Bounds on the Error Probability of Raptor Codes under Maximum Likelihood Decoding

Francisco Lázaro, *Member, IEEE*, Gianluigi Liva, *Senior Member, IEEE*,
Gerhard Bauch, *Fellow, IEEE*, Enrico Paolini, *Senior Member, IEEE*

Abstract

In this paper upper and lower bounds on the probability of decoding failure under maximum likelihood decoding are derived for different (nonbinary) Raptor code constructions. In particular four different constructions are considered; *(i)* the standard Raptor code construction, *(ii)* a multi-edge type construction, *(iii)* a construction where the Raptor code is nonbinary but the generator matrix of the LT code has only binary entries, *(iv)* a combination of (ii) and (iii). The latter construction resembles the one employed by RaptorQ codes, which at the time of writing this article represents the state of the art in fountain codes. The bounds are shown to be tight, and provide an important aid for the design of Raptor codes.

Index Terms

Erasure correction, fountain codes, inactivation decoding, LT codes, maximum likelihood decoding, Raptor codes.

Francisco Lázaro and Gianluigi Liva are with the Institute of Communications and Navigation of the German Aerospace Center (DLR), Muenchner Strasse 20, 82234 Wessling, Germany. Email:{Francisco.LazaroBlasco, Gianluigi.Liva}@dlr.de.

Gerhard Bauch is with the Institute for Telecommunication, Hamburg University of Technology, Hamburg, Germany. E-mail: Bauch@tuhh.de.

Enrico Paolini is with CNIT, DEI, University of Bologna, via Dell'Università 50, 47522 Cesena (FC), Italy. E-mail: e.paolini@unibo.it.

Corresponding Address: Francisco Lázaro, KN-SAN, DLR, Muenchner Strasse 20, 82234 Wessling, Germany. Tel: +49-8153 28-3211, Fax: +49-8153 28-2844, E-mail: Francisco.LazaroBlasco@dlr.de.

This work has been presented in part at IEEE Globecom, Washington DC, USA, December 2016 [1].

I. Introduction

Fountain codes [2] are a class of erasure codes which have the property of being rateless. Thus, they are potentially able to generate an endless amount of encoded (or output) symbols from k information (or input) symbols. This property makes them suitable for application in situations where the channel erasure rate is not a priori known. The first class of practical fountain codes, Luby Transform (LT) codes, was introduced in [3] together with an iterative decoding algorithm that achieves a good performance when the number of input symbols is large. In [3], [4] it was shown how, in order to achieve a low probability of decoding error, the encoding and iterative decoding cost[1] per output symbol is $O\left(\ln(k)\right)$.

Raptor codes were introduced in [4] and outperform LT codes in several aspects. They consist of a serial concatenation of an outer code $\mathcal{C}$ (or *precode*) with an inner LT code. On erasure channels, this construction allows relaxing the design of the LT code, requiring only the recovery of a fraction $1-\sigma$ of the input symbols, with σ small. This can be achieved with linear encoding and decoding complexity (under iterative decoding). The outer code is responsible for recovering the remaining fraction σ of input symbols. If the outer code $\mathcal{C}$ is linear-time encodable and decodable, then the Raptor code has linear encoding and (iterative) decoding complexity over erasure channels.

Most of the existing works on LT and Raptor codes consider iterative decoding and assume large input block lengths (k at least in the order of a few tens of thousands). However, in practice, smaller values of k are more commonly used. For example, for the binary Raptor codes standardized in [5] and [6] the supported values of k range from 4 to 8192. For these input block lengths, iterative decoding performance degrades considerably. In this regime, a different decoding algorithm may be adopted that is an efficient maximum likelihood (ML) decoder, in the form of inactivation decoding [7]–[11]. An inactivation decoder solves a system of equations in several stages. First a set of variables is declared to be *inactive*. Next a system of equations involving only the set of inactive variables needs to be solved, for example using Gaussian elimination. Finally, once the value of the inactive variables is known, all other variables (those which were not inactive) are recovered using iterative decoding (back substitution).

[1] In [4] the cost per output symbol is defined as the encoding/decoding complexity normalized by the number of output symbols. The complexity is defined as the number operations needed to carry out encoding/decoding.

Recently, some works have addressed the complexity of inactivation decoding for Raptor and LT codes [12]–[15]. The probability of decoding failure of LT and Raptor codes under ML decoding has also been subject of study in several works. In [16] upper and lower bounds on the symbol erasure rate were derived for LT codes and Raptor codes with outer codes in which the elements of the parity-check matrix are independent and identically distributed (i.i.d.) Bernoulli random variables. This work was elegantly extended in [17], [18], where upper and lower bounds on the error probability of LT codes under ML decoding were derived. Moreover, [18] introduced an approximation to the probability of error of Raptor codes under ML decoding, that was derived under the assumption that the number of erasures correctable by the outer code is small. Hence, the approximation holds when the rate of the outer code is sufficiently high. In [19] it was shown by means of simulations how the error probability of Raptor codes constructed on $\mathbb{F}_q$, the finite field of order q, is very close to that of linear random fountain codes. In [20] upper and lower bounds on the probability of decoding failure of Raptor codes were derived. The outer codes considered in [20] are binary linear random codes with a systematic encoder. Ensembles of Raptor codes with linear random outer codes were also studied in a fixed-rate setting in [21], [22]. In [23], q-ary Raptor codes are considered, but only for the case in which the outer code is a low-density generator matrix code. Although a number of works have studied the probability of decoding failure of Raptor codes, to the best of the authors' knowledge, up to now the results hold only for specific outer codes (see [16], [20]–[23]).

In this paper upper and lower bounds on the probability of decoding failure of different Raptor code constructions are derived. The upper bounds derived in this paper follow the footsteps of [17], [18], where bounds to the error probability of LT codes were derived. In contrast to other works in literature [16], [20]–[23], the bounds presented in this paper are general since they are valid for any outer code, requiring only the (joint) weight enumerator (or composition enumerator, a quantity to be defined later) of the outer code. Furthermore, simulation results are presented which show how the derived bounds are tight. In particular four different constructions are considered, namely:

i) a Raptor code construction over $\mathbb{F}_q$, where the outer code is built over $\mathbb{F}_q$ as well as the generator matrix of the LT code;

ii) a multi-edge type Raptor construction over $\mathbb{F}_q$, where intermediate symbols of two different types can be distinguished;

iii) a construction where the Raptor code is built over $\mathbb{F}_q$ but the generator matrix of the LT code has only entries belonging to $\{0, 1\} \subseteq \mathbb{F}_q$;

iv) a combination of (ii) and (iii).

The bounds are applicable for the two Raptor codes present in standards. In particular, the R10 Raptor code in its nonsystematic form [5] is an example of construction (i), since binary Raptor codes are simply a special case ($q = 2$). Furthermore, the RaptorQ code in its nonsystematic form [24] is an example of construction (iv). The RaptorQ code is, at the timing, the state of the art fountain code construction, and it is an IETF standard [24]. To the best of the authors' knowledge, this is the first work which analyzes the performance of the RaptorQ construction[2].

The upper bounds on the probability of decoding failure are derived for all the above four constructions and they all result from application of the union bound. As mentioned before, they generalize the results in literature to the case where the outer codes are chosen arbitrarily (with the caveat of having sufficient knowledge of the outer code distance properties). In the same general setting, two types of lower bounds are obtained. A first lower bound is a consequence of the degree-two Bonferroni inequality (as for the lower bounds introduced in [16]). A second, tighter lower bound is obtained by means of the Dawson-Sankoff inequality [25], which generalizes the Bonferroni inequality.[3] The bounds are shown to be remarkably tight at large overheads, and sufficiently tight at overheads approaching zero. Starting from the upper bound on the probability of decoding failure, an error exponent analysis of Raptor codes is presented, which allows characterizing the overhead regions for which an exponential decay (in the input block length) of the expected failure probability can be attained. Examples of the application of the proposed bounds to the design of Raptor codes are finally provided.

The paper is organized as follows. In Section II some preliminary definitions are given. Section III presents a number of results on joint compositions. Section IV addresses the different Raptor code constructions considered in this paper. Section V presents several theorems with upper and lower bounds on the probability of decoding failure for the different Raptor code constructions. Proofs of the bounds are given in Section VI. Section VII introduces the error exponent analysis. Numerical results comparing the bounds with Monte Carlo simulations are illustrated in Section VIII, while code design examples are discussed in Section IX. Section X

[2] In [23] a q-ary Raptor code construction is analyzed, but it does not consider all the peculiarities of the RaptorQ code.

[3] Note that the Dawson-Sankoff inequality was used in [26] to lower bound the expected error probability of regular low-density parity-check (LDPC) code ensembles over the binary erasure channel (BEC).

presents the conclusions of our work.

II. Preliminaries

A. Vector and Matrix Notation

We use boldface letters to denote vectors and matrices. Vectors are conventionally assumed as row vectors with indices starting from 0; matrix row and column indices also start from 0. For any integer matrix $\mathbf{A}$ we denote by $|\mathbf{A}|$ the sum of all matrix elements. We use the same notation for integer vectors, i.e., $|\mathbf{a}|$ represents the sum of all elements of vector $\mathbf{a}$. We also denote by $\mathbf{1}(\mathbf{A})$ the matrix obtained from $\mathbf{A}$ by turning to 1 all its nonzero elements. The transpose of any matrix $\mathbf{A}$ is denoted by $\mathbf{A}^{\mathsf{T}}$.

We say that a zero-one square matrix $\mathbf{A}$ is a circulant permutation matrix when: (i) it is a permutation matrix; (ii) each row of $\mathbf{A}$ is obtained from the previous row by the right cyclic shift of one position. We say that a zero-one square matrix $\mathbf{A}$ is an *incomplete* circulant permutation matrix when: (i) it is nonzero; (ii) it can be obtained from a circulant permutation matrix by turning to 0 some 1 elements.

For a nonnegative integer vector $\mathbf{a} = (a_0, a_1, \dots, a_{n-1})$ such that $|\mathbf{a}| = h$ we denote the multinomial coefficient $\binom{h}{a_0, a_1, \dots, a_{n-1}}$ by $\binom{h}{\mathbf{a}}$. With a slight abuse of notation, for an $m \times n$ nonnegative integer matrix $\mathbf{A} = [a_{s,t}]$ such that $|\mathbf{A}| = h$ we write $\binom{h}{\mathbf{A}}$ as a compact notation for $\binom{h}{a_{0,0}, \dots, a_{0,n-1}, \dots, a_{m-1,0}, \dots, a_{m-1,n-1}}$.

B. Bonferroni-Type Inequalities

Let A_1, …, A_n be events in a probability space and

$$S_k = \sum_{1 \le i_1 < \cdots < i_k \le n} \Pr\{A_{i_1} \cap \cdots \cap A_{i_k}\}.$$

The general Bonferroni inequality states that, for any $1 \le t \le n$, we have [27]

$$(-1)^t \Pr\{A_1 \cup \cdots \cup A_n\} \ge (-1)^t \sum_{i=1}^{t} (-1)^{i-1} S_i. \tag{1}$$

Inequality (1) holds with equality for $t = n$ (inclusion-exclusion identity). Notable special cases are obtained for $t = 1$ and $t = 2$. Specifically, for $t = 1$ it reduces to the union upper bound

$$\Pr\{A_1 \cup \cdots \cup A_n\} \le S_1 = \sum_{i=1}^{n} \Pr\{A_i\} \tag{2}$$

while for $t = 2$ it yields the degree-two Bonferroni lower bound

$$\Pr\{A_1 \cup \cdots \cup A_n\} \geq S_1 - S_2 = \sum_{i=1}^{n} \Pr\{A_i\} - \sum_{1 \leq i < j \leq n} \Pr\{A_i \cap A_j\}. \tag{3}$$

A tighter version of (3) was developed in [25], where it was shown that, for any $r \in \{1, \ldots, n\}$,

$$\Pr\{A_1 \cup \cdots \cup A_n\} \geq \frac{2}{r+1} S_1 - \frac{2}{r(r+1)} S_2. \tag{4}$$

Moreover, maximization with respect to r yields

$$\Pr\{A_1 \cup \cdots \cup A_n\} \geq \frac{\theta S_1^2}{(2-\theta)S_1 + 2S_2} + \frac{(1-\theta)S_1^2}{(1-\theta)S_1 + 2S_2} \tag{5}$$

where $\theta = 2S_2/S_1 - \lfloor 2S_2/S_1 \rfloor$. Indeed, it was proved in [28] that (5) is the sharpest possible lower bound for $\Pr\{A_1 \cup \cdots \cup A_n\}$ based on a linear combination of S_1 and S_2. As such, it is tighter than $S_1 - S_2$. Hereafter, (5) will be referred to as Dawson-Sankoff lower bound.

C. Weight and Composition Enumerators

For any linear block code $\mathcal{C}$ constructed over $\mathbb{F}_q$ and any codeword $\mathbf{v} \in \mathcal{C}$, we let $w(\mathbf{v})$ be the Hamming weight (often referred to simply as the weight) of $\mathbf{v}$. Letting h be the codeword length, we denote the weight enumerator of $\mathcal{C}$ as $A = \{A_0, A_1 \ldots A_h\}$, where A_i denotes the multiplicity of codewords of weight i. Similarly, given an ensemble $\mathscr{C}$ of linear block codes, all with the same block length h, along with a probability distribution on the codes in the ensemble, we denote the expected weight enumerator of a random code in $\mathscr{C}$ as $\mathsf{A} = \{\mathsf{A}_0, \mathsf{A}_1 \ldots \mathsf{A}_h\}$, where A_l denotes the expected multiplicity of codewords of weight l.

Next, consider a linear block code $\mathcal{C} \subset \mathbb{F}_q^h$, whose codeword symbols are partitioned into two different types, namely, type A and type B. Let h_A and h_B be the number of codeword symbols of types A and B, respectively, such that $h_A + h_B = h$. A generic codeword after reordering can be expressed as $\mathbf{v} = (\mathbf{v}_A, \mathbf{v}_B)$, where $\mathbf{v}_A$ and $\mathbf{v}_B$ denote the vectors of encoded symbols of type A and type B respectively. In this context the bivariate weight enumerator polynomial of the code is defined as

$$A(x, z) = \sum_{l=0}^{h_A} \sum_{t=0}^{h_B} A_{l,t} \, x^l z^t \tag{6}$$

where $A_{l,t}$ denotes the multiplicity of codewords with $w(\mathbf{v}_A) = l$ and $w(\mathbf{v}_B) = t$. Similarly, given an ensemble $\mathscr{C}$ of block codes with block length h and with two types of codeword symbols as defined above, along with a probability distribution on the codes in the ensemble, we define its expected bivariate weight enumerator polynomial as

$$\mathsf{A}(x,z)=\sum_{l=0}^{h_A}\sum_{t=0}^{h_B}\mathsf{A}_{l,t}\,x^l z^t$$

where $\mathsf{A}_{l,t}$ denotes the expected multiplicity of codewords with $w(\mathbf{v}_A)=l$ and $w(\mathbf{v}_B)=t$.

Given a vector $\mathbf{r}=(r_0,r_1,\ldots,r_{h-1})\in\mathbb{F}_q^h$, we define its composition $\varsigma(\mathbf{r})$ as

$$\varsigma(\mathbf{r})=(\varsigma_0(\mathbf{r}),\varsigma_1(\mathbf{r}),\ldots,\varsigma_{q-1}(\mathbf{r}))$$

where

$$\varsigma_i(\mathbf{r})=\left|\left\{r_j : r_j=\alpha^{i-1}\right\}\right|,\ \text{for } j\in\{0,\ldots,h-1\} \text{ and } i\in\{1,2,\ldots,q-1\}$$

being α the residue class of the polynomial x, and

$$\varsigma_0(\mathbf{r})=\left|\{r_j : r_j=0\}\right|\ \text{ for } j\in\{1,\ldots,h\}.$$

That is, $\varsigma_i(\mathbf{r})$, $i\in\{1,2,\ldots,q-1\}$, is the number of elements in $\mathbf{r}$ that take value α^{i-1} whereas $\varsigma_0(\mathbf{r})$ is the number of null elements in $\mathbf{r}$. Given a linear block code $\mathcal{C}$, we define its composition enumerator, $\mathcal{Q}_{\mathbf{f}}$, as the number of codewords $\mathbf{v}\in\mathcal{C}$ with composition $\varsigma(\mathbf{v})=\mathbf{f}$. Similarly, for a code ensemble we define its expected composition enumerator $\mathsf{Q}_{\mathbf{f}}$ as the expected multiplicity of codewords with composition $\mathbf{f}$.

Consider also a linear block code $\mathcal{C}$ of length h, with two types of codeword symbols as defined above. We define the bivariate composition enumerator $\mathcal{Q}_{\mathbf{f}_A,\mathbf{f}_B}$ of a code $\mathcal{C}$ as the number of codewords $\mathbf{v}=(\mathbf{v}_A,\mathbf{v}_B)$ in $\mathcal{C}$ for which $\mathbf{v}_A$ has composition $\mathbf{f}_A$ and $\mathbf{v}_B$ has composition $\mathbf{f}_B$. This definition can be easily extended to code ensembles. In particular, we define the expected bivariate composition enumerator $\mathsf{Q}_{\mathbf{f}_A,\mathbf{f}_B}$ of a random code in the ensemble as the expected multiplicity of codewords $\mathbf{v}=(\mathbf{v}_A,\mathbf{v}_B)$ for which $\mathbf{v}_A$ has composition $\mathbf{f}_A$ and $\mathbf{v}_B$ has composition $\mathbf{f}_B$.

Given the composition $\mathbf{f}$ of a vector $\mathbf{r}\in\mathbb{F}_q^h$, $\mathbf{f}=\varsigma(\mathbf{r})$, as defined above, we define $B(\mathbf{f})$ as an indicator function that takes value 1 only if $\sum_{i=1}^h r_i=0$, i.e.,

$$B(\mathbf{f})=\begin{cases}1, & \text{if } \sum_{i=1}^{q-1}\sum_{s=1}^{f_i}\alpha^{i-1}=0\\ 0, & \text{otherwise.}\end{cases}$$

D. Joint Weight and Joint Composition Enumerators

Given two vectors $\mathbf{r}_1 \in \mathbb{F}_q^h$ and $\mathbf{r}_2 \in \mathbb{F}_q^h$, we define the joint weight of $\mathbf{r}_1$ and $\mathbf{r}_2$, denoting it as $\boldsymbol{\tau} = \tau(\mathbf{r}_1, \mathbf{r}_2)$, as the vector $(\tau_0, \tau_1, \tau_2, \tau_3)$ such that:

- There are τ_0 positions in which both $\mathbf{r}_1$ and $\mathbf{r}_2$ are zero;
- There are τ_1 positions in which $\mathbf{r}_1$ is zero and $\mathbf{r}_2$ is nonzero;
- There are τ_2 positions in which $\mathbf{r}_1$ is nonzero and $\mathbf{r}_2$ is zero;
- There are τ_3 positions in which both $\mathbf{r}_1$ and $\mathbf{r}_2$ are nonzero.

The elements of $\boldsymbol{\tau} = \tau(\mathbf{r}_1, \mathbf{r}_2)$ are nonnegative integers and $|\boldsymbol{\tau}| = h$.

Given two vectors $\mathbf{r}_1 \in \mathbb{F}_q^h$ and $\mathbf{r}_2 \in \mathbb{F}_q^h$, we define the joint composition of $\mathbf{r}_1$ and $\mathbf{r}_2$, denoting it as $\boldsymbol{\kappa} = \kappa(\mathbf{r}_1, \mathbf{r}_2)$, as the $q \times q$ matrix $[\kappa_{s,t}]$, $(s,t) \in \{0, \ldots, q-1\} \times \{0, \ldots, q-1\}$, such that:

- There are $\kappa_{0,0}$ positions in which both $\mathbf{r}_1$ and $\mathbf{r}_2$ are zero;
- There are $\kappa_{0,t}$ positions, $t \neq 0$, in which $\mathbf{r}_1$ is zero and $\mathbf{r}_2$ is equal to α^{t-1};
- There are $\kappa_{s,0}$ positions, $s \neq 0$, in which $\mathbf{r}_1$ is equal to α^{s-1} and $\mathbf{r}_2$ is zero;
- There are $\kappa_{s,t}$ positions, $s \neq 0$, $t \neq 0$, in which $\mathbf{r}_1$ is equal to α^{s-1} and $\mathbf{r}_2$ is equal to α^{t-1}.

The elements of $\kappa(\mathbf{r}_1, \mathbf{r}_2)$ are nonnegative integers and $|\boldsymbol{\kappa}| = h$. We write

$$\boldsymbol{\kappa} = \begin{bmatrix} \kappa_{0,0} & \boldsymbol{\kappa}_1 \\ \boldsymbol{\kappa}_2 & \boldsymbol{\kappa}_3 \end{bmatrix} \tag{7}$$

where $\boldsymbol{\kappa}_1$ is the $1 \times (q-1)$ matrix $[\kappa_{0,1}, \ldots, \kappa_{0,q-1}]$, $\boldsymbol{\kappa}_2$ is the $(q-1) \times 1$ matrix $[\kappa_{1,0}, \ldots, \kappa_{q-1,0}]^{\mathsf{T}}$, and $\boldsymbol{\kappa}_3$ is the $(q-1) \times (q-1)$ matrix $[\kappa_{s,t}]$, $(s,t) \in \{1, \ldots, q-1\} \times \{1, \ldots, q-1\}$.

There is a simple relationship between the joint weight $\boldsymbol{\tau} = \tau(\mathbf{r}_1, \mathbf{r}_2)$ of two vectors and their joint composition $\boldsymbol{\kappa} = \kappa(\mathbf{r}_1, \mathbf{r}_2)$. In particular, we have $\tau_0 = \kappa_{0,0}$, $\tau_1 = |\boldsymbol{\kappa}_1|$, $\tau_2 = |\boldsymbol{\kappa}_2|$, and $\tau_3 = |\boldsymbol{\kappa}_3|$. We write $\boldsymbol{\tau} = \tau(\boldsymbol{\kappa})$ to indicate the joint weight $\boldsymbol{\tau}$ associated with the joint composition $\boldsymbol{\kappa}$. There also is a simple relationship between the joint composition $\boldsymbol{\kappa} = \kappa(\mathbf{r}_1, \mathbf{r}_2)$ of two vectors and the composition of each of them. Specifically, denoting the composition of $\mathbf{r}_1$, $\varsigma(\mathbf{r}_1)$, by $\boldsymbol{\gamma}_1(\boldsymbol{\kappa})$ and the composition of $\mathbf{r}_2$, $\varsigma(\mathbf{r}_2)$, by $\boldsymbol{\gamma}_2(\boldsymbol{\kappa})$, we have

$$\boldsymbol{\gamma}_1(\boldsymbol{\kappa}) = \Big(\sum_{t=0}^{q-1} \kappa_{0,t}, \ldots, \sum_{t=0}^{q-1} \kappa_{q-1,t} \Big) \tag{8}$$

$$\boldsymbol{\gamma}_2(\boldsymbol{\kappa}) = \Big(\sum_{s=0}^{q-1} \kappa_{s,0}, \ldots, \sum_{s=0}^{q-1} \kappa_{s,q-1} \Big). \tag{9}$$

Given two linear block codes $\mathcal{C}_1 \subset \mathbb{F}_q^h$ of dimension k_1 and $\mathcal{C}_2 \subset \mathbb{F}_q^h$ of dimension k_2, we define their joint weight enumerator, denoting it by $J_{\boldsymbol{\tau}}$, as the number of codeword pairs $(\mathbf{v}, \mathbf{z}) \in \mathcal{C}_1 \times \mathcal{C}_2$

such that $\tau(\mathbf{v},\mathbf{z}) = \boldsymbol{\tau}$. We also define their joint composition enumerator, denoting it by $\mathcal{S}_{\boldsymbol{\kappa}}$, as the number of codeword pairs $(\mathbf{v},\mathbf{z}) \in \mathcal{C}_1 \times \mathcal{C}_2$, such that $\kappa(\mathbf{v},\mathbf{z}) = \boldsymbol{\kappa}$. If $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}$, then $J_{\boldsymbol{\tau}}$ and $\mathcal{S}_{\boldsymbol{\kappa}}$ are called the biweight and the bicomposition enumerator of $\mathcal{C}$, respectively. For an ensemble $\mathscr{C}$ of linear block codes, all with the same block length, we denote by $\mathsf{J}_{\boldsymbol{\tau}}$ and $\mathsf{S}_{\boldsymbol{\kappa}}$ the expected biweight and bicomposition enumerators, respectively, of a random code in $\mathscr{C}$.[4]

Remark 1. *For $q = 2$, if $\boldsymbol{\tau} = \tau(\boldsymbol{\kappa})$, then $\boldsymbol{\tau} = (\kappa_{0,0}, \kappa_{0,1}, \kappa_{1,0}, \kappa_{1,1})$. Thus, in the binary case there exists a bijection between joint weights and joint compositions so that the two concepts become equivalent and can be used interchangeably. With this bijection in mind we can also write $\mathcal{S}_{\boldsymbol{\kappa}} = J_{\boldsymbol{\tau}}$. This is not the case in the nonbinary case.*

E. Weight Spectral Shape of Code Ensemble Sequences

A code ensemble sequence $\{\mathscr{C}_k\}$ is a sequence of code ensembles, where $\mathscr{C}_k$ is an ensemble of dimension-k codes with block length $h = k/R$ defined over $\mathbb{F}_q$, being R a constant, i.e., not dependent on k. The weight spectral shape of the ensemble sequence $\{\mathscr{C}_k\}$ is given by

$$G(\omega) = \lim_{h\to\infty} \frac{1}{h} \log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h \rfloor}$$

where $\mathsf{A}^{(hR)}$ is the expected weight enumerator of the code ensemble $\mathscr{C}_{hR}$. In the definition above, ω can be regarded as the normalized Hamming weight.

We recall next the definition of uniform convergence, which will become essential for the results derived in Section VII. A sequence f_h of real-valued functions on $D \subseteq \mathbb{R}$ converges uniformly to the function $f : D \mapsto \mathbb{R}$ on $D_0 \subseteq D$ if for any $\varepsilon > 0$ there exists $h_0(\varepsilon)$ such that, for all $h \geq h_0(\varepsilon)$, $|f_h(x) - f(x)| < \varepsilon$ for all $x \in D_0$. We write $f_h \xrightarrow{\mathsf{u}} f$ to indicate that f_h converges to f uniformly.

F. Further Useful Definitions and Results

For a positive integer n and a prime or prime power q, we denote by $\mathcal{K}_i^{n,q}(x)$ the Krawtchouk polynomial of degree i with parameters n and q, which is defined as [29]

$$\mathcal{K}_i^{n,q}(x) = \sum_{j=0}^{i} (-1)^j \binom{x}{j} \binom{n-x}{i-j} (q-1)^{i-j}.$$

[4] The concept of joint weight and joint weight enumerator was introduced in [30], where examples of biweight numerators for some classical codes were obtained.

Moreover, we recall Chu-Vandermonde identity, stating that

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}.$$

III. Results on Joint Weights and Joint Compositions

This section presents a number of results on joint compositions. These results will be useful to develop a lower bound on the error probability of a class of Raptor codes.

Lemma 1. *Let* $\mathbf{r}_1 \in \mathbb{F}_q^h \setminus \{\mathbf{0}\}$ *and* $\mathbf{r}_2 \in \mathbb{F}_q^h \setminus \{\mathbf{0}\}$. *We have*

$$\kappa(\mathbf{r}_1, \mathbf{r}_2) = \begin{bmatrix} \kappa_{0,0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\kappa}_3 \end{bmatrix}$$

in which $\mathbf{1}(\boldsymbol{\kappa}_3)$ *is a (possibly incomplete) circular permutation matrix, if and only if* $\mathbf{r}_1 = \beta \mathbf{r}_2$ *for some* $\beta \in \mathbb{F}_q \setminus \{0\}$.

Proof: Let $\mathbf{r}_1 = \beta \mathbf{r}_2$ for some $\beta \in \mathbb{F}_q \setminus \{0\}$ ($\mathbf{r}_1$ and $\mathbf{r}_2$ are linearly dependent). With reference to (7), since $\mathbf{r}_1$ and $\mathbf{r}_2$ have the same support, both $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ must be null. Let the proportionality factor β be equal to α^s for some $s \in \{0, \dots, q-2\}$. Every element of $\mathbf{r}_1$ equal to α^i corresponds to an element $\alpha^{(i+s) \bmod (q-1)}$ in $\mathbf{r}_2$, making $\kappa_{1+i, 1+(i+s) \bmod (q-1)} > 0$; any other element of $\boldsymbol{\kappa}$ in row of index $1+i$ must be zero. This suffices to conclude that $\mathbf{1}(\boldsymbol{\kappa}_3)$ is a circulant permutation matrix if all elements of $\mathbb{F}_q \setminus \{0\}$ appear in $\mathbf{r}_1$. It is an incomplete circulant permutation matrix otherwise. Conversely, let $\boldsymbol{\kappa}_1 = \boldsymbol{\kappa}_2 = \mathbf{0}$ and $\mathbf{1}(\boldsymbol{\kappa}_3)$ be a (possibly incomplete) circulant permutation matrix. The vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ must have the same support. Moreover, there must exist $s \in \{0, \dots, q-2\}$ such that every nonzero element of $\boldsymbol{\kappa}$, apart from $\kappa_{0,0}$, is in the form $\kappa_{1+i, 1+(i+s) \bmod (q-1)}$ for some $i \in \{1, \dots, q-1\}$. But then $\mathbf{r}_1 = \alpha^s \mathbf{r}_2$. ■

Let $\mathcal{C} \subset \mathbb{F}_q^h$ be a linear block code of dimension k. We partition the codebook of $\mathcal{C}$ into $M_{q,k} = (q^k - 1)/(q-1) + 1$ parts $\mathcal{P}_a$, $a = \{0, 1, \dots, M_{q,k} - 1\}$, as follows. Part $\mathcal{P}_0$ only contains the null codeword, while any other part contains $q-1$ codewords having the same support and being linearly dependent. Moreover, we index the codewords in $\mathcal{C}$ from 0 to $q^k - 1$, as follows. The index 0 is reserved to the null codeword; the indices from $(a-1)(q-1)+1$ to $a(q-1)$ are reserved to the codewords in $\mathcal{P}_a$, $a \in \{1, \dots, M_{q,k} - 1\}$. For every $a \in \{1, \dots, M_{q,k} - 1\}$ we take one representative in $\mathcal{P}_a$, denoting it by $\tilde{\mathbf{v}}_a$. In particular, we choose as $\tilde{\mathbf{v}}_a$ the codeword

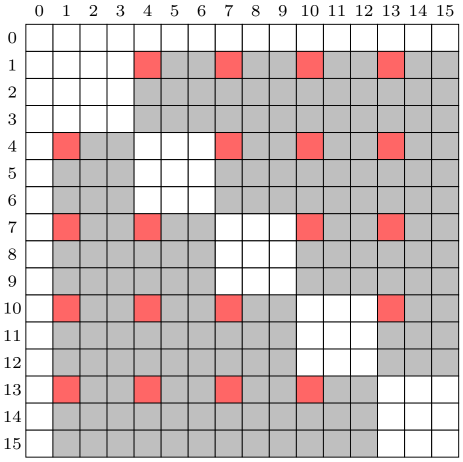

Fig. 1. Graphical interpretation of the set $\mathscr{D}_{q,k}$ for $q = 4$ and $k = 2$.

in $\mathcal{P}_a$ having the smallest index. Letting $\{\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{q^k-1}\}$ be the codebook of $\mathcal{C}$, with the above-mentioned indexing convention we have $\tilde{\mathbf{v}}_a = \mathbf{v}_{(a-1)(q-1)+1}$.

We define the set $\mathscr{D}_{q,k}$ as

$$\mathscr{D}_{q,k} = \{(s,t) \in \{1, q^k - 1\} \times \{1, q^k - 1\} : \lfloor \frac{s-1}{q-1} \rfloor \neq \lfloor \frac{t-1}{q-1} \rfloor \}.$$

Moreover, we define the set $\tilde{\mathscr{D}}_{q,k} \subseteq \mathscr{D}_{q,k}$ as

$$\tilde{\mathscr{D}}_{q,k} = \{(s,t) \in \mathscr{D}_{q,k} : s = (a-1)(q-1)+1; t = (b-1)(q-1)+1;$$
$$a, b \in \{1, \dots, M_{q,k} - 1\}; a \neq b\}\,.$$

The set $\mathscr{D}_{q,k}$ is the set of codeword index pairs (s,t) such that: (i) $\mathbf{v}_s$ and $\mathbf{v}_t$ are both nonzero; (ii) $\mathbf{v}_s$ and $\mathbf{v}_t$ are not linearly dependent. Its cardinality is $(q-1)^2(M_{q,k}-2)(M_{q,k}-1)$. The set $\tilde{\mathscr{D}}_{q,k}$ is a subset of $\mathscr{D}_{q,k}$. It includes all codeword index pairs (s,t) such that $\mathbf{v}_s$ is the representative of part $\mathcal{P}_{(s-1)/(q-1)+1}$, $\mathbf{v}_t$ is the representative of part $\mathcal{P}_{(t-1)/(q-1)+1}$, and $\mathbf{v}_s \neq \mathbf{v}_t$. Its cardinality is $(M_{q,k}-2)(M_{q,k}-1)$.

Example 1. *Let $q = 4$ and $k = 2$. A graphical interpretation of the set $\mathscr{D}_{4,2}$ is provided in Fig. 1. The codebook is partitioned into the $M_{4,2} = 6$ parts $\mathcal{P}_0 = \{\mathbf{0}\}$, $\mathcal{P}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, $\mathcal{P}_2 = \{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}$, $\mathcal{P}_3 = \{\mathbf{v}_7, \mathbf{v}_8, \mathbf{v}_9\}$, $\mathcal{P}_4 = \{\mathbf{v}_{10}, \mathbf{v}_{11}, \mathbf{v}_{12}\}$, $\mathcal{P}_5 = \{\mathbf{v}_{13}, \mathbf{v}_{14}, \mathbf{v}_{15}\}$, where all codewords in the same part are linearly dependent. The set $\mathscr{D}_{4,2}$ is represented by the union of all grey and red cells of the "chessboard", while the set $\tilde{\mathscr{D}}_{4,2}$ is represented only by the red* *cells. White cells, the ones not belonging to $\mathscr{D}_{4,2}$, correspond either to pairs of codewords of which at least one is null or to pairs of linearly dependent codewords.*

We define $\mathscr{K}_{q,h}$ as the set of all joint compositions $\boldsymbol{\kappa}$ such that $|\boldsymbol{\kappa}| = h$ and such that any of the following two conditions holds: (1) at least two matrices out of $\boldsymbol{\kappa}_1$, $\boldsymbol{\kappa}_2$, $\boldsymbol{\kappa}_3$ are nonzero; (2) $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ are null matrices, $\boldsymbol{\kappa}_3$ is nonzero, $\mathbf{1}(\boldsymbol{\kappa}_3)$ is neither a complete nor an incomplete circulant permutation matrix.

Lemma 2. *For any linear block code $\mathcal{C} \subset \mathbb{F}_q^h$ of dimension k and any pair $(\mathrm{v}_s, \mathrm{v}_t) \in \mathcal{C} \times \mathcal{C}$, we have $\kappa(\mathrm{v}_s, \mathrm{v}_t) \in \mathscr{K}_{q,h}$ if and only if $(s,t) \in \mathscr{D}_{q,k}$.*

Proof: Let $\kappa(\mathrm{v}_s, \mathrm{v}_t) \in \mathscr{K}_{q,h}$. If at least two matrices out of $\boldsymbol{\kappa}_1$, $\boldsymbol{\kappa}_2$, and $\boldsymbol{\kappa}_3$ are nonzero, then v_s and v_t are both nonzero and have different supports (so they cannot be linearly dependent). Thus we must have $(s,t) \in \mathscr{D}_{q,k}$. If $\boldsymbol{\kappa}_1 = \boldsymbol{\kappa}_2 = \mathbf{0}$, $\boldsymbol{\kappa}_3 \neq \mathbf{0}$, and $\mathbf{1}(\boldsymbol{\kappa}_3)$ is neither a circulant permutation matrix nor an incomplete one, then v_s and v_t have the same support but are not linearly dependent (Lemma 1). Thus we must have $(s,t) \in \mathscr{D}_{q,k}$ again. Conversely, let $(s,t) \in \mathscr{D}_{q,k}$, meaning that v_s and v_t are both nonzero and they are not linearly dependent. If v_s and v_t have different supports then at least two matrices out of $\boldsymbol{\kappa}_1$, $\boldsymbol{\kappa}_2$, and $\boldsymbol{\kappa}_3$ must nonzero, so $\kappa(\mathrm{v}_s, \mathrm{v}_t) \in \mathscr{K}_{q,h}$. If v_s and v_t have the same support, since they are not linearly dependent, by Lemma 1 $\boldsymbol{\kappa}_3$ can be neither a circulant permutation matrix, nor an incomplete one. Hence $\kappa(\mathrm{v}_s, \mathrm{v}_t) \in \mathscr{K}_{q,h}$ again. ■

A. Binary codes

In Remark 1 we pointed out that over $\mathbb{F}_2$ the concepts of joint composition and joint weight become equivalent. Thus, in the binary case the quantities and results so far introduced in this section can be reformulated in terms of joint weight. Note at first that when $q = 2$ the two sets $\mathscr{D}_{2,k}$ and $\tilde{\mathscr{D}}_{2,k}$ coincide and that $\mathscr{D}_{2,k}$ can be simply defined as

$$\mathscr{D}_{2,k} = \{(s,t) \in \{1, 2^k - 1\} \times \{1, 2^k - 1\} : s \neq t\}.$$

This is the set of all codeword index pairs (s,t) such that $\mathrm{v}_s \neq \mathbf{0}$, $\mathrm{v}_t \neq \mathbf{0}$, and $\mathrm{v}_s \neq \mathrm{v}_t$.

For $q = 2$, $\mathscr{K}_{2,h}$ may be simply defined as the set of all joint compositions $\boldsymbol{\kappa} = [\kappa_{s,t}]$, $s,t \in \{0,1\}$, such that $|\boldsymbol{\kappa}| = h$ and such that at least two parameters out of $\kappa_{0,1}$, $\kappa_{1,0}$, $\kappa_{1,1}$ are positive. Owing to the above-recalled equivalence between joint weights and joint compositions, we introduce the set $\mathscr{T}_{2,h}$ as the equivalent of $\mathscr{K}_{2,h}$ for joint weights. We define $\mathscr{T}_{2,h}$ as the set of all joint weights $\boldsymbol{\tau} = (\tau_0, \tau_1, \tau_2, \tau_3)$ such that $|\boldsymbol{\tau}| = h$ and such that at least two parameters out of τ_1, τ_2, τ_3 are positive. The following result is an immediate corollary of Lemma 2 for $q = 2$.

Lemma 3. *For any binary linear block code $\mathcal{C} \subset \mathbb{F}_2^h$ of dimension k and any pair $(\mathbf{v}_s, \mathbf{v}_t) \in \mathcal{C} \times \mathcal{C}$, we have $\tau(\mathbf{v}_s, \mathbf{v}_t) \in \mathscr{T}_{2,h}$ if and only if $(s, t) \in \mathscr{D}_{2,k}$.*

IV. Raptor Codes

A. Encoding and Decoding

We consider four different Raptor code constructions, all of them over $\mathbb{F}_q$, with $q \geq 2$, being q a prime or prime power. Fig. 2 shows a block diagram of Raptor encoding. In particular we consider an outer linear block code $\mathcal{C}$ whose length and dimension are denoted by h and k, respectively. We denote the k input (or source) symbols of the Raptor code as $\mathbf{u} = (u_0, u_1, \ldots, u_{k-1})$. Out of the k input symbols, the outer code generates a vector of h intermediate symbols $\mathbf{v} = (v_0, v_1, \ldots, v_{h-1})$. The rate of the outer code is hence $R = k/h$. Denoting by $\mathbf{G}_{\text{o}}$ the generator matrix of the outer code, of dimension $(k \times h)$, the intermediate symbol vector can be expressed as

$$\mathbf{v} = \mathbf{u}\mathbf{G}_{\text{o}}.$$

The intermediate symbols serve as input to an LT encoder, which generates the output symbols $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$, where n can grow unbounded. For any n, we have

$$\mathbf{c} = \mathbf{v}\mathbf{G}_{\text{LT}} = \mathbf{u}\mathbf{G}_{\text{o}}\mathbf{G}_{\text{LT}} \tag{10}$$

where $\mathbf{G}_{\text{LT}}$ is an $(h \times n)$ matrix. The different constructions addressed in this paper differ in how matrix $\mathbf{G}_{\text{LT}}$ is built, as we will explain later in this section.

The output symbols are transmitted over a q-ary erasure channel (q-EC). At its output each transmitted symbol is either correctly received or erased.[5] We denote by m the number of output symbols collected by the receiver, and we express it as $m = k + \delta$, where δ is the absolute receiver overhead. Let us denote by $\mathbf{y} = (y_0, y_1, \ldots, y_{m-1})$ the vector of m received output

[5] We remark that, due to the fact that LT output symbols are generated independently of each other, the results developed in this paper remain valid regardless the statistic of the erasures introduced by the channel.

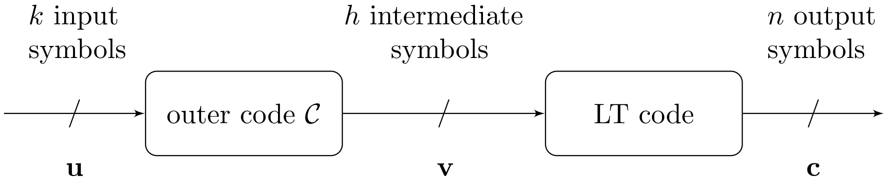

Fig. 2. Block diagram of Raptor encoding.

symbols. Denoting by $\mathcal{I} = \{i_0, i_1, \ldots, i_{m-1}\}$ the set of indices corresponding to the m non-erased symbols, we have $y_j = c_{i_j}$. An ML decoder proceeds by solving the linear system of equations

$$\mathbf{y} = \mathbf{u}\tilde{\mathbf{G}} \tag{11}$$

where

$$\tilde{\mathbf{G}} = \mathbf{G}_{\mathrm{o}}\tilde{\mathbf{G}}_{\mathrm{LT}}$$

and where $\tilde{\mathbf{G}}_{\mathrm{LT}}$ is the submatrix of $\mathbf{G}_{\mathrm{LT}}$ formed by the m columns with indices in $\mathcal{I}$.

B. Raptor Code Constructions

The first construction considered in this paper is referred to as *Raptor code over* $\mathbb{F}_q$. In this construction each column of $\mathbf{G}_{\mathrm{LT}}$ is generated by first randomly drawing an output degree d, according to a probability distribution $\Omega = (\Omega_1, \Omega_2, \ldots, \Omega_{d_{\max}})$, and then by drawing d different indices uniformly at random between 1 and h. The distribution Ω is usually referred to as output degree distribution and its generating function is

$$\Omega(x) = \sum_{d=1}^{d_{\max}} \Omega_d x^d .$$

Finally, the elements of the column in the row positions corresponding to these indices are drawn independently and uniformly at random from $\mathbb{F}_q \backslash \{0\}$, while all other elements of the column are set to zero.

The second considered construction is referred to as *multi-edge type Raptor code*. This construction is characterized by having two different types of intermediate symbols, namely, type A and type B. Thus, the vector of intermediate symbols after reordering can be expressed as $\mathbf{v} = (\mathbf{v}_A, \mathbf{v}_B)$, where $\mathbf{v}_A$ and $\mathbf{v}_B$ denote the vectors of intermediate symbols of types A and B

respectively. Furthermore, we denote the number of intermediate symbols of type A and B as h_A and h_B respectively. We have $h_A + h_B = h$. This Raptor code construction is characterized by a relationship between output symbols and intermediate symbols in the form

$$\mathbf{c} = \mathbf{v}\mathbf{G}_{\text{LT}} = (\mathbf{v}_A, \mathbf{v}_B)\mathbf{G}_{\text{LT}} = (\mathbf{v}_A, \mathbf{v}_B) \left[\begin{array}{c} \mathbf{G}_{\text{LT}}^{A} \\ \hline \mathbf{G}_{\text{LT}}^{B} \end{array} \right].$$

Under the assumption that n output symbols are generated, $\mathbf{G}_{\text{LT}}^{A}$ and $\mathbf{G}_{\text{LT}}^{B}$ have sizes $(h_A \times n)$ and $(h_B \times n)$ respectively. Each column of $\mathbf{G}_{\text{LT}}$ is generated by first drawing two output degrees j and s according to a joint probability distribution $\Omega_{j,s}$ whose bivariate generating function is[6]

$$\Omega(x, z) = \sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s}\, x^j z^s.$$

For each column, j different indices are drawn uniformly at random in $\{1, 2, \ldots, h_A\}$ and the elements of the column in $\mathbf{G}_{\text{LT}}^{A}$ at the rows corresponding to these indices are drawn independently and uniformly from $\mathbb{F}_q \backslash \{0\}$, while all other elements of the column of $\mathbf{G}_{\text{LT}}^{A}$ are set to zero. In a similar way, s different indices are picked uniformly at random in $\{1, 2, \ldots, h_B\}$ and the elements of the column in $\mathbf{G}_{\text{LT}}^{B}$ at the rows corresponding to these indices are drawn independently and uniformly from $\mathbb{F}_q \backslash \{0\}$, while all other elements of the column of $\mathbf{G}_{\text{LT}}^{B}$ are set to zero.

The third construction considered is referred to as *Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code*. This construction is relevant to $q > 2$, since otherwise it collapses to the first construction. It is similar to the first construction (Raptor code over $\mathbb{F}_q$), but all non-zero coefficients of $\mathbf{G}_{\text{LT}}$ are equal to $1 \in \mathbb{F}_q$. Thus, each column of $\mathbf{G}_{\text{LT}}$ is generated by first drawing an output degree d according to the degree distribution $\Omega = (\Omega_1, \Omega_2, \ldots, \Omega_{d_{\max}})$, and then by picking d different indices uniformly at random in $\{1, 2, \ldots, h\}$. Finally, the elements of the column with rows corresponding to these indices are set to 1, while all other elements of the column are set to zero. The relationship between input and output symbols is still given by (10), where vectors $\mathbf{c}$, $\mathbf{v}$ and $\mathbf{u}$ have elements in $\mathbb{F}_q$, matrix $\mathbf{G}_{\text{o}}$ has elements in $\mathbb{F}_q$ as well, and the elements of $\mathbf{G}_{\text{LT}}$ belong to $\{0, 1\} \subset \mathbb{F}_q$. The advantage of this construction is that encoding and decoding complexities are significantly reduced when using a standard computing platform, particularly when q is a power of 2.

[6] This definition implies $\Omega_{0,1} = \Omega_{1,0} = 0$ (besides $\Omega_{0,0} = 0$), which is in line with the distribution used for the RaptorQ code [24]. This assumption is practically motivated but is not strictly necessary.

Finally, the fourth construction considered is referred to as *multi-edge type Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code*. As its name indicates this construction is a combination of the second and third constructions described before. In particular, this construction is the same as the second construction, except for the fact that the non-zero elements in $\mathbf{G}_{\text{LT}}$, and therefore in $\mathbf{G}_{\text{LT}}^{A}$ and $\mathbf{G}_{\text{LT}}^{B}$, take always value 1.

This last construction closely resembles the RaptorQ code [24], representing the state of art fountain code at the time of writing. The RaptorQ code is built over $\mathbb{F}_{256}$. Its outer code is itself obtained as the serial concatenation of two block codes, the first code being a quasi-cyclic nonbinary LDPC code and the second code being a nonbinary code defined by a dense parity-check matrix. In particular, the quasi-cyclic LDPC code has all its nonzero elements in the parity-check matrix equal to $1 \in \mathbb{F}_{256}$, whereas the second code resembles a random code over $\mathbb{F}_{256}$. The intermediate symbols belong to two different classes, which are called *LT symbols* and *permanently inactive symbols*. The LT code is a $0/1$ LT code characterized by the bivariate degree distribution

$$\Omega(x, z) = \Omega(x) \left(\frac{z^2 + z^3}{2} \right)$$

where x and z are, respectively, the dummy variables associated with LT and permanently inactive symbols, and $\Omega(x)$ is a degree distribution with maximum output degree 30. Finally, we remark that the RaptorQ construction can be made systematic.[7] Thus, the RaptorQ code in its non-systematic form[8] is an example of the fourth construction considered in this paper (multi-edge type Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code). For more details about the RaptorQ construction as well as the design choices involved we refer the reader to [34].

V. Bounds on the Error Probability of Raptor Codes

This section contains the main contribution of this paper, a series of bounds on the performance of the different Raptor code constructions presented in Section IV. Proofs of these bounds are deferred to Section VI. The first theorem establishes a bound on the probability of decoding failure of a Raptor code over $\mathbb{F}_q$.

[7] A Raptor code is made systematic by adding a further precoding stage and specifying the seed of the pseudorandom generator which is used to generate the LT output symbols, see [31], [32] for more details.

[8] The RaptorQ code is in non-systematic form when random Encoding Symbol Identifiers (ESI) are used [33].

Theorem 1. *Consider a Raptor code over $\mathbb{F}_q$ with an (h, k) outer code $\mathcal{C}$ characterized by a weight enumerator A, and an inner LT code with output degree distribution Ω. The probability of decoding failure under ML erasure decoding, given that $k + \delta$ output symbols have been collected by the receiver, can be upper bounded as*

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{l=1}^{h} A_l \pi_l^{k+\delta} \tag{12}$$

where π_l is the probability that the generic output symbol y is equal to 0 given that the vector $\mathbf{v}$ of intermediate symbols has Hamming weight l. The expression of π_l is

$$\pi_l = \frac{1}{q} + \frac{q-1}{q} \sum_{j=1}^{d_{\max}} \Omega_j \frac{\mathcal{K}_j^{h,q}(l)}{\mathcal{K}_j^{h,q}(0)}. \tag{13}$$

The upper bound in Theorem 1 also applies to LT codes. In that case, $h = k$ and A_l is simply the total number of sequences of Hamming weight l and length k,

$$A_l = \binom{k}{l} (q-1)^l.$$

The upper bound thus obtained for LT codes coincides with the bound in [17, Theorem 1]. Theorem 1 may be extended to multi-edge type Raptor codes over $\mathbb{F}_q$ as follows.

Theorem 2. *Consider a multi-edge type Raptor code over $\mathbb{F}_q$ with an (h, k) outer code $\mathcal{C}$ characterized by a bivariate weight enumerator polynomial $A(x, z)$ and an inner LT code with bivariate output degree distribution $\Omega(x, z)$. The probability of decoding failure under ML erasure decoding given that $k + \delta$ output symbols have been collected by the receiver can be upper bounded as*

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{\substack{0 \leq l \leq h_A \\ 0 \leq t \leq h_B \\ l+t>0}} A_{l,t} \pi_{l,t}^{k+\delta}$$

where

$$\pi_{l,t} = \frac{1}{q} + \frac{q-1}{q} \sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s} \frac{\mathcal{K}_j^{h_A,q}(l)}{\mathcal{K}_j^{h_A,q}(0)} \frac{\mathcal{K}_s^{h_B,q}(t)}{\mathcal{K}_s^{h_B,q}(0)}. \tag{14}$$

The next result establishes a bound on the probability of decoding failure of a Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code.

Theorem 3. *Consider a Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code having an output degree distribution Ω and with an (h, k) outer code $\mathcal{C}$ characterized by a composition enumerator $\mathcal{Q}_{\mathbf{f}}$.*

The probability of decoding failure under ML erasure decoding given that $k+\delta$ output symbols have been collected by the receiver can be upper bounded as

$$\mathsf{P}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{\mathbf{f} \neq \varsigma(\mathbf{0})} \mathcal{Q}_{\mathbf{f}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\gamma} \in \mathsf{\Gamma}_j} B(\boldsymbol{\gamma}) \, \frac{\binom{f_0}{\gamma_0}\binom{f_1}{\gamma_1} \cdots \binom{f_{q-1}}{\gamma_{q-1}}}{\binom{h}{j}} \right)^{k+\delta} \tag{15}$$

where $\mathsf{\Gamma}_j$ is the set of all possible compositions for vectors in $\mathbb{F}_q^j$.

The upper bound in Theorem 3 can be extended to the multi-edge type case as follows.

Theorem 4. *Consider a multi-edge type Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code having bivariate output degree distribution $\Omega(x,z)$, and with an (h,k) outer code $\mathcal{C}$ characterized by a bivariate composition enumerator $\mathcal{Q}_{\mathbf{f}_A,\mathbf{f}_B}$. The probability of decoding failure under ML erasure decoding given that $k+\delta$ output symbols have been collected by the receiver can be upper bounded as*

$$\begin{aligned}\mathsf{P}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{\substack{\mathbf{f}_A,\mathbf{f}_B \\ \mathbf{f}_A+\mathbf{f}_B \neq \varsigma(\mathbf{0})}} \mathcal{Q}_{\mathbf{f}_A,\mathbf{f}_B} & \left(\sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s} \sum_{\boldsymbol{\gamma}_A \in \mathsf{\Gamma}_j} \sum_{\boldsymbol{\gamma}_B \in \mathsf{\Gamma}_s} B(\boldsymbol{\gamma}_A + \boldsymbol{\gamma}_B) \right. \\ & \left. \times \frac{\binom{f_{A,0}}{\gamma_{A,0}}\binom{f_{A,1}}{\gamma_{A,1}} \cdots \binom{f_{A,q-1}}{\gamma_{A,q-1}}}{\binom{h_A}{j}} \, \frac{\binom{f_{B,0}}{\gamma_{B,0}}\binom{f_{B,1}}{\gamma_{B,1}} \cdots \binom{f_{B,q-1}}{\gamma_{B,q-1}}}{\binom{h_B}{s}} \right)^{k+\delta}\end{aligned}$$

where $\mathsf{\Gamma}_j$ and $\mathsf{\Gamma}_s$ are the set of all possible compositions for vectors in $\mathbb{F}_q^j$ and in $\mathbb{F}_q^s$, respectively.

Each of the above theorems specializes the union bound (2) for a specific Raptor construction, providing an explicit expression for the corresponding S_1 parameter. By developing an expression for S_2, it is also possible to bound the decoding failure probability from below via (3) or (5). Hereafter we provide such a lower bound for a Raptor code over $\mathbb{F}_q$ with a $0/1$ LT code and, as a particular case, for a Raptor code over $\mathbb{F}_2$. The lower bounds exploit the sets $\mathscr{K}_{q,h}$ and $\mathscr{T}_{2,h}$ defined in Section III.

Theorem 5. *Consider a Raptor code over $\mathbb{F}_q$ with a 0/1 LT code having output degree distribution Ω, and an (h,k) outer code $\mathcal{C}$ characterized by a composition enumerator $\mathcal{Q}_{\mathbf{f}}$. The probability of decoding failure under ML erasure decoding, given that $k+\delta$ output symbols have been collected by the receiver, fulfills*

$$\mathsf{P}_{\mathsf{F}} \geq \frac{\theta S_1^2}{(2-\theta)S_1 + 2S_2} + \frac{(1-\theta)S_1^2}{(1-\theta)S_1 + 2S_2} \geq S_1 - S_2 \tag{16}$$

where $\theta = 2S_2/S_1 - \lfloor 2S_2/S_1 \rfloor$*,* S_1 *equals the right-hand side of* (15)*, and*

$$S_2 = \frac{1}{2(q-1)^2} \sum_{\boldsymbol{\kappa} \in \mathscr{K}_{q,h}} \mathcal{S}_{\boldsymbol{\kappa}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\upsilon} \in \Upsilon_j} B(\boldsymbol{\gamma}_1(\boldsymbol{\upsilon})) B(\boldsymbol{\gamma}_2(\boldsymbol{\upsilon})) \frac{\prod_{s,t} \binom{\kappa_{s,t}}{\upsilon_{s,t}}}{\binom{h}{j}} \right)^{k+\delta} . \tag{17}$$

In (17)*,* Υ_j *is the set of all possible joint compositions for vector pairs in* $\mathbb{F}_q^j \times \mathbb{F}_q^j$*.* *Moreover, for* $q = 2$*: (i) the parameter* S_1 *equals the right-hand side of* (12) *(expressed with* $q = 2$*); (ii) the parameter* S_2 *reduces to*

$$S_2 = \frac{1}{2} \sum_{\boldsymbol{\tau} \in \mathscr{T}_{2,h}} J_{\boldsymbol{\tau}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{(i_1,i_2,i_3)} \frac{\binom{\tau_0}{j-i_1-i_2-i_3}\binom{\tau_1}{i_1}\binom{\tau_2}{i_1}\binom{\tau_3}{i_3}}{\binom{h}{j}} \right)^{k+\delta} \tag{18}$$

where $J_{\boldsymbol{\tau}}$ *is the biweight enumerator of the outer code and where the most inner sum in* (18) *is over all integer triplets* (i_1, i_2, i_3) *such that* $i_1 + i_2 + i_3 = j$*; both* $i_1 + i_3$ *and* $i_2 + i_3$ *are even;* $0 \le i_1 \le \min\{\tau_1, j\}$*,* $0 \le i_2 \le \min\{\tau_2, j\}$*,* $0 \le i_3 \le \min\{\tau_3, j\}$*.*

Theorems 1-5 apply to Raptor codes with a given outer code. Next we extend these results to the case of a random outer code drawn from an ensemble of codes. Specifically, we consider a parity-check based ensemble of outer codes, denoted by $\mathscr{C}$, defined by a random matrix of size $(h-k) \times h$ whose elements belong to $\mathbb{F}_q$ (here, k may not coincide with the dimension of a specific code in the ensemble, as it will be discussed later). A linear block code of length h belongs to $\mathscr{C}$ if and only if at least one of the instances of the random matrix is a valid parity-check matrix for it. Moreover, the probability measure of each code in the ensemble is the sum of the probabilities of all instances of the random matrix which are valid parity-check matrices for that code. Note that all codes $\mathcal{C}$ in $\mathscr{C}$ are linear, have length h, and have dimension $k_{\mathcal{C}} \ge k$. In the following we use the expression *Raptor code ensemble* to refer to the set of Raptor codes obtained by concatenating an outer code belonging to the ensemble $\mathscr{C}$ with an LT code. Given a Raptor code ensemble we define its expected probability of decoding failure as

$$\bar{\mathsf{P}}_{\mathsf{F}} = \mathbb{E}_{\mathcal{C}}[\mathsf{P}_{\mathsf{F}}(\mathcal{C})] \tag{19}$$

where the expectation is taken over all codes $\mathcal{C}$ in the ensemble of outer codes $\mathscr{C}$.

The following corollary extends the result of Theorem 1 to Raptor code ensembles.

Corollary 1. *Consider a Raptor code ensemble over* $\mathbb{F}_q$ *with an outer code randomly drawn from the ensemble* $\mathscr{C}$*, characterized by an expected weight enumerator* $\mathsf{A} = \{\mathsf{A}_0, \mathsf{A}_1, \ldots, \mathsf{A}_h\}$ *and an LT code with degree distribution* Ω*. Under ML erasure decoding and given that* $k + \delta$

output symbols have been collected by the receiver, the expected probability of the decoding failure can be upper bounded as

$$\bar{\mathsf{P}}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{l=1}^{h} \mathsf{A}_l \pi_l^{k+\delta} .$$

The following three corollaries extend Theorems 2, 3, 4 and to Raptor code ensembles.

Corollary 2. *Consider a multi-edge type Raptor code ensemble over $\mathbb{F}_q$, whose outer code is randomly drawn from a code ensemble characterized by an expected bivariate weight enumerator polynomial $\mathsf{A}(x, z)$ and an inner LT code with bivariate output degree distribution $\Omega(x, z)$. The expected probability of decoding failure under ML erasure decoding given that $k + \delta$ output symbols have been collected by the receiver can be upper bounded as*

$$\bar{\mathsf{P}}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{\substack{0 \leq l \leq h_A \\ 0 \leq t \leq h_B \\ l+t>0}} \mathsf{A}_{l,t} \pi_{l,t}^{k+\delta}$$

where $\pi_{l,t}$ is defined in (14).

Corollary 3. *Consider an ensemble of Raptor codes over $\mathbb{F}_q$ with a $0/1$ LT code with degree distribution Ω and where the outer code is randomly drawn from a code ensemble $\mathscr{C}$ characterized by an expected composition enumerator $\mathsf{Q}_{\mathbf{f}}$. The expected probability of decoding failure under ML erasure decoding given that $k + \delta$ output symbols have been collected by the receiver can be upper bounded as*

$$\bar{\mathsf{P}}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{\mathbf{f} \neq \varsigma(\mathbf{0})} \mathsf{Q}_{\mathbf{f}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\gamma} \in \mathsf{\Gamma}_j} B(\boldsymbol{\gamma}) \, \frac{\binom{f_0}{\gamma_0}\binom{f_1}{\gamma_1} \cdots \binom{f_{q-1}}{\gamma_{q-1}}}{\binom{h}{j}} \right)^{k+\delta}$$

where $\mathsf{\Gamma}_j$ is the set of all possible compositions for vectors in $\mathbb{F}_q^j$.

Corollary 4. *Consider a multi-edge type Raptor code ensemble over $\mathbb{F}_q$ with a $0/1$ LT code with bivariate output degree distribution $\Omega(x, z)$ and where the outer code is randomly drawn from an ensemble $\mathscr{C}$ characterized by an expected bivariate composition enumerator $\mathsf{Q}_{\mathbf{f}_A, \mathbf{f}_B}$. The expected probability of decoding failure under ML erasure decoding given that $k + \delta$ output* *symbols have been collected by the receiver can be upper bounded as*

$$\bar{\mathsf{P}}_{\mathsf{F}} \le \frac{1}{q-1} \sum_{\substack{\mathbf{f}_A, \mathbf{f}_B \\ \mathbf{f}_A + \mathbf{f}_B \neq \varsigma(\mathbf{0})}} \mathsf{Q}_{\mathbf{f}_A, \mathbf{f}_B} \left(\sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s} \sum_{\boldsymbol{\gamma}_A \in \Gamma_j} \sum_{\boldsymbol{\gamma}_B \in \Gamma_s} B(\boldsymbol{\gamma}_A + \boldsymbol{\gamma}_B) \right.$$
$$\left. \times \frac{\binom{f_{A,0}}{\gamma_{A,0}} \binom{f_{A,1}}{\gamma_{A,1}} \cdots \binom{f_{A,q-1}}{\gamma_{A,q-1}}}{\binom{h_A}{j}} \frac{\binom{f_{B,0}}{\gamma_{B,0}} \binom{f_{B,1}}{\gamma_{B,1}} \cdots \binom{f_{B,q-1}}{\gamma_{B,q-1}}}{\binom{h_B}{s}} \right)^{k+\delta}$$

where Γ_j and Γ_s are the set of all possible compositions for vectors in $\mathbb{F}_q^j$ and in $\mathbb{F}_q^s$, respectively.

Theorem 5 can also be extended to Raptor code ensembles where the outer code is drawn from an ensemble of linear block codes all with the same block length.

Corollary 5. *Consider an ensemble of Raptor codes over $\mathbb{F}_q$ with a 0/1 LT code with degree distribution Ω, where the outer code is drawn randomly from a code ensemble $\mathscr{C}$ characterized by an expected composition enumerator $\mathcal{Q}_{\mathbf{f}}$ and an expected bicomposition enumerator $\mathsf{S}_{\boldsymbol{\kappa}}$. The probability of decoding failure under ML erasure decoding, given that m output symbols have been collected by the receiver, fulfills*

$$\bar{\mathsf{P}}_{\mathsf{F}} \ge \frac{\bar{\theta}\,[\bar{S}_1(m)]^2}{(2-\bar{\theta})\bar{S}_1(m) + 2\bar{S}_2(m)} + \frac{(1-\bar{\theta})[\bar{S}_1(m)]^2}{(1-\bar{\theta})\bar{S}_1(m) + 2\bar{S}_2(m)} \ge \bar{S}_1(m) - \bar{S}_2(m) \tag{20}$$

where $\bar{\theta} = 2\bar{S}_2(m)/\bar{S}_1(m) - \lfloor 2\bar{S}_2(m)/\bar{S}_1(m) \rfloor$ and

$$\bar{S}_1(m) = \frac{1}{q-1} \sum_{\mathbf{f} \neq \varsigma(\mathbf{0})} \mathsf{Q}_{\mathbf{f}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\gamma} \in \Gamma_j} B(\boldsymbol{\gamma}) \frac{\binom{f_0}{\gamma_0} \binom{f_1}{\gamma_1} \cdots \binom{f_{q-1}}{\gamma_{q-1}}}{\binom{h}{j}} \right)^m \tag{21}$$

$$\bar{S}_2(m) = \frac{1}{2(q-1)^2} \sum_{\boldsymbol{\kappa} \in \mathscr{K}_{q,h}} \mathcal{S}_{\boldsymbol{\kappa}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\upsilon} \in \Upsilon_j} B(\boldsymbol{\gamma}_1(\boldsymbol{\upsilon})) B(\boldsymbol{\gamma}_2(\boldsymbol{\upsilon})) \frac{\prod_{s,t} \binom{\kappa_{s,t}}{\upsilon_{s,t}}}{\binom{h}{j}} \right)^m . \tag{22}$$

In (22), Υ_j is the set of all possible joint compositions for vector pairs in $\mathbb{F}_q^j \times \mathbb{F}_q^j$. Moreover, in the particular case $q = 2$ we have $\bar{S}_1(m) = \sum_{l=1}^{h} \mathsf{A}_l \pi_l^m$ where π_l is given by (13) (with $q = 2$) and

$$\bar{S}_2(m) = \frac{1}{2} \sum_{\boldsymbol{\tau} \in \mathscr{T}_{2,h}} \mathsf{J}_{\boldsymbol{\tau}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{(i_1, i_2, i_3)} \frac{\binom{\tau_0}{j - i_1 - i_2 - i_3} \binom{\tau_1}{i_1} \binom{\tau_2}{i_1} \binom{\tau_3}{i_3}}{\binom{h}{j}} \right)^m . \tag{23}$$

In (23), $\mathsf{J}_{\boldsymbol{\tau}}$ is the average bicomposition enumerator of the outer code ensemble. Furthermore, the most inner sum is over all integer triplets (i_1, i_2, i_3) such that $i_1 + i_2 + i_3 = j$; both $i_1 + i_3$ and $i_2 + i_3$ are even; $0 \le i_1 \le \min\{\tau_1, j\}$, $0 \le i_2 \le \min\{\tau_2, j\}$, $0 \le i_3 \le \min\{\tau_3, j\}$.

Remark 2. *Note that the bounds provided in Corollaries* (1) *to* (5) *hold also for Raptor code ensembles based on outer codes of fixed dimension k (e.g., systematic-form generator-based outer code ensembles). The proof for this case is trivial, and follows from the linearity of the expectation. The proofs for the case where the outer code is drawn from a parity-check ensemble require some more care, as illustrated in the following section.*

VI. Derivation of the Bounds

This section contains the proofs of the results presented in Section V.

1) Proof of Theorem 1: The proof follows the same approach as for [17, Theorem 1]. An ML decoder solves the linear system of equations in (11). Decoding fails whenever the system does not admit a unique solution, that is, if and only if $\mathsf{rank}(\tilde{\mathbf{G}}) < k$, i.e., if $\exists\, \mathbf{u} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}$ s.t. $\mathbf{u}\tilde{\mathbf{G}} = \mathbf{0}$. For any two vectors $\mathbf{u} \in \mathbb{F}_q^k$ and $\mathbf{v} \in \mathbb{F}_q^h$, we define $E_{\mathbf{u}}$ as the event $\mathbf{u}\mathbf{G}_{\text{o}}\tilde{\mathbf{G}}_{\text{LT}} = \mathbf{0}$, and $E_{\mathbf{v}}$ as the event $\mathbf{v}\tilde{\mathbf{G}}_{\text{LT}} = \mathbf{0}$. We have

$$\mathsf{P_F} = \Pr\left\{ \bigcup_{\mathbf{u} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}} E_{\mathbf{u}} \right\} = \Pr\left\{ \bigcup_{\mathbf{v} \in \mathcal{C} \backslash \{\mathbf{0}\}} E_{\mathbf{v}} \right\} \tag{24}$$

where we made use of the fact that due to outer code linearity, the all zero intermediate word is only generated by the all zero input vector.

Due to linearity of the outer code, if $\mathbf{v} \in \mathcal{C}$, then $\beta\mathbf{v} \in \mathcal{C}$ for any $\beta \in \mathbb{F}_q \backslash \{0\}$. Furthermore, for any $\beta \in \mathbb{F}_q \backslash \{0\}$, $\mathbf{v}\tilde{\mathbf{G}}_{\text{LT}} = 0$ if and only if $\beta\mathbf{v}\tilde{\mathbf{G}}_{\text{LT}} = 0$. Thus, for any two outer codewords $\mathbf{v}_1$ and $\mathbf{v}_2$ such that $\mathbf{v}_1 = \beta\mathbf{v}_2$ for some $\beta \in \mathbb{F}_q \setminus \{0\}$, the event $E_{\mathbf{v}_1}$ holds if and only if $E_{\mathbf{v}_2}$ does, and we have $E_{\mathbf{v}_1} \cup E_{\mathbf{v}_2} = E_{\mathbf{v}_1}$. If we take a union bound on (24), this allows us dividing it by a factor $q-1$, leading to

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{\mathbf{v} \in \mathcal{C} \backslash \{\mathbf{0}\}} \Pr\{E_{\mathbf{v}}\}. \tag{25}$$

Defining $\mathcal{C}_l$ as $\mathcal{C}_l = \{\mathbf{v} \in \mathcal{C} : w(\mathbf{v}) = l\}$, the expression can be developed as

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{l=1}^{h} \sum_{\mathbf{v} \in \mathcal{C}_l} \Pr\{E_{\mathbf{v}}\} = \frac{1}{q-1} \sum_{l=1}^{h} A_l \Pr\{E_{\mathbf{v}} | w(\mathbf{v}) = l\}$$

where we made use of the fact that, since the neighbors of an output symbol are chosen uniformly at random, $\Pr\{E_{\mathbf{v}}\}$ does not depend on the specific vector $\mathbf{v}$, but only on its Hamming weight.

Observing that the output symbols are independent of each other, we have

$$\Pr\{E_{\mathbf{v}} | w(\mathbf{v}) = l\} = \pi_l^{k+\delta}$$

where $\pi_l = \Pr\{y = 0 | w(\mathbf{v}) = l\}$.

Let J and I be discrete random variables representing the number of intermediate symbols which are linearly combined to generate the generic output symbol y, and the number of non-zero such intermediate symbols, respectively. Note that $I \leq \min\{J, w(\mathbf{v})\}$. An expression for π_l may be obtained as

$$\begin{aligned}
\pi_l &= \sum_{j=1}^{d_{\max}} \Pr\{y = 0 | w(\mathbf{v}) = l, J = j\} \Pr\{J = j | w(\mathbf{v}) = l\} \\
&\overset{\text{(a)}}{=} \sum_{j=1}^{d_{\max}} \Omega_j \Pr\{y = 0 | w(\mathbf{v}) = l, J = j\} \\
&\overset{\text{(b)}}{=} \sum_{j=1}^{d_{\max}} \Omega_j \sum_{i=0}^{\min\{j,l\}} \Pr\{y = 0 | I = i\} \Pr\{I = i | w(\mathbf{v}) = l, J = j\}
\end{aligned}$$

where (a) is due to

$$\Pr\{J = j | w(\mathbf{v}) = l\} = \Pr\{J = j\} = \Omega_j$$

and (b) to

$$\Pr\{y = 0 | w(\mathbf{v}) = l, J = j, I = i\} = \Pr\{y = 0 | I = i\}.$$

Letting $\vartheta_{i,l,j} = \Pr\{I = i | w(\mathbf{v}) = l, J = j\}$, since the j intermediate symbols are chosen uniformly at random by the LT encoder we have

$$\vartheta_{i,l,j} = \frac{\binom{l}{i}\binom{h-l}{j-i}}{\binom{h}{j}}. \tag{26}$$

Let us denote $\Pr\{y = 0 | I = i\}$ by φ_i and let us observe that the non-zero elements of $\tilde{\mathbf{G}}_{\mathrm{LT}}$ are i.i.d. and uniformly drawn in $\mathbb{F}_q \setminus \{0\}$. On invoking Lemma 4 in Appendix A,[9] we have

$$\varphi_i = \frac{1}{q}\left(1 + \frac{(-1)^i}{(q-1)^{i-1}}\right). \tag{27}$$

We conclude that π_l is given by

$$\pi_l = \sum_{j=1}^{d_{\max}} \Omega_j \sum_{i=0}^{\min\{j,l\}} \vartheta_{i,l,j}\, \varphi_i$$

[9] The proof in Appendix A is only valid for fields with characteristic 2, the case of most interest for practical purposes. The proof of the general case is a simple extension of Lemma 4.

where $\vartheta_{i,l,j}$ and φ_i are given by (26) and (27), respectively. Expanding this expression and rewriting it using Krawtchouk polynomials and making use of the Chu-Vandermonde identity, one obtains (13).[10] ■

We remark that (25) holds not only for Raptor codes over $\mathbb{F}_q$, but also for the other three considered constructions. Hence, (25) represents the starting point in all subsequent proofs.

2) Proof of Theorem 2: For this construction we may develop (25) as

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{\substack{0 \leq l \leq h_A \\ 0 \leq t \leq h_B \\ l+t>0}} \sum_{\mathbf{v} \in \mathcal{C}_{l,t}} \Pr\left\{E_{\mathbf{v}}\right\} \tag{28}$$

where $\mathcal{C}_{l,t}$ is the set of codewords in $\mathcal{C}$ with l non-zero elements in $\mathbf{v}_A$ and t non-zero elements in $\mathbf{v}_B$, formally $\mathcal{C}_{l,t} = \{\mathbf{v} = (\mathbf{v}_A, \mathbf{v}_B) \in \mathcal{C} : w(\mathbf{v}_A) = l, w(\mathbf{v}_B) = t\}$. Making use of the bivariate weight enumerator of the outer code, we can rewrite (28) as

$$\mathsf{P_F} \leq \frac{1}{q-1} \sum_{\substack{0 \leq l \leq h_A \\ 0 \leq t \leq h_B \\ l+t>0}} A_{l,t} \Pr\{E_{\mathbf{v}} | w(\mathbf{v}_A) = l, w(\mathbf{v}_B) = t\}$$

where we made use of the fact that since the neighbors of an output symbol are chosen uniformly at random, $\Pr\left\{E_{\mathbf{v}}\right\}$ does not depend on the particular vector $\mathbf{v}$, but only on its split Hamming weight, $w(\mathbf{v}_A) = l$ and $w(\mathbf{v}_B) = t$.

Since output symbols are generated independently of each other

$$\Pr\{E_{\mathbf{v}} | w(\mathbf{v}_A) = l, w(\mathbf{v}_B) = t\} = \pi_{l,t}^{k+\delta}$$

where $\pi_{l,t} = \Pr\{y = 0 | w(\mathbf{v}_A) = l, w(\mathbf{v}_B) = t\}$.

Let J and I be two discrete random variables representing, respectively, the number of intermediate symbols of type A which are linearly combined to generate output symbol y, and the number of non-zero such intermediate symbols. Similarly, let S and D be two discrete random variables representing, respectively, the number of intermediate symbols of type B which are linearly combined to generate output symbol y, and the number of non-zero such intermediate symbols. Note that we have $I \leq \min\{J, w(\mathbf{v}_A)\}$ and $D \leq \min\{S, w(\mathbf{v}_B)\}$. The expression of $\pi_{l,t}$ can be obtained as

[10] The expression of π_l was derived in [17], where an upper bound on the performance of LT codes was derived. However, the derivation of π_l in [17] is different from the one we provide in this paper.

$$
\begin{aligned}
\pi_{l,t} &= \sum_{j=1}^{h_A}\sum_{s=1}^{h_B} \Pr\{y=0|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t, J=j, S=s\} \\
&\quad \times \Pr\{J=j, S=s|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t\} \\
&\overset{\text{(a)}}{=} \sum_{j=1}^{h_A}\sum_{s=1}^{h_B} \Omega_{j,s} \Pr\{y=0|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t, J=j, S=s\} \\
&\overset{\text{(b)}}{=} \sum_{j=1}^{h_A}\sum_{s=1}^{h_B} \Omega_{j,s} \sum_{i=0}^{\min(j,l)} \sum_{d=0}^{\min(s,t)} \Pr\{y=0|I=i, D=d\} \\
&\quad \times \Pr\{I=i, D=d|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t, J=j, S=s\} \\
&\overset{\text{(c)}}{=} \sum_{j=1}^{h_A}\sum_{s=1}^{h_B} \Omega_{j,s} \sum_{i=0}^{\min(j,l)} \sum_{d=0}^{\min(s,t)} \Pr\{y=0|I=i, D=d\} \\
&\quad \times \Pr\{I=i|w(\mathbf{v}_A)=l, J=j\} \Pr\{D=d|w(\mathbf{v}_B)=t, S=s\}
\end{aligned}
$$

where (a) is due to

$$\Pr\{J=j, S=s|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t\} = \Pr\{J=j, S=s\} = \Omega_{j,s}$$

(b) is due to

$$\Pr\{y=0|w(\mathbf{v}_A)=l, w(\mathbf{v}_B)=t, J=j, S=s, I=i, D=d\} = \Pr\{y=0|I=i, D=d\}$$

and (c) follows from independence of I and D. Let us denote $\Pr\{y=0|I=i, D=d\}$ by $\varphi_{i,d}$. Since the non-zero elements of $\tilde{\mathbf{G}}_{\text{LT}}$ are i.i.d. and uniformly drawn in $\mathbb{F}_q \setminus \{0\}$, on invoking Lemma 4 in the Appendix we have

$$\varphi_{i,d} = \frac{1}{q}\left(1 + \frac{(-1)^{i+d}}{(q-1)^{i+d-1}}\right).$$

Similarly, letting $\vartheta^{(A)}_{i,l,j} = \Pr\{I=i|w(\mathbf{v}_A)=l, J=j\}$, we have

$$\vartheta^{(A)}_{i,l,j} = \frac{\binom{l}{i}\binom{h_A-l}{j-i}}{\binom{h_A}{j}}.$$

If we now define $\vartheta^{(B)}_{d,t,s} = \Pr\{D=d|w(\mathbf{v}_B)=t, S=s\}$ and use the same reasoning for the intermediate symbols of type B, we have

$$\vartheta^{(B)}_{d,t,s} = \frac{\binom{t}{d}\binom{h_B-t}{s-d}}{\binom{h_B}{s}}.$$

Hence, the expression of $\pi_{l,t}$ is given by

$$\pi_{l,t} = \sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s} \sum_{i=0}^{\min(j,l)} \sum_{d=0}^{\min(s,t)} \varphi_{i,d}\, \vartheta^{(A)}_{i,l,j}\, \vartheta^{(B)}_{d,t,s}$$

Expanding and rewriting this expression using Krawtchouk polynomials yields (14). ■

3) Proof of Theorem 3: Starting again from (25) and defining $\mathcal{C}_{\mathbf{f}}$ as the set of codewords with composition $\mathbf{f}$ in the outer code $\mathcal{C}$, i.e., $\mathcal{C}_{\mathbf{f}} = \{\mathbf{v} \in \mathcal{C} : \varsigma(\mathbf{v}) = \mathbf{f}\}$, we have

$$\mathsf{P}_{\mathsf{F}} \leq \frac{1}{q-1} \sum_{\mathbf{f} \neq \varsigma(\mathbf{0})} \sum_{\mathbf{v} \in \mathcal{C}_{\mathbf{f}}} \Pr\{E_{\mathbf{v}}\} = \frac{1}{q-1} \sum_{\mathbf{f} \neq \varsigma(\mathbf{0})} \mathcal{Q}_{\mathbf{f}} \Pr\{E_{\mathbf{v}} | \varsigma(\mathbf{v}) = \mathbf{f}\}$$

where we made use of the fact that since the neighbors of an output symbol are chosen uniformly at random, any two codewords having the same composition are characterized by the same probability $\Pr\{E_{\mathbf{v}}\}$.

Due to independence among the output symbols, we have

$$\Pr\{E_{\mathbf{v}} | \varsigma(\mathbf{v}) = \mathbf{f}\} = \left(\Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}\}\right)^{k+\delta}.$$

Let us now introduce again an auxiliary discrete random variable J to represent the output symbol degree, i.e., the number of intermediate symbols which are summed to generate the generic output symbol y. We have

$$\Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}\} = \sum_{j=1}^{d_{\max}} \Omega_j \Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\}.$$

Next, let us introduce the random vector $\mathbf{\Gamma}$ representing the composition of the j intermediate output symbols that are added to obtain output symbol y. Recalling that $\mathsf{\Gamma}_j$ is the set of possible compositions of length-j vectors, we can recast $\Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\}$ as

$$\begin{aligned}
\Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\} &= \sum_{\boldsymbol{\gamma} \in \mathsf{\Gamma}_j} \Pr\{y = 0 | \varsigma(\mathbf{v}) = \mathbf{f}, J = j, \mathbf{\Gamma} = \boldsymbol{\gamma}\} \Pr\{\mathbf{\Gamma} = \boldsymbol{\gamma} | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\} \\
&= \sum_{\boldsymbol{\gamma} \in \mathsf{\Gamma}_j} \Pr\{y = 0 | \mathbf{\Gamma} = \boldsymbol{\gamma}\} \Pr\{\mathbf{\Gamma} = \boldsymbol{\gamma} | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\} \\
&= \sum_{\boldsymbol{\gamma} \in \mathsf{\Gamma}_j} B(\boldsymbol{\gamma}) \Pr\{\mathbf{\Gamma} = \boldsymbol{\gamma} | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\}
\end{aligned}$$

where the indicator function B has been defined in Section II. The term $\Pr\{\mathbf{\Gamma} = \boldsymbol{\gamma} | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\}$ can easily be computed making use of a multivariate hypergeometric distribution. In particular:

$$\Pr\{\mathbf{\Gamma} = \boldsymbol{\gamma} | \varsigma(\mathbf{v}) = \mathbf{f}, J = j\} = \frac{\binom{f_0}{\gamma_0}\binom{f_1}{\gamma_1} \cdots \binom{f_{q-1}}{\gamma_{q-1}}}{\binom{h}{j}}.$$

■

4) Proof of Theorem 4: The proof tightly follows the proofs of Theorems 2 and 3. Let us define $\mathcal{C}_{\mathbf{f}_A,\mathbf{f}_B}$ as the set of codewords in $\mathcal{C}$ where $\mathbf{v}_A$ $\mathbf{v}_B$ have, respectively, composition $\mathbf{f}_A$ and $\mathbf{f}_B$, formally $\mathcal{C}_{\mathbf{f}_A,\mathbf{f}_B} = \{\mathbf{v} = (\mathbf{v}_A, \mathbf{v}_B) \in \mathcal{C} : \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\}$. From (25) we obtain

$$\begin{aligned}
\mathsf{P_F} &\leq \frac{1}{q-1} \sum_{\substack{\mathbf{f}_A,\mathbf{f}_B \\ \mathbf{f}_A+\mathbf{f}_B \neq \varsigma(\mathbf{0})}} \sum_{\mathbf{v}\in\mathcal{C}_{\mathbf{f}_A,\mathbf{f}_B}} \Pr\{E_\mathbf{v}\} \\
&= \frac{1}{q-1} \sum_{\substack{\mathbf{f}_A,\mathbf{f}_B \\ \mathbf{f}_A+\mathbf{f}_B \neq \varsigma(\mathbf{0})}} \mathcal{Q}_{\mathbf{f}_A,\mathbf{f}_B} \Pr\{E_\mathbf{v} | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\}.
\end{aligned}$$

Again we exploited the fact that since the neighbors of an output symbol are chosen uniformly at random, $\Pr\{E_\mathbf{v}\}$ depends only on the split composition of $\mathbf{v}$, $\varsigma(\mathbf{v}_A) = \mathbf{f}_A$ and $\varsigma(\mathbf{v}_B) = \mathbf{f}_B$.

Due to independence among the output symbols, we have

$$\Pr\{E_\mathbf{v} | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\} = \left(\Pr\{y = 0 | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\}\right)^{k+\delta}.$$

Introducing the two auxiliary discrete random variables, J and S representing, respectively, the number of intermediate symbols of type A and B which are summed to generate the generic output symbol y, we have

$$\begin{aligned}
&\Pr\{y = 0 | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\} \\
&\qquad = \sum_{j=1}^{h_A} \sum_{s=1}^{h_B} \Omega_{j,s} \Pr\{y = 0 | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B, J = j, S = s\}.
\end{aligned}$$

Next, let the two random vectors $\mathbf{\Gamma}_A$ and $\mathbf{\Gamma}_B$ represent, respectively, the composition of the j intermediate symbols of type A and s intermediate symbols of type B that are added to obtain output symbol y. Let us also recall that $\mathsf{\Gamma}_j$ and $\mathsf{\Gamma}_s$ represent the set of possible compositions of length-j and s vectors, respectively. We can recast the rightmost term in the last expression as

$$\begin{aligned}
&\Pr\{y = 0 | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B, J = j, S = s\} \\
&\quad = \sum_{\boldsymbol{\gamma}_A \in \mathsf{\Gamma}_j} \sum_{\boldsymbol{\gamma}_B \in \mathsf{\Gamma}_s} \Pr\{y = 0 | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B, J = j, S = s, \mathbf{\Gamma}_A = \boldsymbol{\gamma}_A, \mathbf{\Gamma}_B = \boldsymbol{\gamma}_B\} \\
&\qquad \times \Pr\{\mathbf{\Gamma}_A = \boldsymbol{\gamma}_A, \mathbf{\Gamma}_B = \boldsymbol{\gamma}_B | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B, J = j, S = s\} \\
&\quad = \sum_{\boldsymbol{\gamma}_A \in \mathsf{\Gamma}_j} \sum_{\boldsymbol{\gamma}_B \in \mathsf{\Gamma}_s} \Pr\{y = 0 | \mathbf{\Gamma}_A = \boldsymbol{\gamma}_A, \mathbf{\Gamma}_B = \boldsymbol{\gamma}_B\} \\
&\qquad \times \Pr\{\mathbf{\Gamma}_A = \boldsymbol{\gamma}_A, \mathbf{\Gamma}_B = \boldsymbol{\gamma}_B | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B, J = j, S = s\}
\end{aligned}$$

$$= \sum_{\boldsymbol{\gamma}_A \in \Gamma_j} \sum_{\boldsymbol{\gamma}_B \in \Gamma_s} B(\boldsymbol{\gamma}_A + \boldsymbol{\gamma}_B) \Pr\left\{\boldsymbol{\Gamma}_A = \boldsymbol{\gamma}_A | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, J = j\right\} \Pr\left\{\boldsymbol{\Gamma}_B = \boldsymbol{\gamma}_B | \varsigma(\mathbf{v}_B) = \mathbf{f}_B, S = s\right\}.$$

The term $\Pr\left\{\boldsymbol{\Gamma}_A = \boldsymbol{\gamma}_A | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, J = j\right\}$ can easily be computed making use of a multivariate hypergeometric distribution. Concretely, we have

$$\Pr\{\boldsymbol{\Gamma}_A = \boldsymbol{\gamma}_A | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, J = j\} = \frac{\binom{f_{A,0}}{\gamma_{A,0}}\binom{f_{A,1}}{\gamma_{A,1}} \cdots \binom{f_{A,q-1}}{\gamma_{A,q-1}}}{\binom{h_A}{j}}$$

and the same holds for

$$\Pr\{\boldsymbol{\Gamma}_B = \boldsymbol{\gamma}_B | \varsigma(\mathbf{v}_B) = \mathbf{f}_B, S = s\} = \frac{\binom{f_{B,0}}{\gamma_{B,0}}\binom{f_{B,1}}{\gamma_{B,1}} \cdots \binom{f_{B,q-1}}{\gamma_{B,q-1}}}{\binom{h_B}{s}}.$$

■

5) Proof of Theorem 5: Applying to the outer codebook the indexing and partition described in Section III we can write

$$\begin{aligned}
\mathsf{P_F} = \Pr\left\{ \bigcup_{\mathbf{v} \in \mathcal{C} \setminus \{\mathbf{0}\}} E_{\mathbf{v}} \right\} &\overset{\text{(a)}}{=} \Pr\left\{ \bigcup_{a=1}^{M_{q,k}-1} E_{\tilde{\mathbf{v}}_a} \right\} \\
&\overset{\text{(b)}}{\geq} \sum_{a=1}^{M_{q,k}-1} \Pr\{E_{\tilde{\mathbf{v}}_a}\} - \sum_{0<a<b<M_{q,k}} \Pr\{E_{\tilde{\mathbf{v}}_a} \cap E_{\tilde{\mathbf{v}}_b}\} \\
&\overset{\text{(c)}}{=} \sum_{a=1}^{M_{q,k}-1} \Pr\{E_{\tilde{\mathbf{v}}_a}\} - \frac{1}{2} \sum_{(s,t) \in \tilde{\mathscr{D}}_{q,k}} \Pr\{E_{\mathbf{v}_s} \cap E_{\mathbf{v}_t}\} \\
&\overset{\text{(d)}}{=} \frac{1}{q-1} \sum_{\mathbf{v} \in \mathcal{C} \setminus \{0\}} \Pr(E_{\mathbf{v}}) - \frac{1}{2(q-1)^2} \sum_{(s,t) \in \mathscr{D}_{q,k}} \Pr\{E_{\mathbf{v}_s} \cap E_{\mathbf{v}_t}\}
\end{aligned}$$

where: (a) is due to the fact that, if two codewords $\mathbf{v}$ and $\mathbf{z}$ belong to the same part $\mathcal{P}_a$ (i.e., they are linearly dependent), then $E_{\mathbf{v}}$ occurs if and only if $E_{\mathbf{z}}$ occurs; (b) is a direct application of degree-two Bonferroni inequality (3); (c) follows from the definition of $\tilde{\mathscr{D}}_{q,k}$ given in Section III and from $\Pr\{E_{\mathbf{v}_s} \cap E_{\mathbf{v}_t}\} = \Pr\{E_{\mathbf{v}_t} \cap E_{\mathbf{v}_s}\}$ for any s and t; (d) is due the definition of $\mathscr{D}_{q,k}$ given in Section III and to the fact that, if $\mathbf{v}_1$ and $\mathbf{v}_2$ belong to some part $\mathcal{P}_a$ and $\mathbf{z}_1$ and $\mathbf{z}_2$ belong to another part $\mathcal{P}_b$, then $E_{\mathbf{v}_1} \cap E_{\mathbf{z}_1}$ occurs if and only if $E_{\mathbf{v}_2} \cap E_{\mathbf{z}_2}$ occurs. The last obtained expression is a degree-two Bonferroni lower bound for $\mathsf{P_F}$ in the form $\mathsf{P_F} \geq S_1 - S_2$. The term

S_1 has been developed in Theorem 3 and equals the right-hand side of (15). The term S_2 can be further developed as

$$\begin{aligned} S_2 &= \frac{1}{2(q-1)^2} \sum_{(s,t)\in\mathscr{D}_{q,k}} \Pr\{E_{\mathbf{v}_s} \cap E_{\mathbf{v}_t}\} \overset{\text{(e)}}{=} \frac{1}{2(q-1)^2} \sum_{\boldsymbol{\kappa}\in\mathscr{K}_{q,h}} \mathcal{S}_{\boldsymbol{\kappa}} \Pr\{E_{\mathbf{v}} \cap E_{\mathbf{z}} | \kappa(\mathbf{v},\mathbf{z}) = \boldsymbol{\kappa}\} \\ &\overset{\text{(f)}}{=} \frac{1}{2(q-1)^2} \sum_{\boldsymbol{\kappa}\in\mathscr{K}_{q,h}} \mathcal{S}_{\boldsymbol{\kappa}} \left(\Pr\{\{y_{\mathbf{v}}=0\}\cap\{y_{\mathbf{z}}=0\}|\, \kappa(\mathbf{v},\mathbf{z})\}\right)^{k+\delta}. \end{aligned} \tag{29}$$

In the previous equation array, (e) holds since the probability $\Pr\{E_{\mathbf{v}} \cap E_{\mathbf{z}}\}$ is the same for all codeword pairs $(\mathbf{v},\mathbf{z})$ with the same bicomposition. In (f) we have denoted by $y_{\mathbf{v}}$ the output symbol given that $\mathbf{v}$ is the intermediate codeword and we have exploited independence of output symbols.

Next, let the random variable J represent the output symbol degree. Moreover, for given bicomposition $\kappa(\mathbf{v},\mathbf{z}) = \boldsymbol{\kappa}$ and given $J=j$, define $\boldsymbol{\Upsilon}$ as the joint composition of the the two vectors in $\mathbb{F}_q^j$ representing the j symbols selected in $\mathbf{v}$ and $\mathbf{z}$. We have

$$\begin{aligned} \Pr\{\{y_{\mathbf{v}}=0\}\cap\{y_{\mathbf{z}}=0\}|\, \kappa(\mathbf{v},\mathbf{z}) = \boldsymbol{\kappa}\} &= \sum_{j=1}^{d_{\max}} \Omega_j \Pr\{\{y_{\mathbf{v}}=0\}\cap\{y_{\mathbf{z}}=0\}|\, \kappa(\mathbf{v},\mathbf{z}) = \boldsymbol{\kappa}, J=j\} \\ &= \sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\upsilon}\in\Upsilon_j} \Pr\{\{y_{\mathbf{v}}=0\}\cap\{y_{\mathbf{z}}=0\}|\boldsymbol{\Upsilon}=\boldsymbol{\upsilon}\} \Pr\{\boldsymbol{\Upsilon}=\boldsymbol{\upsilon}|J=j, \kappa(\mathbf{v},\mathbf{z})=\boldsymbol{\kappa}\} \\ &= \sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\upsilon}\in\Upsilon_j} B(\boldsymbol{\gamma}_1(\boldsymbol{\upsilon}))B(\boldsymbol{\gamma}_2(\boldsymbol{\upsilon})) \frac{\prod_{\substack{0\le s\le q-1\\0\le t\le q-1}} \binom{\kappa_{s,t}}{\upsilon_{s,t}}}{\binom{h}{j}}. \end{aligned} \tag{30}$$

where $\boldsymbol{\gamma}_1(\boldsymbol{\upsilon})$ and $\boldsymbol{\gamma}_2(\boldsymbol{\upsilon})$, defined in (8) and (9), are the compositions corresponding to $\boldsymbol{\upsilon}$. Expression (17) is obtained by substituting (30) into (29). The two bounds in (16) then follow as a direct application of degree-two Bonferroni and Dawson-Sankoff bounds, and from the observation that Dawson-Sankoff bound is tighter than the $S_1 - S_2$ one.

For $q=2$, the right-hand sides of (12) and (15) coincide. The S_1 term is therefore equal to right-hand side of (12) expressed with $q=2$. Next, recall from Remark 1 that for $q=2$ there is a one-to-one correspondence between joint compositions and joint weights. With this correspondence in mind we can write $\mathcal{S}_{\boldsymbol{\kappa}} = J_{\boldsymbol{\tau}}$. Again owing to this correspondence, we can establish a bijection between the set of joint compositions $\mathscr{K}_{2,k}$ and the set of joint weights $\mathscr{T}_{2,h}$. The right-hand side of (17) may thus be recast as

$$\frac{1}{2} \sum_{\boldsymbol{\tau}\in\mathscr{T}_h} J_{\boldsymbol{\tau}} \left(\sum_{j=1}^{d_{\max}} \Omega_j \sum_{\boldsymbol{\upsilon}\in\Upsilon_j} B(\boldsymbol{\gamma}_1(\boldsymbol{\upsilon}))B(\boldsymbol{\gamma}_2(\boldsymbol{\upsilon})) \frac{\binom{\tau_0}{\upsilon_{0,0}}\binom{\tau_1}{\upsilon_{0,1}}\binom{\tau_2}{\upsilon_{1,0}}\binom{\tau_3}{\upsilon_{1,1}}}{\binom{h}{j}} \right)^{k+\delta}$$

which yields the statement by simply letting

$$\boldsymbol{v} = \begin{bmatrix} j - i_1 - i_2 - i_3 & i_1 \\ i_2 & i_3 \end{bmatrix}.$$

■

6) Proof of Corollary 1: Due to Theorem 1 we may write

$$\bar{\mathsf{P}}_{\mathsf{F}} \leq \mathbb{E}_{\mathcal{C}}\left[\frac{1}{q-1}\sum_{l=1}^{h} A_l(\mathcal{C})\pi_l^{k_{\mathcal{C}}+\delta}\right]. \tag{31}$$

For all outer codes $\mathcal{C} \in \mathscr{C}$ we have $k_{\mathcal{C}} \geq k$. Since $\pi_l \leq 1$ we can write $\pi_l^{k_{\mathcal{C}}+\delta} \leq \pi_l^{k+\delta}$ which allows us to upper bound (31) as

$$\bar{\mathsf{P}}_{\mathsf{F}} \leq \mathbb{E}_{\mathcal{C}}\left[\frac{1}{q-1}\sum_{l=1}^{h} A_l(\mathcal{C})\pi_l^{k+\delta}\right] = \frac{1}{q-1}\sum_{l=1}^{h} \mathsf{A}_l \pi_l^{k+\delta}$$

where the last equality follows from linearity of expectation. ■

The proofs of Corollaries 2, 3 and 4 follow closely that of Corollary 1. Thus, they are omitted for the sake of brevity.

7) Proof of Corollary 5: Let m be the number of symbols collected by the receiver. Denote by $\mathcal{C}$ the generic outer code in the ensemble. Denote by $S_1(\mathcal{C}, m)$ and $S_2(\mathcal{C}, m)$ the parameters S_1 and S_2 for code $\mathcal{C}$ for a fixed number m of collected symbols. Using (4) we can write

$$\begin{aligned} \bar{\mathsf{P}}_{\mathsf{F}} &= \sum_{\mathcal{C}\in\mathscr{C}} \Pr\{\mathcal{C}\}\mathsf{P}_{\mathsf{F}}(\mathcal{C}) \geq \sum_{\mathcal{C}\in\mathscr{C}} \Pr\{\mathcal{C}\}\left[\frac{2}{r+1}S_1(\mathcal{C}, m) - \frac{2}{r(r+1)}S_2(\mathcal{C}, m)\right] \\ &= \frac{2}{r+1}\mathbb{E}_{\mathcal{C}}[S_1(\mathcal{C}, m)] - \frac{2}{r(r+1)}\mathbb{E}_{\mathcal{C}}[S_2(\mathcal{C}, m)] = \frac{2}{r+1}\bar{S}_1(m) - \frac{2}{r(r+1)}\bar{S}_2(m) \end{aligned}$$

for any $r \in \{1, \ldots, M_{q,k}\}$, where $\bar{S}_1(m)$ and $\bar{S}_2(m)$ are given by (21) and (22), respectively. Taking $r = 1$ we obtain the looser bound in (20) (i.e., $\bar{\mathsf{P}}_{\mathsf{F}} \geq \bar{S}_1(m) - \bar{S}_2(m)$). Maximization with respect to r leads us to the tighter bound in (20). (The calculation is the same as that used in [25] to obtain (5) from (4) via maximization with respect to r.)[11] ■

[11] In the extension of the upper bounds to Raptor ensembles, we expressed the number of collected symbols at the receiver as $k_{\mathcal{C}} + \delta$ for each randomly drawn outer code $\mathcal{C}$, i.e., we considered a fixed absolute overhead with respect to the outer code dimension. In the extension of the lower bounds, instead, the number of collected symbols was expressed as a fixed m for all outer codes. Note that we can also write $\bar{\mathsf{P}}_{\mathsf{F}} = \bar{\mathsf{P}}_{\mathsf{F}|k_{\mathcal{C}}=k} \Pr\{k_{\mathcal{C}} = k\} + \bar{\mathsf{P}}_{\mathsf{F}|k_{\mathcal{C}}>k} \Pr\{k_{\mathcal{C}} > k\}$. Since $\Pr\{k_{\mathcal{C}} = k\} < 1$ and $\bar{\mathsf{P}}_{\mathsf{F}|k_{\mathcal{C}}>k} < 1$, we obtain $\bar{\mathsf{P}}_{\mathsf{F}|k_{\mathcal{C}}=k} > \bar{\mathsf{P}}_{\mathsf{F}} - \Pr\{k_{\mathcal{C}} > k\}$. If $\Pr\{k_{\mathcal{C}} > k\}$ is small compared to $\bar{\mathsf{P}}_{\mathsf{F}}$ (as an example, for a linear random outer code defined by m equations we have $\Pr\{k_{\mathcal{C}} > k\} < 2^{-(h-m)}$) then Corollary 5 with $m = k + \delta$ may be regarded as an approximate lower bound for the average error probability when the outer code ensemble is expurgated from all codes with dimension larger than k.

VII. Error Exponent Analysis

In this section, we aim at deriving an error exponent analysis of Raptor code. In particular, a lower bound to the error exponent is obtained for Raptor code ensembles as a function of the outer code ensemble weight spectral shape and of the inner LT code distribution. The focus in on both binary and nonbinary Raptor codes.[12] Before proceeding with the derivation, we need to introduce a few definitions.

Following the definitions of Section II-E above, we refer to a Raptor code ensemble sequence as a sequence of Raptor code ensembles indexed by the code dimension k, where the kth Raptor code ensemble is defined by an outer code ensemble $\mathscr{C}_k$ and an inner LT code with degree distribution $\Omega(x)$, both over $\mathbb{F}_q$. To emphasize the role of the code dimension, we re-write next (19) as $\bar{\mathsf{P}}_{\mathsf{F}}^{(k)} = \mathbb{E}_{\mathcal{C}}[\mathsf{P}_{\mathsf{F}}(\mathcal{C})]$ where the average is over the outer code ensemble $\mathscr{C}_k$. For a given relative overhead $\epsilon = \delta/k$, with $\epsilon \geq 0$, the error exponent of the Raptor code ensemble sequence is

$$E(\epsilon) = \lim_{k\to\infty} -\frac{1}{k}\log_2 \bar{\mathsf{P}}_{\mathsf{F}}^{(k)}(\epsilon). \tag{32}$$

Before proceeding with the derivation of a lower bound to the error exponent for general Raptor code ensemble sequences, we illustrate the case of linear random fountain codes as an example.

Example 2. *The probability of decoding failure for a dimension-k linear random fountain code over $\mathbb{F}_q$ can be tightly upper bounded as [19]*

$$\bar{\mathsf{P}}_{\mathsf{F}}^{(k)} < \frac{1}{q-1} q^{-\epsilon k}.$$

For linear random fountain codes we hence have

$$\begin{aligned} E(\epsilon) &= \lim_{k\to\infty} -\frac{1}{k}\log_2 \bar{\mathsf{P}}_{\mathsf{F}}^{(k)}(\epsilon) \\ &> \lim_{k\to\infty} -\frac{1}{k}\log_2 \left(\frac{1}{q-1} q^{-\epsilon k}\right) \\ &= \epsilon \log_2 q. \end{aligned} \tag{33}$$

Note that (33) *is positive for positive ϵ, i.e., a positive relative overhead is sufficient to achieve an exponential (in k) decay of the decoding failure probability.*

[12] The analysis of Raptor code ensemble sequences over $\mathbb{F}_q$ with $0/1$ LT codes is omitted due to the lack of a definition of an equivalent of the weight spectral shape for (bivariate) composition enumerators.

For general Raptor code ensemble sequences, the following theorem provides a lower bound to the error exponent (under mild conditions on the outer code ensemble sequence).

Theorem 6. *Consider a Raptor code ensemble sequence over $\mathbb{F}_q$ defined by an outer code ensemble sequence $\{\mathscr{C}_k\}$ and an inner LT code degree distribution $\Omega(x)$. Let the outer code ensemble sequence spectral shape $G(\omega)$ be well-defined in $[0,1]$. If $\frac{1}{h}\log_2 \mathsf{A}^{(k)}_{\lfloor \omega h \rfloor} \overset{\mathsf{u}}{\to} G(\omega)$ then the Raptor code ensemble sequence error exponent can be lower bounded as*

$$E(\epsilon) \geq - \sup_{\omega \in (0,1]} \left[\frac{1}{R} G(\omega) + (1+\epsilon)\log_2 \varrho_\omega \right] \tag{34}$$

where $\varrho_\omega = \frac{1}{2}\sum_{j=1}^{d_{\max}} \Omega_j \left[1-(1-2\omega)^j\right]$.

Proof. For a general Raptor code ensemble sequence, we re-write the upper bound of on the decoding failure probability from Corollary 1 as

$$\bar{\mathsf{P}}_{\mathsf{F}}^{(k)} \leq \frac{1}{q-1} \sum_{\omega \in \mathcal{F}_h} \mathsf{A}^{(k)}_{\lfloor \omega h \rfloor} \pi^{k(1+\epsilon)}_{\lfloor \omega h \rfloor}$$

where $\mathcal{F}_h = \left\{ \frac{l}{h} \right\}$ with $l = 1, \ldots, h$. Following (32), we have that

$$\begin{aligned}
E(\epsilon) &= \lim_{k\to\infty} -\frac{1}{k}\log_2 \bar{\mathsf{P}}_{\mathsf{F}}^{(k)}(\epsilon) \\
&\geq \lim_{h\to\infty} -\frac{1}{hR}\log_2 \frac{1}{q-1}\sum_{\omega\in\mathcal{F}_h} \mathsf{A}^{(hR)}_{\lfloor \omega h\rfloor}\pi^{hR(1+\epsilon)}_{\lfloor \omega h\rfloor} \\
&= \lim_{h\to\infty} -\frac{1}{hR}\log_2 \frac{1}{q-1}\sum_{\omega\in\mathcal{F}_h} 2^{\log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h\rfloor} + hR(1+\epsilon)\log_2 \pi_{\lfloor \omega h\rfloor}} \\
&\geq \lim_{h\to\infty} -\frac{1}{hR}\log_2 \left\{ h \sup_{\omega\in\mathcal{F}_h} \left[2^{\log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h\rfloor} + hR(1+\epsilon)\log_2 \pi_{\lfloor \omega h\rfloor}} \right] \right\} \\
&= -\lim_{h\to\infty} \sup_{\omega\in\mathcal{F}_h} \left[\frac{1}{R}\log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h\rfloor} + (1+\epsilon)\log_2 \pi_{\lfloor \omega h\rfloor} \right] \\
&= -\lim_{h\to\infty} \sup_{\omega\in(0,1]} \left[\frac{1}{R}\log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h\rfloor} + (1+\epsilon)\log_2 \pi_{\lfloor \omega h\rfloor} \right].
\end{aligned} \tag{35}$$

If $\frac{1}{h}\log_2 \mathsf{A}^{(hR)}_{\lfloor \omega h \rfloor}$ converges uniformly to $G(\omega)$ in $[0,1]$, by observing that $\pi_{\lfloor \omega h\rfloor} \overset{\mathsf{u}}{\to} \varrho_\omega$ (see [22, Sec. III]), the order of the limit and the supremum operations in (35) can be inverted, yielding (34). ■

Remark 3. *Observe that the error exponent lower bound is monotonically increasing with ϵ. Let us assume next that, for a given Raptor code ensemble sequence, there exist an $\epsilon^\star > 0$ s.t. the right-hand side of* (34) *is strictly positive for all $\epsilon > \epsilon^\star$. We can conclude that the Raptor code* *ensemble sequence is characterized by a decoding failure probability that decays exponentially fast in k for $\epsilon > \epsilon^\star$. The value of $\epsilon^\star$ can be regarded as an upper bound on the ML decoding threshold of the Raptor code ensemble. It is important to stress that this bound on the ML decoding threshold may not be tight since it does not capture the performance in the region $\epsilon \leq \epsilon^\star$. In this latter region, the decoding failure probability may still become vanishing small as k grows large at a sub-exponential rate (e.g., only polynomially-fast in k).*

VIII. Examples of Application to Raptor Codes and Raptor Code Ensembles

In this section, we apply the results of Sections V and VII to Raptor codes and Raptor code ensembles. For the analysis, we use the LT output degree distribution employed by standard R10 Raptor codes [5], [6], given by

$$\Omega_{\mathsf{A}}(x) = 0.0098x + 0.4590x^2 + 0.2110x^3 + 0.1134x^4 + 0.1113x^{10} + 0.0799x^{11} + 0.0156x^{40}. \quad (36)$$

A. Raptor Code over $\mathbb{F}_2$ with a Hamming Outer Code

Consider a binary Raptor code over $\mathbb{F}_2$ with a Hamming outer code. The weight enumerator of a binary Hamming code of length $h = 2^t - 1$ and dimension $k = h - t$ can be derived easily using the recursion $(i+1)\, A_{i+1} + A_i + (h - i + 1)\, A_{i-1} = \binom{h}{i}$ with $A_0 = 1$ and $A_1 = 0$ [29]. The weight distribution obtained from this recursion can then be incorporated in Theorem 1 to derive the corresponding upper bound on the failure probability. The lower bounds established by Theorem 5 (binary case) can also be derived, by employing the Hamming code biweight enumerator, an expression of which was developed in [29].

Fig. 3 shows the decoding failure rate for a Raptor code over $\mathbb{F}_2$ employing a $(63, 57)$ binary Hamming outer code as a function of the absolute overhead, δ, together with the upper bound from Theorem 1 and the lower bounds from Theorem 5 (binary case). In order to obtain the values of failure rate, Monte Carlo simulations were run for each δ until 200 errors were collected using inactivation decoding. It can be observed how the upper bound is very tight and how the gap between the upper and lower bounds is very small already for values of δ in the order

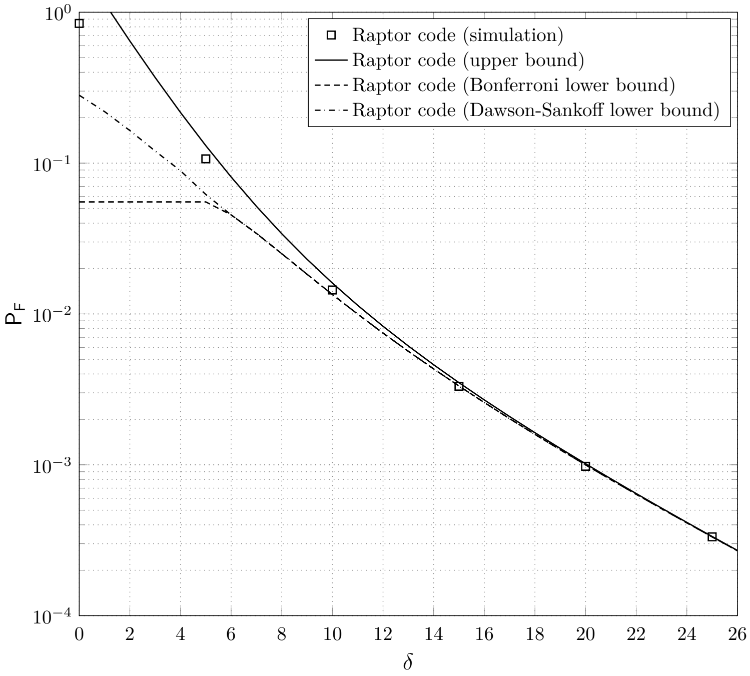

Fig. 3. Decoding failure probability $\mathsf{P_F}$ versus the absolute overhead δ for a binary Raptor code with a $(63, 57)$ Hamming outer code and LT distribution $\Omega_{\mathsf{A}}(x)$. Markers: simulation results. Solid: upper bound (Theorem 1). Dotted: Degree-two Bonferroni lower bound (Theorem 5). Dot-dashed: Dawson-Sankoff lower bound (Theorem 5).

of 10. Interestingly, the order-two Bonferroni and the Dawson-Sankoff bounds are practically coincident for $\delta \geq 6$ while for $\delta < 6$ the Dawson-Sankoff bound turns to be remarkably tighter.[13]

B. Raptor Code Ensembles with Linear Random Outer Codes

Next, consider a Raptor code ensemble over $\mathbb{F}_q$, with LT degree distribution $\Omega_{\mathsf{A}}(x)$ and in which the outer code is picked from the uniform parity-check ensemble, with parity-check matrix of size $(h-k)\times h$ and characterized by i.i.d. entries with uniform distribution in $\mathbb{F}_q$. The expected weight enumerator for an outer code drawn randomly in $\mathscr{C}$ is known to be $\mathsf{A}_l = \binom{h}{l} q^{-(h-k)}(q-1)^l$. The expected composition enumerator can be obtained from the expected weight enumerator, as discussed in Appendix C, while the expected bicomposition enumerator can be obtained as shown in Appendix D.

[13] The difference $S_1 - S_2$ is actually increasing for $\delta \in \{0, \dots, 5\}$, it reaches a maximum at $\delta = 5$ and then decreases. For $\delta \in \{0, \dots, 4\}$ the difference is even negative. However, since the failure probability cannot increase as δ increases, we can apply the value taken by $S_1 - S_2$ at $\delta = 5$ to all $\delta < 5$. In contrast, Dawson-Sankoff bound decreases monotonically over the whole range of δ.

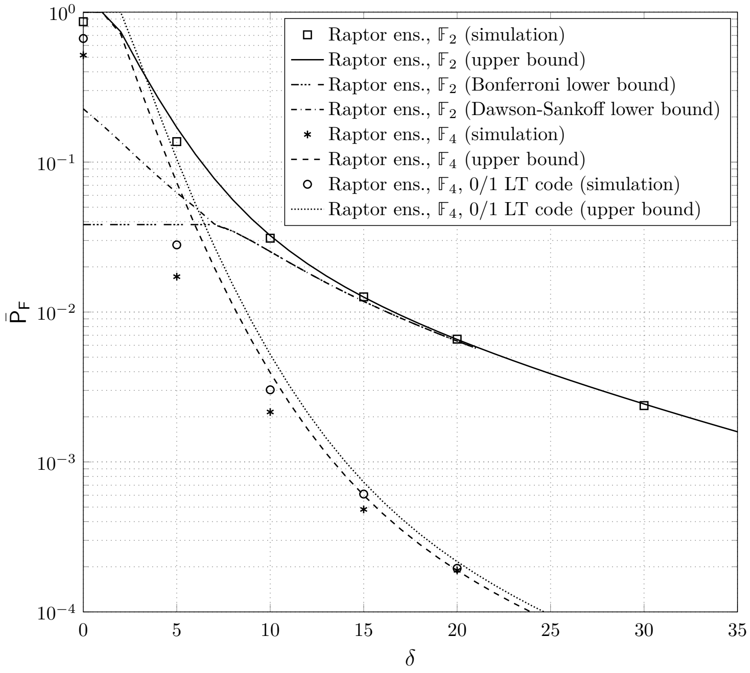

Fig. 4. Expected probability of decoding failure $\bar{\mathsf{P}}_{\mathsf{F}}$ vs absolute overhead for Raptor code ensembles where the outer code is drawn randomly from the uniform parity-check ensemble with $k = 64$ and $h = 70$. LT distribution: $\Omega_{\mathsf{A}}(x)$. Lines: upper and lower bounds. Markers: simulation results.

To obtain the experimental values of the expected decoding failure rate, 6000 different outer codes were generated. For each outer code and for each overhead value, 1000 inactivation decoding attempts were carried out. The average failure rate was calculated by averaging the failure rates of the individual Raptor codes. To generate an outer code, an $(h-k)\times h$ parity-check matrix was drawn randomly by picking its elements independently and uniformly in $\mathbb{F}_q$.

In Fig. 4 we show simulation results for $k = 64$ and $h = 70$. Three different Raptor code ensembles were considered, one constructed over $\mathbb{F}_2$, one constructed over $\mathbb{F}_4$, and one constructed over $\mathbb{F}_4$ with a $0/1$ LT code. We can observe how in all cases the upper bounds are tight, even for small values of δ. Comparing the two ensembles over $\mathbb{F}_4$, it is remarkable that employing a $0/1$ LT code results only in a small performance degradation, which vanishes as δ increases. Both order-two Bonferroni and Dawson-Sankoff lower bounds are displayed for the binary ensemble. Again, the Dawson-Sankoff bound turns out to be remarkably tighter for small δ.

In Fig. 5 lower bounds on the error exponents of various binary Raptor code ensemble sequences are provided. The Raptor code ensemble sequences are defined by the degree distribution $\Omega_{\mathsf{A}}(x)$ and linear random outer code sequences with (outer) code rates $R = 0.90, 0.95$ and 0.98.

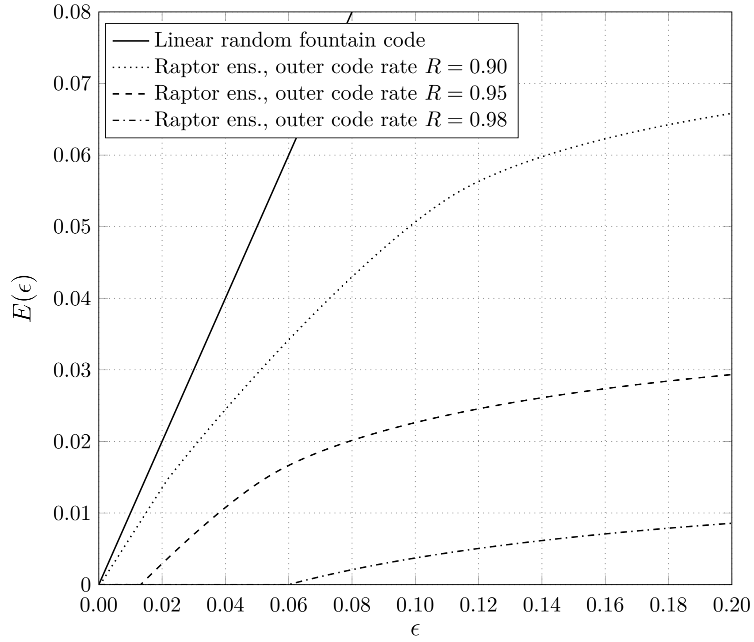

Fig. 5. Lower bounds on the error exponent vs. relative overhead ϵ for binary Raptor code ensemble sequences defined by the degree distribution $\Omega_{\mathsf{A}}(x)$ and linear random outer code sequences with (outer) code rates $R = 0.90, 0.95$ and 0.98. The error exponent lower bound for linear random fountain codes of (33) is provided as reference.

When the outer code is picked from a binary linear random code ensemble, the error exponent lower bound of (34) reduces to

$$E(\epsilon) \geq - \sup_{\omega \in (0,1]} \left[\frac{H_b(\omega) + R - 1}{R} + (1+\epsilon) \log_2 \varrho_\omega \right]$$

where $H_b(\omega) = -\omega \log_2 \omega - (1-\omega) \log_2 (1-\omega)$ is the binary entropy function. The error exponent lower bound for linear random fountain codes of (33) is provided as a reference. As intuition suggests, the error exponent lower bound for Raptor codes approaches the one of linear random fountain codes as the outer code rate decreases. The upper bounds on the ML decoding thresholds are $\epsilon^\star \approx 6 \times 10^{-2}$ for $R = 0.98$, $\epsilon^\star \approx 1.33 \times 10^{-2}$ for $R = 0.95$, and $\epsilon^\star \approx 5 \times 10^{-4}$ for $R = 0.90$.

C. Raptor Code Ensembles with Regular LDPC Outer Codes

We now consider ensembles of Raptor codes in which the outer code is drawn from a (d_v, d_c) regular low-density parity-check (LDPC) code ensemble, where d_v and d_c are the variable and check node degrees, respectively. In order to draw a code from this ensemble we first generate

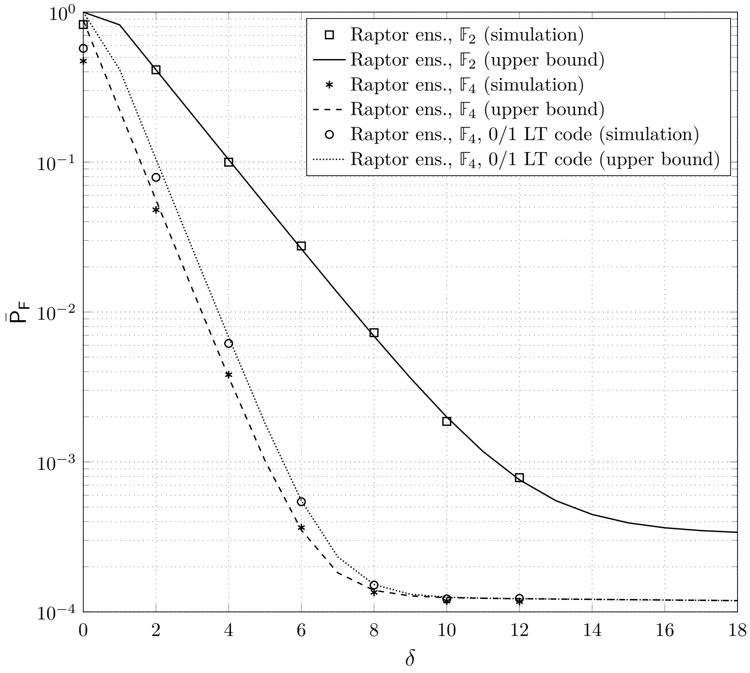

Fig. 6. Average probability of decoding failure P_F vs absolute overhead for two Raptor code ensembles where the outer code is randomly drawn from the $(d_v = 3, d_c = 15)$ regular LDPC ensemble with $k = 1000$ input symbols and $h = 1250$ intermediate symbols. LT distribution: $\Omega_\mathsf{A}(x)$. Lines: upper bounds. Markers: simulation results.

a random permutation of the $hd_v = (h - k)d_c$ edges between check and variable nodes. Then we assign to each edge a non-binary label picked uniformly at random in $\mathbb{F}_q \backslash \{0\}$. The average weight enumerator for this ensemble is reviewed in Appendix C, where an expression of its expected composition enumerator is also derived.

In order to simulate the average probability of decoding failure of the ensemble, 10000 different outer codes were generated. For each outer code and overhead value, 100 decoding attempts were carried out. The average probability of decoding failure was obtained averaging the probabilities of decoding failure obtained with the different outer codes.

Fig. 6 shows the average probability of decoding failure for three ensembles of Raptor codes where the outer code is randomly drawn from the $(d_v = 3, d_c = 15)$ regular LDPC ensemble with $k = 1000$ input symbols and $h = 1250$ intermediate symbols. The first ensemble is constructed over $\mathbb{F}_2$, the second over $\mathbb{F}_4$ and the third is also constructed over $\mathbb{F}_4$ but with a $0/1$ LT code. It can be observed how the upper bounds are very tight. Furthermore, as δ increases the performance of the ensemble with a $0/1$ LT code quickly converges to that of the ordinary ensemble over $\mathbb{F}_4$.

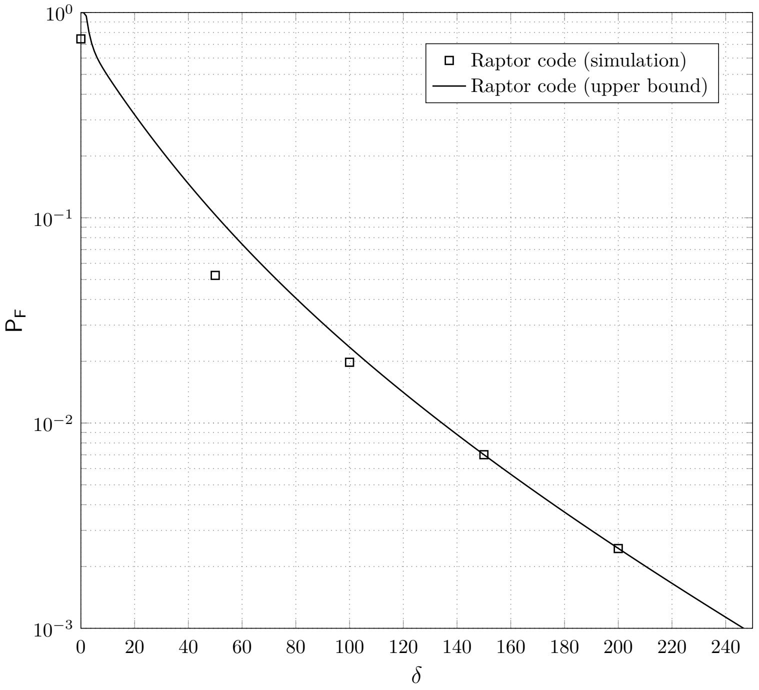

Fig. 7. Decoding failure probability $\mathsf{P_F}$ vs absolute overhead for a multi-edge type Raptor code where the outer code is a $(1023, 1013)$ Hamming code with $h_A = 900$ and $h_B = 123$. LT distribution: $\Omega_{\mathsf{A}}(x)(z^2 + z^3)/2$. Line: upper bound. Markers: simulation results.

D. Multi-Edge Type Raptor Code Ensembles

Next we consider multi-edge type Raptor codes with a bivariate LT output degree distribution given by $\Omega_{\mathsf{A}}(x)\left(z^2 + z^3\right)/2$.[14]

We consider first a multi-edge type Raptor code over $\mathbb{F}_2$ where the outer code is a $(1023, 1013)$ Hamming code, with $h_A = 900$ intermediate symbols of type A and $h_B = 123$ intermediate symbols of type B. In order to obtain the bivariate weight enumerator of the Hamming code, the bivariate weight enumerator of the dual code was first obtained by enumerating all its codewords. Then, the extension of the MacWilliams identity developed in Appendix B was applied. Fig. 7 shows the average decoding failure probability, as well as its upper bound. It can be observed how the upper bound is tight.

[14] This degree distribution is inspired by the one used in RaptorQ codes [24], where for type A intermediate symbols (called LT symbols in [24]) a conventional LT output degree distribution is used, whereas for type B intermediate symbols (referred to as permanently inactivated symbols in [24]) degrees 2 and 3 are chosen with probability $1/2$. See [24] for more details.

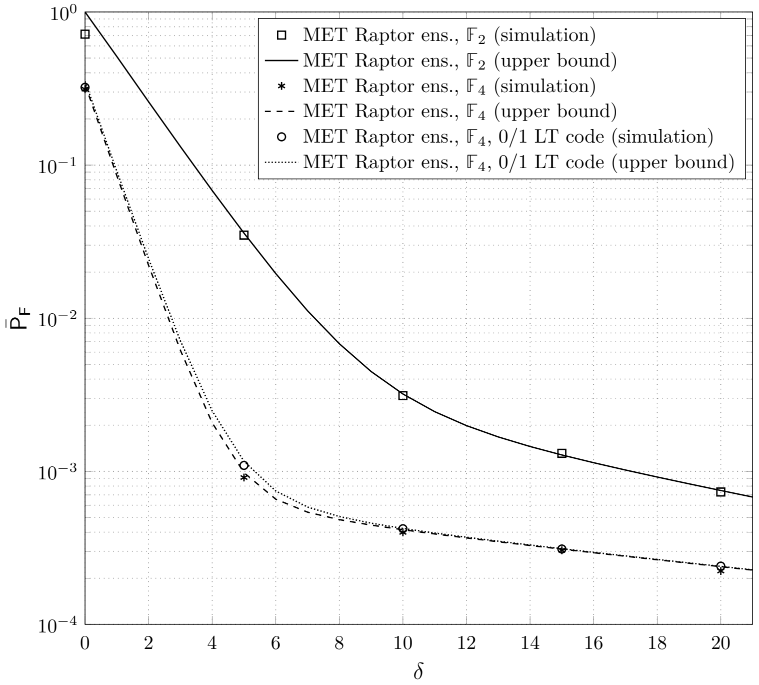

Fig. 8. Average probability of decoding failure $\bar{\mathsf{P}}_\mathsf{F}$ vs absolute overhead for three multi-edge type Raptor code ensembles where the outer code is randomly drawn from the $(5, 55)$ regular LDPC ensemble with $k = 100$ and $h = 110$, with $h_A = 100$ and $h_B = 10$. LT distribution: $\Omega_\mathsf{A}(x)(z^2 + z^3)/2$. Lines: upper bounds. Markers: simulation results.

Next, we consider multi-edge type Raptor code ensembles where the outer code is again drawn from the (d_v, d_c) regular LDPC code ensemble. In particular, the outer code is randomly drawn from the $(5, 55)$ regular LDPC ensemble with $k = 100$ input symbols and $h = 110$ intermediate symbols. Out of the 110 intermediate symbols, 100 are of class A and 10 of class B. The average bivariate weight enumerator for this ensemble is given by

$$\mathsf{A}_{a,b} = \frac{\binom{h_A}{a}\binom{h_B}{b}}{\binom{h}{a+b}} \mathsf{A}_{a+b}.$$

from which the average bivariate composition enumerator can be obtained through Proposition 4 in Appendix C.

Fig. 8 shows the average probability of decoding failure for three ensembles of multi-edge type Raptor codes, one constructed over $\mathbb{F}_2$, another over $\mathbb{F}_4$, and a third one also constructed over $\mathbb{F}_4$ but with a $0/1$ LT code. It can be observed how the upper bounds are very tight in this case too. If we compare the the probability of failure of the two ensembles built over $\mathbb{F}_4$, we can see how their performance is almost the same. It is remarkable how restricting the LT code to use only binary labels does not result in an appreciable performance loss.

IX. Code Design Examples

In this section we provide several code design examples that illustrate the practical impact of the derived bounds.

A. Design of a Binary Raptor code with an LDPC Outer Code

We consider the case in which the outer code ensemble is given and run a computer search in order to find an LT output degree distribution that optimizes a given metric subject to some design constraints. In particular, we consider Raptor code ensembles where the outer code is picked from the $(d_v = 3, d_c = 33)$ binary regular LDPC ensemble with $k = 1000$ and $h = 1100$, and we set as requirement minimizing the inactivation decoding complexity subject to a decoding failure probability not exceeding 10^{-3}.

Inactivation decoding [11] is the efficient ML decoding algorithm used to decode standardized Raptor codes [5], [6]. It can be seen as an extension of iterative (peeling) decoding where, whenever the iterative decoding process stops, an input symbol is declared as inactive, so that iterative decoding is resumed. At the end, one is left with a number of input symbols that have been inactivated, and whose values have to be recovered by means of Gaussian elimination. After doing so, all input symbols can be resolved by back-substitution (i.e., using iterative decoding). The complexity of inactivation decoding is generally dominated by the Gaussian elimination step, whose complexity is cubic on the number of inactivations. Thus, minimizing the number of inactivations can be used as a proxy for minimizing the decoding complexity.

The degree distribution Ω_{A}, given in (36), has been designed for inactivation decoding. However, as it can be observed in Fig. 9, if we use Ω_{A} we do not fulfill the probability of failure constraint, since there is an error floor around 2×10^{-3}. Thus, we need carry out an ad-hoc design.

The analysis presented in [15] can be used to determine the expected number of inactivations for LT codes. Extending the analysis to Raptor codes is not easy, but, as it was shown in [15], when the parity-check matrix of the outer code is considerably denser than the generator matrix of the inner LT code, it is possible to design Raptor codes that require few inactivations by optimizing the LT output degree distribution in isolation.[15] In other words, if we design an LT

[15] Note that this heuristic observation holds true also for the case where the outer code parity-check matrix is not dense, e.g., to the case where the outer code is an LDPC code, provided that the average check node degree of the LDPC code is considerably larger than the average output degree of the LT code.

degree distribution that requires few inactivations, and then construct a Raptor code using this degree distribution for the inner LT code, we obtain a Raptor code that requires few inactivations.

Following this approach, we can use simulated annealing [35] to design an LT degree distribution that minimizes the number of inactivations for the LT code in isolation, under the constraint on the decoding failure probability for the resulting Raptor code, estimated using the upper bounds derived in this paper. By using this approach we obtained the following degree distribution

$$\Omega_{\mathsf{B}} = 0.0108x + 0.4557x^2 + 0.1959x^3 + 0.1195x^4 + 0.0245x^5 + 0.0243x^6 + 0.0357x^{10} \\ + 0.0412x^{11} + 0.0440x^{15} + 0.0196x^{21} + 0.0115x^{26} + 0.0088x^{30} + 0.0085x^{40}.$$

Fig. 9 shows the average probability of decoding failure and its upper bound in Corollary 1 for the designed ensemble based on Ω_{B}. We can observe how the Raptor code ensemble meets the design requirement, since $\bar{\mathsf{P}}_{\mathsf{F}} < 10^{-3}$ for $\delta = 15$.

If we now consider the number of inactivations, we have that the designed Raptor code ensemble, which employs Ω_{B}, needs in average 94 inactivations for an absolute overhead $\delta = 15$. In constrast, the Raptor code ensemble employing Ω_{A} needs 87. This confirms how a reduction in the number of inactivations forces the failure rate to jump above the maximum tolerable value.

B. Design of a Nonbinary Raptor code with an LDPC Outer Code

This design example is similar to the previous one, but this time we focus on a nonbinary Raptor code ensemble. In particular we aim at designing a Raptor code ensemble over $\mathbb{F}_4$, where the outer code is taken from the $(d_v = 3, d_c = 63)$ regular LDPC ensemble with $k = 200$ and $h = 210$. The goal is minimizing the number of inactivations[16] subject to $\bar{\mathsf{P}}_{\mathsf{F}} \leq 10^{-4}$ at $\delta = 10$. Using simulated annealing, the following degree distribution is obtained:

$$\Omega_{\mathsf{C}}(x) = 0.0214x + 0.3213x^2 + 0.2971x^3 + 0.0276x^4 + 0.0252x^5 + 0.0418x^9 + 0.0458x^{13} \\ + 0.0654x^{18} + 0.0457x^{23} + 0.0612x^{30} + 0.0295x^{35} + 0.0180x^{40}.$$

[16] The analysis in [15] is also valid for non-binary codes. The number of inactivations is a product of the first phase of inactivation decoding, triangulation, which is equivalent to column and row swapping and does not carry out any operations over the finite field. Thus, the number of inactivations only depends on the elements of the generator matrix of the LT code being zero or nonzero, and not on the particular value in $\mathbb{F}_q \setminus \{0\}$ that the elements take.

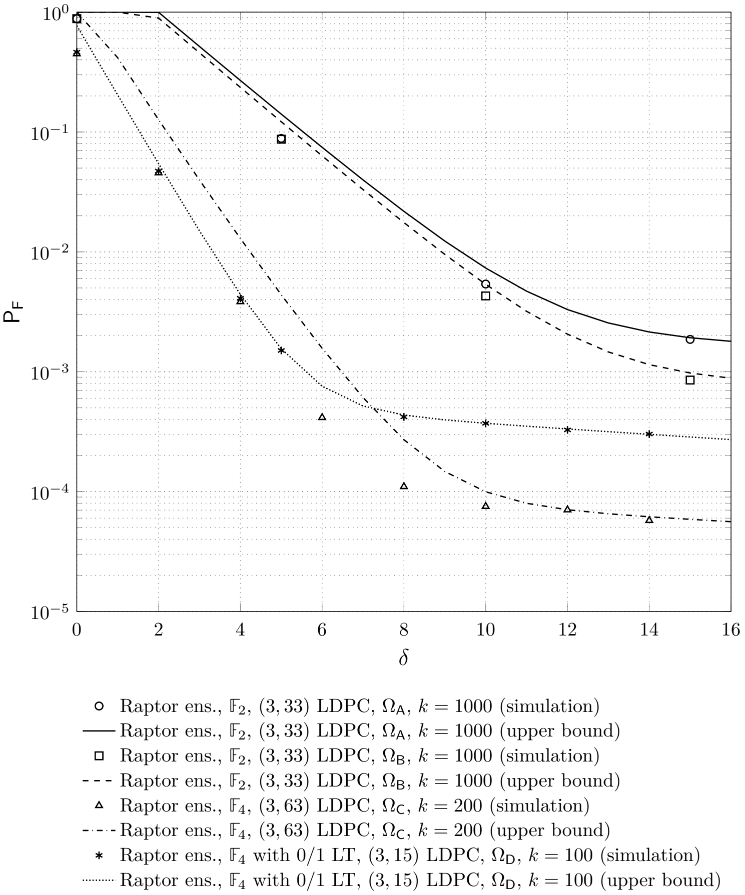

Fig. 9. Average probability of decoding failure $\bar{\mathsf{P}}_\mathsf{F}$ vs absolute overhead for 4 different Raptor code ensembles. The first and second ensemble have outer codes randomly drawn from the binary $(d_v = 3, d_c = 33)$ regular LDPC ensemble with $k = 1000$ input symbols. For the second and third ensembles the outer codes are randomly drawn from the $(d_v = 3, d_c = 63)$ regular LDPC ensemble with $k = 200$ and the $(d_v = 3, d_c = 15)$ regular LDPC ensemble with $k = 100$. The LT degree distributions are Ω_A, Ω_B, Ω_C and Ω_D, respectively. Lines: upper bounds. Markers: simulation results.

Fig. 9 shows the average probability of decoding failure for the ensemble obtained from the code design. We can observe how the constraint on $\bar{\mathsf{P}}_\mathsf{F}$ is fulfilled. The average number of inactivation needed for decoding at $\delta = 10$ is approximately 32.

C. Design of a Raptor Code with a $0/1$ LT Code

We now address the design of a nonbinary Raptor code ensemble with a $0/1$ LT code. We aim at designing a Raptor code ensemble over $\mathbb{F}_4$, where the outer code is taken from the $(d_v = 3, d_c = 15)$ regular LDPC ensemble with $k = 100$ and $h = 125$. The goal is minimizing the number of inactivations subject to $\bar{\mathsf{P}}_{\mathsf{F}} \leq 2 \times 10^{-3}$ at $\delta = 5$. Using simulated annealing, the following degree distribution is obtained:

$$\Omega_{\mathsf{D}}(x) = 0.0095x + 0.3896x^2 + 0.3159x^3 + 0.0843x^4 + 0.0611x^{10} + 0.0585x^{15} + 0.0811x^{22}.$$

Fig. 9 shows the average probability of decoding failure for the designed ensemble. We can observe how the constraint on $\bar{\mathsf{P}}_{\mathsf{F}}$ is fulfilled. The average number of inactivations needed for decoding at $\delta = 5$ is approximately 22.

X. Conclusions

In this paper we have considered different Raptor code constructions over $\mathbb{F}_q$ under ML decoding, deriving tight upper and lower bounds to the probability of decoding failure. The bounds are first derived for Raptor codes with a deterministic outer code, and then they are extended to Raptor code ensembles in which the outer code is drawn at random from an ensemble of linear block codes. In all cases the upper bounds require the knowledge of the weight enumerator of the outer code (ensemble) or its composition enumerator, whereas the lower bounds require the knowledge of the joint weight/composition enumerators of the outer code (ensemble). By means of extensive simulations we have illustrated how the bounds presented in this paper are tight. A framework for the analysis of the error exponent of Raptor code ensemble sequences is introduced, which allows deriving a lower bound on the error exponent. The result allows gaining further insights on the performance of Raptor code ensemble sequences, by identifying relative overhead regions where an exponential (in the input block size) decay of the error probability can be achieved. The work is completed by selected examples of Raptor code design based on the bounds derived in this paper. To the best of the authors' knowledge, this is the first work which considers Raptor codes with a generic q-ary outer code. An open question relates to the concentration properties of Raptor code ensembles.

Appendix A
Sum of Random Uniform Variables in $\mathbb{F}_{2^m}\backslash\{0\}$

The following lemma is used in the proof of Theorem 1.

Lemma 4. *Let X_1, X_2 ... X_l be discrete i.i.d random variables uniformly distributed over $\mathbb{F}_{2^m}\backslash\{0\}$. Then*

$$\Pr\{X_1 + X_2 + \ldots + X_l = 0\} = \frac{1}{q}\left(1 + \frac{(-1)^l}{(q-1)^{l-1}}\right)$$

where $q = 2^m$.

Proof. Observe that the additive group of $\mathbb{F}_{2^m}$ is isomorphic to the vector space $\mathbb{Z}_2^m$. Thus, we may let X_1, X_2 ... X_l be i.i.d random variables with uniform probability mass function over the vector space $\mathbb{Z}_2^m \backslash \{0\}$.

Let us introduce the auxiliary random variable $W = X_1 + X_2 + \ldots + X_l$ and let us denote by $P_W(w)$ and by $P_X(x)$ the probability mass functions of W and X_i, respectively, where

$$P_X(x) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{1}{q-1} & \text{otherwise.} \end{cases}$$

Due to independence we have $P_W = P_X * P_X * \ldots * P_X$ which, taking the m-dimensional two-points discrete Fourier transform (DFT) $\mathscr{J}\{\cdot\}$ of both sides, yields $\mathscr{J}\{P_W(w)\} = (\mathscr{J}\{P_X(x)\})^l$. Next, since

$$\hat{P}_X(t) = \mathscr{J}\{P_X(x)\} = \begin{cases} 1 & \text{if } t = 0 \\ \frac{-1}{q-1} & \text{otherwise} \end{cases}$$

we have

$$\hat{P}_W(t) = \mathscr{J}\{P_W(w)\} = \begin{cases} 1 & \text{if } t = 0 \\ \frac{(-1)^l}{(q-1)^l} & \text{otherwise.} \end{cases}$$

We are interested in $P_W(0)$ whose expression corresponds to

$$P_W(0) = \frac{1}{q} \sum_t \hat{P}_W(t) = \frac{1}{q} + \frac{1}{q}(q-1)\frac{(-1)^l}{(q-1)^l}$$

from which the statement follows. ■

The result in this lemma appears in [17]. However, the proof in [17] uses a different approach based on a known result on the number of closed walks of length l in a complete graph of size q from a fixed but arbitrary vertex back to itself.

APPENDIX B
AN EXTENSION OF THE MACWILLIAMS IDENTITY

Consider a linear block code $\mathcal{C} \subset \mathbb{F}_q^h$. The same way we defined its bivariate weight enumerator in (6), we can define its h-variate enumerator polynomial as

$$A(x_1, \ldots, x_h) = \sum_{i_1=0}^{1} \cdots \sum_{i_h=0}^{1} A_{i_1,\ldots,i_h} \prod_{j=1}^{h} x_j^{i_j}$$

where $A_{i_1,\ldots,i_h}$ denotes the multiplicity of codewords with $w(v_1) = i_1$, $w(v_2) = i_2$, ... and $w(v_h) = i_h$, i.e., the number of codewords with support $(i_i, i_2, \ldots, i_h)$. The following proposition establishes an extension of the MacWilliams identity for h-variate weight enumerators.

Proposition 1. *Let $\mathcal{C}$ be an (h, k) linear block code over $\mathbb{F}_q$ with h-variate weight enumerator $A(x_1, \ldots, x_h)$. Let $\mathcal{C}^\perp$ be the dual of $\mathcal{C}$ and denote its h-variate weight enumerator by $B(x_1, \ldots, x_h)$. Then*

$$B(x_1, \ldots, x_h) = q^{-k} \prod_{i=1}^{h} (1 + (q-1)x_i)\, A\left(\frac{1 - x_1}{1 + (q-1)x_1} \ldots, \frac{1 - x_h}{1 + (q-1)x_h}\right).$$

Proof. The proof builds on that that of the MacWilliams identity for linear block codes over $\mathbb{F}_q$ [36]. We start by rewriting $A(x_1, \ldots, x_h)$ as

$$A(x_1, \ldots, x_h) = \sum_{\mathbf{v}\in\mathcal{C}} \prod_{i=1}^{h} {x_i}^{w(v_i)}$$

Let us now define function $g(\mathbf{u})$ as follows

$$g(\mathbf{u}) = \sum_{\mathbf{v}\in\mathbb{F}_q^h} \chi\left(\langle \mathbf{u}, \mathbf{v}\rangle\right) \prod_{i=1}^{h} {x_i}^{w(v_i)}$$

where χ is a non-trivial character of $(\mathbb{F}_q, +)$.

We have

$$\begin{aligned}
\sum_{\mathbf{u}\in\mathcal{C}} g(\mathbf{u}) &= \sum_{\mathbf{u}\in\mathcal{C}} \sum_{\mathbf{v}\in\mathbb{F}_q^h} \chi\left(\langle \mathbf{u}, \mathbf{v}\rangle\right) \prod_{i=1}^{h} {x_i}^{w(v_i)} = \sum_{\mathbf{v}\in\mathbb{F}_q^h} \prod_{i=1}^{h} {x_i}^{w(v_i)} \sum_{\mathbf{u}\in\mathcal{C}} \chi\left(\langle \mathbf{u}, \mathbf{v}\rangle\right) \qquad (37)\\
&= \sum_{\mathbf{v}\in\mathcal{C}^\perp} \prod_{i=1}^{h} {x_i}^{w(v_i)} \sum_{\mathbf{u}\in\mathcal{C}} \chi\left(\langle \mathbf{u}, \mathbf{v}\rangle\right) + \sum_{\mathbf{v}\notin\mathcal{C}^\perp} \prod_{i=1}^{h} {x_i}^{w(v_i)} \sum_{\mathbf{u}\in\mathcal{C}} \chi\left(\langle \mathbf{u}, \mathbf{v}\rangle\right)\\
&= \sum_{\mathbf{v}\in\mathcal{C}^\perp} \prod_{i=1}^{h} {x_i}^{w(v_i)} \sum_{\mathbf{u}\in\mathcal{C}} \chi\left(0\right) = \sum_{\mathbf{v}\in\mathcal{C}^\perp} \prod_{i=1}^{h} {x_i}^{w(v_i)} \left|\mathcal{C}\right|\\
&= \left|\mathcal{C}\right| B(x_1, \ldots, x_n)
\end{aligned}$$

Let us now rewrite $g(\mathbf{u})$ as follows

$$\begin{aligned}
g(\mathbf{u}) &= \sum_{\mathbf{v}\in\mathbb{F}_q^h}\prod_{i=1}^{h} {x_i}^{w(v_i)}\,\chi\left(u_1v_1+\ldots+u_hv_h\right)\\
&= \sum_{\mathbf{v}\in\mathbb{F}_q^h}\prod_{i=1}^{h} {x_i}^{w(v_i)}\chi\left(u_iv_i\right)\\
&= \prod_{i=1}^{h}\sum_{v\in\mathbb{F}_q} {x_i}^{w(v)}\chi\left(u_iv\right)
\end{aligned}$$

Let us now look at the inner summation, we have

$$\sum_{v\in\mathbb{F}_q} {x_i}^{w(v)}\chi\left(u_iv\right) = \begin{cases} 1+(q-1)x_i, & \text{if } u_i=0\\ 1+x_i\sum_{\alpha\in\mathbb{F}_q\setminus\{0\}}\chi\left(\alpha\right)=1-x, & \text{otherwise.}\end{cases}$$

Thus, we can write

$$g(\mathbf{u}) = \prod_{i=1}^{h}(1-x_i)^{w(v_i)}\left(1+(q-1)x_i\right)^{1-w(v_i)} \tag{38}$$

Finally, if we replace (38) into (37) we obtain

$$\begin{aligned}
B(x,z) &= \frac{1}{|\mathcal{C}|}\sum_{\mathbf{u}\in\mathcal{C}} g(\mathbf{u})\\
&= \frac{1}{|\mathcal{C}|}\sum_{\mathbf{u}\in\mathcal{C}}\prod_{i=1}^{h}(1-x_i)^{w(v_i)}\left(1+(q-1)x_i\right)^{1-w(v_i)}\\
&= q^{-k}\prod_{i=1}^{h}\left(1+(q-1)x_i\right)A\left(\frac{1-x_1}{1+(q-1)x_1}\ \ldots,\frac{1-x_h}{1+(q-1)x_h}\right)
\end{aligned}$$

■

The result in Proposition 1 is strongly related to the result derived in [37, Appendix], where a similar analysis is used to derive a maximum-a-posteriori decoding algorithm for a code based on its dual. However, for the sake of completeness, we decided to include the result in the form of a Theorem with its corresponding proof.

Now that we have a MacWilliams identity for h-variate weight enumerators it is easy to derive a similar result for bi-variate weight enumerators.

Proposition 2. *Let $\mathcal{C}$ be an (h,k) linear block code over $\mathbb{F}_q$ in which the h codeword symbols are divided into h_A symbols of class A and $h_B = h - h_A$ of class B, with bivariate weight* *enumerator of $A(x,z)$. Let $\mathcal{C}^{\perp}$ be the dual of $\mathcal{C}$ and denote its bivariate weight enumerator by $B(x,z)$. Then*

$$B(x,z) = q^{-k}\left(1+(q-1)x\right)^{h_A}\left(1+(q-1)z\right)^{h_B} A\left(\frac{1-x}{1+(q-1)x}, \frac{1-z}{1+(q-1)z}\right).$$

Proof. We just need to introduce the variable changes $x_i = x$ for $i = 1, \ldots, h_A$ and $x_i = z$ for $i = h_A + 1, \ldots, h$ in Proposition 1. ■

Note that the special case of Proposition 2 for $h_B = h_A$ is proposed in [29, Chapter 5.6] as an exercise.

APPENDIX C
AVERAGE COMPOSITION ENUMERATORS OF SOME CODES ENSEMBLES

This appendix provides results on the average composition enumerator of some code ensembles. The following proposition states that, in some cases, the average composition enumerator can be easily derived from the average weight enumerator.

Proposition 3. *Consider an ensemble $\mathscr{C}$ of linear block codes, all with block length h, along with a probability measure on each such code. Let A_l be the expected weight enumerator of a random code $\mathcal{C} \in \mathscr{C}$. Assume that $\Pr\{\mathbf{v} \in \mathcal{C} | \varsigma(\mathbf{v}) = \mathbf{f}\} = \Pr\left\{\mathbf{v} \in \mathcal{C} | w(\mathbf{v}) = \sum_{i=1}^{q-1} f_i\right\}$ for all $\mathbf{v} \in \mathbb{F}_q^h$. Then*

$$\mathsf{Q}_{\mathbf{f}} = \mathsf{A}_l \binom{l}{f_1, f_2, \ldots, f_{q-1}} (q-1)^{-l} \qquad (39)$$

where $l = \sum_{i=1}^{q-1} f_i$.

Proof: We can express $\mathsf{Q}_{\mathbf{f}}$ as the number of vectors of composition $\mathbf{f}$ times the probability that each such vector is a codeword. Letting $l = \sum_{i=1}^{q-1} f_i = w(\mathbf{v})$ we can write

$$\begin{aligned}\mathsf{Q}_{\mathbf{f}} &= \binom{h}{\mathbf{f}} \Pr\{\mathbf{v} \in \mathcal{C} | \varsigma(\mathbf{v}) = \mathbf{f}\} = \binom{h}{\mathbf{f}} \Pr\{\mathbf{v} \in \mathcal{C} | w(\mathbf{v}) = l\} \\ &= \binom{h}{\mathbf{f}} \frac{\mathsf{A}_l}{\binom{h}{l}(q-1)^l}\end{aligned}$$

The last obtained expression yields (39) by applying the identity $\binom{h}{\mathbf{f}} = \binom{h}{l}\binom{l}{f_1, f_2, \ldots, f_{q-1}}$. ■

Examples of ensembles for which the assumption on Proposition 3 holds are the uniform parity-check ensemble and the (regular and irregular) LDPC code ensembles.

1) Uniform parity-check ensemble: For a uniform parity-check ensemble defined by a random parity-check matrix of size $(h-k)\times h$ with i.i.d. entries uniformly distributed in $\mathbb{F}_q$ we have $\mathsf{A}_l = \binom{h}{l}(q-1)^l q^{-(h-k)}$ and therefore (39) leads to

$$\mathsf{Q}_{\mathbf{f}} = \binom{h}{\mathbf{f}} q^{-(h-k)}.$$

2) Regular LDPC ensemble: Consider a (d_v, d_c) regular LDPC code ensemble of length h, where d_v and d_c are the variable and check node degrees, respectively. The ensemble is defined by all possible permutations of the $hd_v = (h-k)d_c$ edges between check and variable node sockets and by all possible ways to label the edges with nonzero symbols. Each edge permutation is picked with uniform probability and the label of each edge is drawn uniformly at random in $\mathbb{F}_q\backslash\{0\}$. The average weight enumerator for this ensemble is given by [38], [39]

$$\mathsf{A}_l = \binom{h}{l}\frac{\text{coeff}\left(p(x)^{h\,d_v/d_c}, x^{l\,d_v}\right)}{\binom{h\,d_v}{l\,d_v}(q-1)^{l(d_v-1)}}$$

where $p(x) = \frac{1}{q}\left(1+(q-1)x\right)^{d_c} + \frac{q-1}{q}(1-x)^{d_c}$. Hence, applying (39) we obtain

$$\mathsf{Q}_{\mathbf{f}} = \binom{h}{\mathbf{f}}\frac{\text{coeff}\left(p(x)^{h\,d_v/d_c}, x^{ld_v}\right)}{\binom{h\,d_v}{l\,d_v}}(q-1)^{-ld_v}$$

Proposition 3 can be extended to bivariate enumerators using the same proof argument.

Proposition 4. *Consider an ensemble $\mathscr{C}$ of linear block codes, all with block length $h = h_A + h_B$, along with a probability measure on each such code. Let $\mathsf{A}_{l,s}$ be the expected bivariate weight enumerator of a random code $\mathcal{C} \in \mathscr{C}$. Assume that* $\Pr\{\mathbf{v} \in \mathcal{C} | \varsigma(\mathbf{v}_A) = \mathbf{f}_A, \varsigma(\mathbf{v}_B) = \mathbf{f}_B\} = \Pr\left\{\mathbf{v} \in \mathcal{C} | w(\mathbf{v}_A) = \sum_{i=1}^{q-1} f_{A,i}, w(\mathbf{v}_B) = \sum_{i=1}^{q-1} f_{B,i}\right\}$ *for all* $\mathbf{v} = (\mathbf{v}_A, \mathbf{v}_B) \in \mathbb{F}_q^h$. *Then*

$$\mathsf{Q}_{\mathbf{f}_A,\mathbf{f}_B} = \mathsf{A}_{l,s}\binom{l}{f_{A,1}, f_{A,2}, \ldots, f_{A,q-1}}(q-1)^{-l}\binom{s}{f_{B,1}, f_{B,2}, \ldots, f_{B,q-1}}(q-1)^{-s}$$

where $\sum_{i=1}^{q-1} l = f_{A,i}$ *and* $s = \sum_{i=1}^{q-1}$.

APPENDIX D

AVERAGE BICOMPOSITION ENUMERATOR OF UNIFORM PARITY-CHECK ENSEMBLES

This appendix provides results on the average bicomposition and biweight enumerators of some ensembles.

Proposition 5. *Consider the uniform parity-check ensemble defined by a random parity-check matrix of size $(h-k)\times h$ with i.i.d. entries with uniform distribution in $\mathbb{F}_q$. For all $\boldsymbol{\kappa}\in\mathscr{K}_{q,h}$, the expected joint composition enumerator for a random code drawn for the ensemble is*

$$\mathsf{S}_{\boldsymbol{\kappa}}=\binom{h}{\boldsymbol{\kappa}}q^{-2(h-k)}.$$

Proof: The parameter $\mathsf{S}_{\boldsymbol{\kappa}}$ may be expressed as the total number of pairs $(\mathbf{r}_1,\mathbf{r}_2)\in\mathbb{F}_q^h\times\mathbb{F}_q^h$ with joint composition $\boldsymbol{\kappa}$, times the probability that both $\mathbf{r}_1$ and $\mathbf{r}_2$ are codewords given that their joint composition is $\boldsymbol{\kappa}$. Hence, we can write

$$\mathsf{S}_{\boldsymbol{\kappa}}=\binom{h}{\boldsymbol{\kappa}}\Pr\{\{\mathbf{r}_1\mathbf{H}^\mathsf{T}=\mathbf{0}\}\cap\{\mathbf{r}_2\mathbf{H}^\mathsf{T}=\mathbf{0}\}|\kappa(\mathbf{r}_1,\mathbf{r}_2)=\boldsymbol{\kappa}\}=\binom{h}{\boldsymbol{\kappa}}\mathsf{p}_{\boldsymbol{\kappa}}^{h-k}$$

where, letting $\mathbf{h}$ be the generic row of $\mathbf{H}$, $\mathsf{p}_{\boldsymbol{\kappa}}=\Pr\{\{\mathbf{r}_1\mathbf{h}^\mathsf{T}=0\}\cap\{\mathbf{r}_2\mathbf{h}^\mathsf{T}=0\}|\kappa(\mathbf{r}_1,\mathbf{r}_2)=\boldsymbol{\kappa}\}$. If $\boldsymbol{\kappa}\in\mathscr{K}_{q,h}$ then five different cases may occur; next we show that in all of them we have $\mathsf{p}_{\boldsymbol{\kappa}}=q^{-2}$. We repeatedly exploit the following property: if $\mathbf{r}\in\mathbb{F}_q^h$ and $\mathbf{h}$ is a random vector in $\mathbb{F}_q^h$ whose elements are uniform i.i.d. random variables in $\mathbb{F}_q$, then $\Pr\{\mathbf{r}\,\mathbf{h}^\mathsf{T}=\beta\}=q^{-1}$ for all $\beta\in\mathbb{F}_q$. For the sake of notational simplicity, we denote by $E_{\boldsymbol{\kappa}}$ the event that $\kappa(\mathbf{r}_1,\mathbf{r}_2)=\boldsymbol{\kappa}$.

Case 1: $|\boldsymbol{\kappa}_1|>0$, $|\boldsymbol{\kappa}_2|>0$, $|\boldsymbol{\kappa}_3|>0$ ($\mathbf{r}_1$ and $\mathbf{r}_2$ have partially overlapping supports). Without loss of generality, assume $\mathbf{r}_1=(\mathbf{r}_{1,1}|\mathbf{r}_{1,2}|\mathbf{0}|\mathbf{0})$ and $\mathbf{r}_2=(\mathbf{0}|\mathbf{r}_{2,1}|\mathbf{r}_{2,2}|\mathbf{0})$, where $\mathbf{r}_{1,1}$, $\mathbf{r}_{1,2}$, $\mathbf{r}_{2,1}$, and $\mathbf{r}_{2,2}$ are nonzero and all subvectors occupying the same position have the same length. Letting $\mathbf{h}=(\mathbf{h}_1|\mathbf{h}_2|\mathbf{h}_3|\mathbf{h}_4)$ we have $\mathsf{p}_{\boldsymbol{\kappa}}=\Pr\{\{\mathbf{r}_{1,1}\mathbf{h}_1^\mathsf{T}+\mathbf{r}_{1,2}\mathbf{h}_2^\mathsf{T}=0\}\cap\{\mathbf{r}_{2,1}\mathbf{h}_2^\mathsf{T}+\mathbf{r}_{2,2}\mathbf{h}_3^\mathsf{T}=0\}|E_{\boldsymbol{\kappa}}\}=\Pr\{\mathbf{r}_{1,1}\mathbf{h}_1^\mathsf{T}+\mathbf{r}_{1,2}\mathbf{h}_2^\mathsf{T}=0|E_{\boldsymbol{\kappa}}\}\Pr\{\mathbf{r}_{2,1}\mathbf{h}_2^\mathsf{T}+\mathbf{r}_{2,2}\mathbf{h}_3^\mathsf{T}=0|E_{\boldsymbol{\kappa}}\}=(q^{-1})(q^{-1})=q^{-2}$, where we exploited independence of $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$.

Case 2: $|\boldsymbol{\kappa}_1|>0$, $|\boldsymbol{\kappa}_2|=0$, $|\boldsymbol{\kappa}_3|>0$ (the support of $\mathbf{r}_2$ includes that of $\mathbf{r}_1$). Same argument with $\mathbf{r}_{1,1}=\mathbf{0}$.

Case 3: $|\boldsymbol{\kappa}_1|=0$, $|\boldsymbol{\kappa}_2|>0$, $|\boldsymbol{\kappa}_3|>0$ (the support of $\mathbf{r}_1$ includes that of $\mathbf{r}_2$). Same argument with $\mathbf{r}_{2,2}=\mathbf{0}$.

Case 4: $|\boldsymbol{\kappa}_1|>0$, $|\boldsymbol{\kappa}_2|>0$, $|\boldsymbol{\kappa}_3|=0$ ($\mathbf{r}_1$ and $\mathbf{r}_2$ have disjoint supports). Same argument with $\mathbf{r}_{1,2}=\mathbf{r}_{2,1}=\mathbf{0}$.

Case 5: $|\boldsymbol{\kappa}_1|=|\boldsymbol{\kappa}_2|=0$, $|\boldsymbol{\kappa}_3|>0$, $\kappa_{0,0}+\sum_{i=1}^{q-1}\kappa_{i,(i+b)\mathrm{mod}q}<h$ for all $b\in\{0,\dots,q-2\}$ ($\mathbf{r}_1$ and $\mathbf{r}_2$ have the same support but are not linearly dependent). Let $\mathbf{r}_1=(r_{1,0},\dots,r_{1,h-1})$, $\mathbf{r}_2=(r_{2,0},\dots,r_{2,h-1})$ and $\mathbf{h}=(\mathrm{h}_0,\dots,\mathrm{h}_{h-1})$. Since $\mathbf{r}_1$ and $\mathbf{r}_2$ are nonzero and not linearly dependent, there exist s and t such that the vectors $(r_{1,s},r_{1,t})$ and $(r_{2,s},r_{2,t})$ are linearly independent. Letting $\beta_1=-\sum_{i=0,i\neq s,t}^{h-1}r_{1,i}\mathrm{h}_i$ and $\beta_2=-\sum_{i=0,i\neq s,t}^{h-1}r_{2,i}\mathrm{h}_i$ we obtain $\mathsf{p}_{\boldsymbol{\kappa}}=\Pr\{\{r_{1,s}\mathrm{h}_s+r_{1,t}\mathrm{h}_t=$

$\beta_1\} \cap \{r_{2,s}\mathrm{h}_s + r_{2,t}\mathrm{h}_t = \beta_2\}|E_{\boldsymbol{\kappa}}\}$. Linear independence of $(r_{1,s}, r_{1,t})$ and $(r_{2,s}, r_{2,t})$ implies that for any β_1 and β_2 there exists a unique pair $(\mathrm{h}_s, \mathrm{h}_t)$ fulfilling the two equations. Since all pairs are equiprobable and their number is q^2 we have $\mathsf{p}_{\boldsymbol{\kappa}} = q^{-2}$. ■

The following result is a direct consequence of Proposition 5 in the binary case.

Proposition 6. *Consider the uniform parity-check ensemble defined by a random parity-check matrix of size $(h-k) \times h$ with i.i.d. entries with uniform distribution in $\mathbb{F}_2$. For all $\boldsymbol{\tau} \in \mathscr{T}_{2,h}$, the expected joint composition enumerator for a random code drawn for the ensemble is*

$$\mathsf{J}_{\boldsymbol{\tau}} = \binom{h}{\boldsymbol{\tau}} 4^{-(h-k)}.$$

Proof: Recall from Remark 1 that for $q = 2$ the two concepts of joint composition and joint weight become equivalent so that, letting $\boldsymbol{\tau} = \tau(\boldsymbol{\kappa})$, we can write $\mathsf{J}_{\boldsymbol{\tau}} = \mathsf{S}_{\boldsymbol{\kappa}}$. ■

Acknowledgment

The authors would like to thank the Associate Editor and the anonymous reviewers for their valuable comments, which helped to substantially improve the paper.

References

[1] F. Lázaro, G. Liva, E. Paolini, and G. Bauch, “Bounds on the error probability of Raptor codes,” in *Proc. IEEE Global Commun. Conf.*, Washington DC, USA, Dec. 2016.

[2] J. Byers, M. Luby, and M. Mitzenmacher, “A digital fountain approach to reliable distribution of bulk data,” *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, pp. 1528–1540, Oct. 2002.

[3] M. Luby, “LT codes,” in *Proc. 43rd Annual IEEE Symp. on Foundations of Computer Science*, Vancouver, Canada, Nov. 2002, pp. 271–282.

[4] M. Shokrollahi, “Raptor codes,” *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.

[5] ETSI TS 26.346 V13.3.0, “UMTS; LTE; Multimedia Broadcast / Multicast Service; Protocols and Codecs,” Jan. 2016.

[6] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, “RFC 5053: Raptor forward error correction scheme: Scheme for object delivery,” IETF, Tech. Rep., Oct. 2007.

[7] E. Berlekamp, *Algebraic coding theory*. New York: McGraw-Hill, 1968.

[8] B. A. LaMacchia and A. M. Odlyzko, “Solving large sparse linear systems over finite fields,” *Advances in Cryptology-CRYPT0’90*, pp. 109–133, 1991.

[9] H. Pishro-Nik and F. Fekri, “On decoding of low-density parity-check codes over the binary erasure channel,” *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 439–454, Mar. 2004.

[10] D. Burshtein and G. Miller, “An efficient maximum likelihood decoding of LDPC codes over the binary erasure channel,” *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2837–2844, Nov. 2004.

[11] M. Shokrollahi, S. Lassen, and R. Karp, “Systems and processes for decoding chain reaction codes through inactivation,” Feb. 2005, US Patent 6,856,263.

[12] K. Mahdaviani, M. Ardakani, and C. Tellambura, “On Raptor code design for inactivation decoding,” *IEEE Commun. Lett.*, vol. 60, no. 9, pp. 2377–2381, Sep. 2012.

[13] F. Lázaro Blasco, G. Liva, and G. Bauch, “LT code design for inactivation decoding,” in *Proc. 2014 IEEE Inf. Theory Workshop*, Hobart, Tasmania, Australia, Nov. 2014, pp. 441–445.

[14] ——, “Enhancing the LT component of Raptor codes,” in *Proc. of the 10th Int. ITG Conf. Systems, Commun. and Coding*, Hamburg, Germany, Feb. 2015.

[15] F. Lázaro, G. Liva, and G. Bauch, “Inactivation decoding of LT and Raptor codes: Analysis and code design,” *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4114–4127, Oct. 2017.

[16] N. Rahnavard, B. Vellambi, and F. Fekri, “Rateless codes with unequal error protection property,” *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1521–1532, Apr. 2007.

[17] B. Schotsch, G. Garrammone, and P. Vary, “Analysis of LT codes over finite fields under optimal erasure decoding,” *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1826–1829, Sep. 2013.

[18] B. E. Schotsch, “Rateless coding in the finite length regime,” Ph.D. dissertation, Inst. of Commun. Systems and Data Proc., RWTH Aachen, Aachen, Germany, Jul. 2014.

[19] G. Liva, E. Paolini, and M. Chiani, “Performance versus overhead for fountain codes over $\mathbb{F}_q$,” *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 178–180, Feb. 2010.

[20] P. Wang, G. Mao, Z. Lin, M. Ding, W. Liang, X. Ge, and Z. Lin, “Performance analysis of Raptor codes under maximum likelihood decoding,” *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 906–917, Mar. 2016.

[21] F. Lázaro Blasco, E. Paolini, G. Liva, and G. Bauch, “On the weight distribution of fixed-rate Raptor codes,” in *Proc. 2015 IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015, pp. 2880–2884.

[22] F. Lázaro, E. Paolini, G. Liva, and G. Bauch, “Distance spectrum of fixed-rate Raptor codes with linear random precoders,” *IEEE J. Select. Areas Commun.*, vol. 34, no. 2, pp. 422–436, Feb. 2016.

[23] K. Zhang, Q. Zhang, and J. Jiao, “Bounds on the reliability of RaptorQ codes in the finite-length regime,” *IEEE Access*, vol. 5, no. 5, pp. 24 766–24 774, Oct. 2017.

[24] RFC 6330, “Network working group; Request for Comments: 5053; RaptorQ Forward Error Correction Scheme for Object Delivery,” Aug. 2011.

[25] D. A. Dawson and D. Sankoff, “An inequality for probabilities,” *Proc. American Math. Society*, vol. 18, no. 3, pp. 504–507, Jun. 1967.

[26] O. Barak and D. Burshtein, “Lower bounds on the error rate of LDPC code ensembles,” *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4225–4236, Nov 2007.

[27] C. Bonferroni, “Teoria statistica classi e calcolo delle probabilità,” *Pubbl. R. Ist. Super. Sci. Econ. Comm. Firenze*, vol. 8, pp. 3–62, 1936.

[28] S. M. Kwerel, “Most stringent bounds on aggregated probabilities of partially specified dependent probability systems,” *J. Amer. Statist. Assoc.*, vol. 70, no. 350, pp. 472–479, Jun. 1975.

[29] F. Mac Williams and N. Sloane, *The theory of error-correcting codes*. North Holland Mathematical Libray, 1977, vol. 16.

[30] F. MacWilliams, C. Mallows, and N. Sloane, “Generalizations of Gleason’s theorem on weight enumerators of self-dual codes,” *IEEE Trans. Inf. Theory*, vol. 18, no. 6, pp. 794–805, Nov. 1972.

[31] M. Shokrollahi and M. Luby, “Systematic encoding and decoding of chain reaction codes,” Jun. 2005, US Patent 6,909,383.

[32] F. Lázaro, “Fountain codes under maximum likelihood decoding,” Ph.D. dissertation, Institute for Telecommunications, Hamburg University of Technology, Hamburg, Germany, 2017.

[33] M. Luby, A. Shokrollahi, M. Watson, T. Stockhammer, and L. Minder, “RFC 6330: RaptorQ forward error correction scheme for object delivery,” IETF, Tech. Rep., Aug. 2011.

[34] A. Shokrollahi and M. Luby, “Raptor codes,” *Foundations and Trends in Commun. and Inf. Theory*, vol. 6, no. 3-4, pp. 213–322, 2011.

[35] S. Kirkpatrick, D. Gelatt, and M. Vecchi, “Optimization by simmulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[36] J. van Lint, *Introduction to Coding Theory*, ser. Graduate Texts in Mathematics. Springer Berlin Heidelberg, 1998.

[37] G. Battail, M. Decouvelaere, and P. Godlewski, “Replication decoding,” *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 332–345, May 1979.

[38] D. Burshtein and G. Miller, “Asymptotic enumeration methods for analyzing LDPC codes,” *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1115–1131, Jun. 2004.

[39] K. Kasai, C. Poulliat, D. Declercq, T. Shibuya, and K. Sakaniwa, “Weight distribution of non-binary LDPC codes,” in *Proc. 2008 Int. Symp. Inf. Theory and App*, Dec. 2008, pp. 1–6.