, IL, USA, June–July 2004.
- [4] B. Sivasubramanian and H. Leib, "Fixed-rate raptor codes over Rician fading channels," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3905–3911, 2008.
- [5] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2033–2051, 2006.
- [6] Z. Cheng, J. Castura, and Y. Mao, "On the design of raptor codes for binary-input Gaussian channels," *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3269–3277, 2009.
- [7] A. Venkiah, C. Poulliat, and D. Declercq, "Jointly decoded raptor codes: analysis and design for the BIAWGN channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 657970, 11 pages, 2009.
- [8] I. Hussain, X. Xiao, and L. K. Rasmussen, "Error floor analysis of LT codes over the additive white Gaussian noise channel," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, pp. 1–5, Houston, TX, USA, December 2011.
- [9] 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 Communications Letters*, vol. 18, no. 3, pp. 491–494, 2014.
- [10] A. Kharel and L. Cao, "Analysis and design of physical layer raptor codes," *IEEE Communications Letters*, vol. 22, no. 3, pp. 450–453, 2018.
- [11] X. Yuan and L. Ping, "On systematic LT codes," *IEEE Communications Letters*, vol. 12, no. 9, pp. 681–683, 2008.
- [12] T. D. Nguyen, L. L. Yang, and L. Hanzo, "Systematic Luby transform codes and their soft decoding," in *2007 IEEE Workshop on Signal Processing Systems*, pp. 67–72, Shanghai, China, October 2007.
- [13] A. Kharel and L. Cao, "Asymptotic analysis and optimization design of physical layer systematic rateless codes," in *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, Las Vegas, NV, USA, January 2018.
- [14] D. Deng, D. Xu, and S. Xu, "Optimisation design of systematic LT codes over AWGN multiple access channel," *IET Communications*, vol. 12, no. 11, pp. 1351–1358, 2018.
- [15] S. Xu and D. Xu, "Optimization design and asymptotic analysis of systematic Luby transform codes over BIAWGN channels," *IEEE Transactions on Communications*, vol. 64, no. 8, pp. 3160–3168, 2016.
- [16] M. Ardakani and F. R. Kschischang, "A more accurate one-dimensional analysis and design of irregular LDPC codes," *IEEE Transactions on Communications*, vol. 52, no. 12, pp. 2106–2114, 2004.
- [17] J. Dai, P. Zhao, F. Zhang, and X. Chen, "Analysis and design of systematic rateless codes in FH/BFSK system with interference," *Mobile Information Systems*, vol. 2020, Article ID 9049284, 8 pages, 2020.
- [18] F. Brännström, L. K. Rasmussen, and A. J. Grant, "Convergence analysis and optimal scheduling for multiple concatenated codes," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3354–3364, 2005.