Analysis of the Second Moment of the LT Decoder

Ghid Maatouk, *Student Member, IEEE,* and Amin Shokrollahi, *Fellow, IEEE*

***Abstract*—In this paper, the second moment of the ripple size during the LT decoding process is analyzed. The standard deviation of the ripple size for an LT code with length k is shown to be of the order of $\sqrt{k}$. Together with a result by Karp et. al (2004) stating that the expectation of the ripple size is of the order of k, this gives bounds on the error probability of the LT decoder. Further, an analytic expression for the variance of the ripple size up to terms of constant order is given, and the expression of Karp et. al for the expectation of the ripple size is refined up to terms of the order of $1/k$. This provides a first step towards an analytic finite-length analysis of LT decoding.**

***Index Terms*—LT decoder, second moment, ripple, finite-length analysis.**

I. Introduction

We start with a brief introduction to Fountain codes, LT codes and belief propagation (BP) decoding. For details, the reader is referred to [1], [2].

LT codes belong to the class of Fountain codes. A Fountain code generates, given a set of k input symbols, a potentially infinite stream of output symbols $z_1, z_2, \ldots$, where each output symbol is produced independently from the addition of some subset of the input symbols, chosen according to a distribution on $\mathbb{F}_2^k$. We assume that the symbols can be either bits or binary vectors, and that there is a way for the receiver to know, for each output symbol, which input symbols it is produced from. A number of such ways are described in [1]. A good Fountain code is one for which the receiver can decode the k input symbols with high probability after collecting n output symbols, with n close to k.

Such codes are well-suited for reliable transmission of data packets over the Internet. The Internet can be modeled as an erasure channel, where each packet is either lost, discarded or delivered to the receiver. Fountain codes are well adapted for many transmission scenarios, such as the existence of multiple receivers on one or multiple channels. Then each receiver can recover the k input symbols independently of the other receivers, as soon as it has gathered n output symbols. Good Fountain codes operate close to capacity for any erasure channel.

LT codes are *universal* Fountain codes [1] in that the decoding process can recover with high probability a set of k input symbols from n output symbols with n arbitrarily close to k. LT codes have the following distribution for generating output symbols: for each output symbol, first sample a number d (the "degree" of this symbol) from a distribution $\Omega = (\Omega_1, \cdots, \Omega_k)$ on the integers $1, \ldots, k$. Then pick d distinct input symbols uniformly at random and XOR them to produce the corresponding output symbol. An LT code that encodes k input symbols and uses a distribution Ω with generating function $\Omega(x) = \sum_i \Omega_i x^i$ is said to have parameters $(k, \Omega(x))$.

LT codes are decoded using BP decoding. The decoding starts when the receiver has gathered $n = (1+\epsilon)k$ output symbols, for some predetermined *overhead* ϵ. Define the *decoding graph* [2] to be an undirected bipartite graph with k nodes on one side, representing the k input symbols, and n nodes on the other, representing the output symbols. An input node is connected to an output node if the corresponding input symbol contributes to the value of the output symbol. The decoding process is as follows: if the receiver can find an output symbol connected to only one input symbol, then the value of this input symbol can be recovered directly. This value is XORed to the value of any other output symbols connected to this input symbol, then the input symbol and all its outgoing edges are removed from the graph. The decoder then repeats the operation by finding another output symbol connected to only one input symbol. If at any stage before the recovery of all symbols, no such output symbol is found, the decoder reports an error.

An important set to consider is the set of output symbols of degree 1 (the *ripple*). The size of the ripple varies during the decoding process, as high-degree output symbols become of degree 1 after the removal of their edges, and as ripple elements become useless after the recovering of their unique neighbor. The decoding is in error if and only if the ripple becomes empty before all the input symbols are recovered. A natural question is thus whether we can track the size of the ripple, in the expectation, during the decoding process. Karp et al. [3] proved that the expected ripple size is linear in k throughout most of the decoding process. Their asymptotic analytic expressions for the expected ripple size can be found in section II. They also derive an expression for the expected *cloud* size throughout decoding, where the cloud is defined at each decoding step as the set of output symbols of degree strictly higher than 1. We are interested in the cloud size inasmuch as the cloud "feeds" the ripple during the decoding process, as higher-degree symbols lose edges. Thus, expressions for the expectation and higher moments of the ripple size depend on the corresponding expressions for the cloud size.

In this paper, we extend the analysis of [3] in two ways.

This research was supported by Grant 228021-ECCSciEng of the European Research Council.

The authors are with the School of Computer and Communications Sciences, EPFL, 1015 Lausanne, Switzerland (email: {amin.shokrollahi, ghid.maatouk}@epfl.ch).

First, we consider higher moments of the cloud and ripple size in order to upper bound the error probability of the LT decoder. More specifically, we use similar methods to derive an expression for the variance of the ripple size and prove that it is also linear in k throughout most of the decoding process. We can then use this expression together with the expression for the expectation to offer a guarantee for successful decoding, as follows: if, for fixed LT code parameters, $R(u)$ is the expectation and $\sigma_R(u)$ is the standard deviation of the ripple size when u symbols are unrecovered, then if the function

$$h_c(u) = R(u) - c \cdot \sigma_R(u) \tag{1}$$

for some parameter c never takes negative values, we can upper bound the error probability of the LT decoder by the probability that the ripple size deviates from its mean by more than c standard deviations. This is easily done using Chebyshev's inequality.

Second, we take the first step towards an analytic finite-length analysis of the LT decoder, by providing exact expressions for the expectation (variance) of the ripple size up to $O(1/k)$ (constant) terms. This is done by considering lower-order terms in the difference equations, but also by getting tight bounds on the discrepancy introduced by approximating difference equations by differential equations.

It is worthy to note that the expressions we deal with are valid for "most of the decoding process," that is, the analysis breaks down when the number of unrecovered symbols is no longer a constant fraction of k. This is no issue, however, when one considers Raptor codes, which need only a constant fraction of the input symbols to be recovered by the LT decoder [2].

II. Preliminaries - an expression for the expected ripple size

Let u be the number of unrecovered (*undecoded*) input symbols at a given decoding step. Define the decoder to be in state (c, r, u) if the cloud size is c and the ripple size is r at this decoding step. To each state (c, r, u), we can associate the probability $p_{c,r,u}$ of the decoder being in this state. Define the *state generating function* of the LT decoder when u symbols are undecoded as

$$P_u(x, y) = \sum_{c \geq 0, r \geq 1} p_{c,r,u} x^c y^{r-1}.$$

The following theorem by Karp et al. gives a recursion for the state generating function of the LT decoder.

Theorem 1: [3] Suppose that the original code has k input symbols and that $n = k(1 + \delta)$ output symbols have been collected for decoding. Further, denote by Ω_i, $i = 2, \ldots, D$, the probability that an output symbol is of degree i, where D is the maximum degree of an output symbol. Then we have for $u = k + 1, k, \ldots, 1$

$$\begin{aligned} P_{u-1}(x, y) = \frac{1}{y}\Bigg[& P_u\left(x(1 - p_u) + y p_u, \frac{1}{u} + y\left(1 - \frac{1}{u}\right)\right) \\ & - P_u\left(x(1 - p_u), \frac{1}{u}\right)\Bigg], \end{aligned} \tag{2}$$

where for $u \leq k$,

$$p_u = \frac{\frac{u-1}{k(k-1)} \sum_{d=1}^{D} \Omega_d d(d-1) \frac{\begin{bmatrix} k-u \\ d-2 \end{bmatrix}}{\begin{bmatrix} k-2 \\ d-2 \end{bmatrix}}}{1 - u \sum_{d=1}^{D} \Omega_d d \frac{\begin{bmatrix} k-u \\ d-1 \end{bmatrix}}{\begin{bmatrix} k \\ d \end{bmatrix}} - \sum_{d=1}^{D} \Omega_d \frac{\begin{bmatrix} k-u \\ d \end{bmatrix}}{\begin{bmatrix} k \\ d \end{bmatrix}}},$$

and

$$\begin{bmatrix} a \\ b \end{bmatrix} := \binom{a}{b} b!,$$

and $p_{k+1} := \Omega_1$. Further, $P_{k+1}(x, y) := x^n$.

This recursion gives a way to compute the probability of a decoding error at each step of the BP decoding as

$$P_{err}(u) = \sum_{c \geq 0} p_{c,0,u} = 1 - \sum_{c \geq 0, r \geq 1} p_{c,r,u} = 1 - P_u(1, 1),$$

and the overall error probability of the decoder as

$$P_{err} = \sum_{u=1}^{k} P_{err}(u).$$

If we approximate the LT process by allowing output symbols to choose their neighbors with replacement during encoding, p_u becomes:

$$p_u = \frac{1}{k} f\left(\frac{u}{k}\right) - \frac{1}{k^2} g\left(\frac{u}{k}\right) = \frac{1}{k} f\left(\frac{u}{k}\right) + O(1/k^2),$$

where

$$f(x) := \frac{x\Omega''(1 - x)}{1 - x\Omega'(1 - x) - \Omega(1 - x)} \tag{3}$$

and

$$g(x) := \frac{f(x)}{x}. \tag{4}$$

For simplicity, the process that we analyze in what follows is this modified LT process. Intuitively, the modified process is "worse" than the original LT process in that it allows for multiple, "useless" edges in the decoding graph. With this assumption, Karp et al. [3] use the recursion to derive difference equations for the expected size of the ripple and the cloud, and further approximate these difference equations by differential equations that they solve to get closed-form expressions for the expected ripple and cloud size. Formally, let

$$R(u) := \sum_{c \geq 0, r \geq 1} (r - 1) p_{c,r,u}$$

denote the expected number of output symbols in the ripple when u symbols are undecoded, and let

$$C(u) := \sum_{c \geq 0, r \geq 1} c p_{c,r,u}$$

denote the expected number of output symbols in the cloud when u input symbols are undecoded, where u is assumed to be a constant fraction of the total number of input symbols k. Then Karp et al. [3] derive closed-form expressions for continuous approximations of $R(u)$ and $C(u)$. More precisely, let $x := u/k$ denote the *fraction* of undecoded symbols, and let $C(x) := C(u)/n$ be a normalized version of $C(u)$. Then the continuous function $\hat{C}(x)$ given by

$$\hat{C}(x) = c_0 \left(1 - x\Omega'(1-x) - \Omega(1-x)\right),$$

with

$$c_0 = 1 - (1-\Omega_1)^{n-1} \tag{5}$$

is a "good" approximation for $C(x)$.

Similarly, let $R(x) := R(u)/n$ be a normalized version of $R(u)$. Then the continuous function $\hat{R}(x)$ given by

$$\hat{R}(x) = x\left(c_0\Omega'(1-x) + \frac{1}{1+\epsilon}\ln x + r_0\right), \tag{6}$$

with

$$r_0 = \Omega_1(1-\Omega_1)^{n-1} - \frac{1-(1-\Omega_1)^n}{n} \tag{7}$$

is a "good" approximation for $R(x)$.[1] Theorem 2 formalizes this notion of a "good" approximation.

Theorem 2: [3] Consider an LT code with parameters $(k, \Omega(x))$ and assume $n = (1+\epsilon)k$ symbols have been collected for decoding. During BP decoding, let $C(u)$ and $R(u)$ be respectively the expected size of the cloud and ripple as a function of the number u of undecoded input symbols. Then, under the assumptions that u is a constant fraction of k and $\Omega_1 > 0$, we have

$$\begin{aligned} C(u) &= n\hat{C}(u/k) + O(1) \\ &= n\left(1 - \frac{u}{k}\Omega'(1-u/k) - \Omega(1-u/k)\right) + O(1) \end{aligned}$$

and

$$\begin{aligned} R(u) &= n\hat{R}(u/k) + O(1) \\ &= (1+\epsilon)u\left(\Omega'(1-u/k) + \frac{1}{1+\epsilon}\ln\frac{u}{k}\right) + O(1). \end{aligned} \tag{8}$$

III. An Expression for the Variance of the Ripple Size

Let $\sigma_R^2(u)$ be the variance of the ripple size as a function of the number of undecoded symbols u. In what follows we will always assume that u is a constant fraction of k. By definition, $\sigma_R^2(u)$ is given by

$$\sigma_R^2(u) = \sum_{c\geq 0, r\geq 1} (r-1)^2 p_{c,r,u} - R(u)^2.$$

If we define

$$\begin{aligned} N(u) &:= \frac{\partial^2 P_u}{\partial y^2}(1,1) \\ &= \sum_{c\geq 0, r\geq 1} (r-1)(r-2) p_{c,r,u} \\ &= \sum_{c\geq 0, r\geq 1} (r-1)^2 p_{c,r,u} - R(u), \end{aligned} \tag{9}$$

we can relate $\sigma_R^2(u)$, $N(u)$ and $R(u)$ as follows:

$$\sigma_R^2(u) = N(u) - R(u)^2 + R(u). \tag{10}$$

It is thus enough to find an expression for $N(u)$ to get an expression for $\sigma_R^2(u)$. We start by differentiating both sides of the recursion (2) twice with respect to y and evaluating at $(1,1)$. This gives us a recursion for $N(u)$:

$$\begin{aligned} N(u-1) = &\left(1-\frac{1}{u}\right)^2 N(u) - 2p_u C(u) - 2\left(1-\frac{1}{u}\right)R(u) \\ &+ p_u^2 \frac{\partial^2 P_u}{\partial x^2}(1,1) + 2p_u\left(1-\frac{1}{u}\right)\frac{\partial^2 P_u}{\partial x \partial y}(1,1) \\ &- 2\left[-P_u(1,1) + P_u\left(1-p_u, \frac{1}{u}\right)\right]. \end{aligned} \tag{11}$$

Before we can proceed with solving this difference equation, we need to find expressions for the second-order derivatives $\frac{\partial^2 P_u}{\partial x^2}(1,1)$ and $\frac{\partial^2 P_u}{\partial x \partial y}(1,1)$. We do so by following exactly the same method that we are currently outlining for an expression for $N(u)$. Define

$$\begin{aligned} M(u) &:= \frac{\partial^2 P_u}{\partial x^2}(1,1) \\ L(u) &:= \frac{\partial^2 P_u}{\partial x \partial y}(1,1). \end{aligned}$$

Then Theorems 3 and 4 gives closed-form expressions for $M(u)$ and $L(u)$, respectively.

Theorem 3: Let $M(x) := M(u)/n^2$ be a normalized version of $M(u)$ (where x denotes, as before, the fraction u/k of undecoded symbols). Then

$$M(x) = \hat{M}(x) + O(1/k),$$

where

$$\hat{M}(x) = m_0\left(1 - x\Omega'(1-x) - \Omega(1-x)\right)^2,$$

with

$$m_0 = \left(1-\frac{1}{n}\right)\left(1-(1-\Omega_1)^{n-2}\right). \tag{12}$$

Proof: See Appendix F. ■

Theorem 4: Let $L(x) := L(u)/n^2$ be a normalized version of $L(u)$. Then

$$L(x) = \hat{L}(x) + O(1/k),$$

where

$$\begin{aligned} \hat{L}(x) = &\, x\left(1 - x\Omega'(1-x) - \Omega(1-x)\right) \\ &\cdot \left(m_0\Omega'(1-x) + \frac{c_0}{1+\epsilon}\ln x + l_0\right), \end{aligned}$$

with

$$l_0 = \frac{-1}{n} + (1-\Omega_1)^{n-2}\left(\Omega_1 + \frac{1-2\Omega_1}{n}\right). \tag{13}$$

Proof: See Appendix G. ■

As for the "dirt" term

$$-2\left[-P_u(1,1) + P_u\left(1-p_u, \frac{1}{u}\right)\right], \tag{14}$$

it does not involve derivatives and we cannot use the same method to find an expression for it independent of the state generating function. However, we can bound

[1] The expressions we present here for $\hat{C}(x)$ and $\hat{R}(x)$ correct some slight typos in [3].

it under an assumption on the ripple size, as Theorem 5 shows.

Theorem 5: If $r \geq 4$,

$$2\left[-P_u(1,1) + P_u\left(1 - p_u, \frac{1}{u}\right)\right] = O(1).$$

Proof: See Appendix H. ■

In what follows, we assume that the size of the ripple does not go below the constant 4.[2]

Replacing $M(u)$ and $L(u)$ by their expressions and bounding the dirt term in the recursion (11), we obtain the following difference equation for $N(u)$:

$$\begin{aligned} N(u) - N(u-1) = &\left(\frac{2}{u} - \frac{1}{u^2}\right) N(u) - p_u^2 M(u) \\ &- 2p_u\left(1 - \frac{1}{u}\right) L(u) + 2p_u C(u) \qquad (15) \\ &+ 2\left(1 - \frac{1}{u}\right) R(u) + O(1). \end{aligned}$$

Note that $N(u)$ as defined in equation (9) can be as large as a constant fraction of k^2. We thus need to normalize $N(u)$ if we want to say something meaningful about the difference $N(u) - N(u-1)$. We let $N(x) := N(u)/n^2$ be a normalized version of $N(u)$, where x denotes, as before, the fraction u/k of undecoded symbols. Normalizing equation (15) and replacing the functions $M(x), L(x), C(x)$ and $R(x)$ by their continuous approximations, we obtain

$$\begin{aligned} N(x) - N(x - 1/k) = &\frac{2}{kx}N(x) - \frac{2}{k}f(x)\hat{L}(x) \\ &+ \frac{2}{(1+\epsilon)k}\hat{R}(x) + O(1/k^2). \end{aligned}$$

Neglecting lower-order terms, we approximate $N(x)$ by the function $\tilde{N}(x)$ which satisfies

$$\tilde{N}(x) - \tilde{N}(x - 1/k) = \frac{2}{kx}\tilde{N}(x) - \frac{2}{k}f(x)\hat{L}(x) + \frac{2}{(1+\epsilon)k}\hat{R}(x),$$

with initial condition $\tilde{N}(1) = N(1)$.

Claim 1: For any x on which $N(x)$ is defined, $N(x)$ and $\tilde{N}(x)$ differ by a term of the order of $1/k$.

Proof: See Appendix A. ■

We further approximate the discrete function $\tilde{N}(x)$ by the continuous function $\hat{N}(x)$, and

$$\frac{\tilde{N}(x) - \tilde{N}(x - 1/k)}{1/k}$$

by the first-order derivative of $\hat{N}(x)$. $\hat{N}(x)$ satisfies the differential equation

$$\hat{N'}(x) = \frac{2}{x}\hat{N}(x) - 2f(x)\hat{L}(x) + \frac{2}{1+\epsilon}\hat{R}(x) \qquad (16)$$

with initial condition $\hat{N}(1) = \tilde{N}(1)$.

Claim 2: For any x on which $\tilde{N}(x)$ is defined, $\tilde{N}(x)$ and $\hat{N}(x)$ differ by a term of the order of $1/k$.

Proof: See Appendix B. ■

The following theorem gives the solution of the differential equation (16).

Theorem 6: An analytic expression for $\hat{N}(x)$ is

$$\begin{aligned} \hat{N}(x) = x^2\Big(&m_0\Omega'(1-x)^2 + 2l_0\Omega'(1-x) + \\ &\frac{2c_0}{1+\epsilon}\Omega'(1-x)\ln x + \frac{2r_0}{1+\epsilon}\ln x + \frac{1}{(1+\epsilon)^2}(\ln x)^2 + n_0\Big), \end{aligned} \qquad (17)$$

where the constants c_0, m_0 and l_0 are given by equations (5), (12) and (13), respectively, and the value of the constant n_0 is

$$n_0 = \frac{2}{n^2}\left(1 - (1-\Omega_1)^n\right) - (1-\Omega_1)^{n-2}\left(\Omega_1^2 + \frac{2\Omega_1 - 3\Omega_1^2}{n}\right).$$

Proof: See Appendix C. ■

By Claims 1 and 2 we thus have

$$N(x) = \hat{N}(x) + O(1/k),$$

where $\hat{N}(x)$ is given by equation (17). This gives us an expression for $N(u)$, up to a term of the order of k:

$$\begin{aligned} N(u) = (1+\epsilon)^2 u^2\Big(&\Omega'(1-u/k)^2 + \frac{2}{1+\epsilon}\Omega'(1-u/k)\ln\frac{u}{k} \\ &+ \frac{1}{(1+\epsilon)^2}\left(\ln\frac{u}{k}\right)^2\Big) + O(k). \end{aligned}$$

Comparing this expression to that for $R(u)^2$ given by equations (6) and (8), it is easy to see that these two expressions agree up to terms of the order of k, so that the variance of the ripple size

$$\sigma_R^2(u) = N(u) - R(u)^2 + R(u)$$

is of the order of k.

Theorem 7: Consider an LT code with parameters $(k, \Omega(x))$ and let $\sigma_R(u)$ be the standard deviation of the ripple size throughout BP decoding. Then

$$\sigma_R(u) = O(\sqrt{k}).$$

IV. Toward a Finite-Length Analysis of the LT Decoder

Our ultimate goal is to be able to bound the error probability of the decoder as a function of k, without the assumption that k goes to infinity. We thus need to find an expression for the variance of the ripple size, instead of simply determining its order. For this purpose, we must find an expression for $N(u)$ up to terms of constant order, and an expression for $R(u)$ up to terms of the order of $1/k$. We illustrate the analysis for $N(u)$. From the recursion given by equation (11), we proceed by first, bounding the“dirt” term more carefully as

$$2\left(1 - P_u(1,1) + P_u\left(1 - p_u, \frac{1}{u}\right)\right) = 2 + O(1/k),$$

[2] It is not difficult to check at the end of the analysis, and using an inductive reasoning, that this assumption holds with high probability.

as the derivation in Appendix H shows. We then replace $C(x)$, $R(x)$, $M(x)$ and $L(x)$ by finer approximations. For this, we define the following discrepancy terms.

Definition 1: Let

$$d_C(x) := \hat{C}(x) - C(x)$$

denote the discrepancy between $C(x)$ and its continuous approximation $\hat{C}(x)$ when an x-fraction of symbols is undecoded. Similarly, let

$$\begin{array}{rcl} d_R(x) & := & \hat{R}(x) - R(x) \\ d_M(x) & := & \hat{M}(x) - M(x) \\ d_L(x) & := & \hat{L}(x) - L(x) \end{array}$$

denote the corresponding discrepancies for $R(x)$, $M(x)$ and $L(x)$, respectively.

Then the following theorem gives expressions for these discrepancy terms.

Theorem 8: The discrepancy terms $d_C(x)$, $d_R(x)$, $d_M(x)$ and $d_L(x)$ are of the order of $1/k$ and are given by the following expressions:

$$\begin{array}{rcl} d_C(x) & = & \dfrac{1}{k^2} \displaystyle\sum_{i=0}^{k(1-x)-1} C_i \prod_{j=i+1}^{k(1-x)-1} \left(1 - \frac{c_j}{k}\right) + O(1/k^2) \\ d_R(x) & = & \dfrac{1}{k^2} \displaystyle\sum_{i=0}^{k(1-x)-1} R_i \prod_{j=i+1}^{k(1-x)-1} \left(1 - \frac{r_j}{k}\right) + O(1/k^2) \\ d_M(x) & = & \dfrac{1}{k^2} \displaystyle\sum_{i=0}^{k(1-x)-1} M_i \prod_{j=i+1}^{k(1-x)-1} \left(1 - \frac{2c_j}{k}\right) + O(1/k^2) \\ d_L(x) & = & \dfrac{1}{k^2} \displaystyle\sum_{i=0}^{k(1-x)-1} L_i \prod_{j=i+1}^{k(1-x)-1} \left(1 - \frac{r_j + c_j}{k}\right) + O(1/k^2), \end{array}$$

with the constants C_i, R_i, M_i, L_i and c_j, r_j given by

$$\begin{array}{rcl} C_i & = & \hat{C}''(1-i/k) - g(1-i/k)\hat{C}(1-i/k) \\ c_j & = & f(1-j/k) \\ & & \\ R_i & = & \hat{R}''(1-i/k) + g(1-i/k)\hat{C}(1-i/k) \\ & & +kf(1-i/k)d_C(1-i/k) \\ r_j & = & \dfrac{1}{1-j/k} \\ & & \\ M_i & = & \hat{M}''(1-i/k) \\ & & -\left(2g(1-i/k) + f(1-i/k)^2\right)\hat{M}(1-i/k) \\ L_i & = & \hat{L}''(1-i/k) - 2g(1-i/k)\hat{L}(1-i/k) \\ & & +\left(g(1-i/k) + f(1-i/k)^2\right)\hat{M}(1-i/k) \\ & & +kf(1-i/k)d_M(1-i/k) \\ & & -\dfrac{1}{1+\epsilon}f(1-i/k)\hat{C}(1-i/k) - \dfrac{k}{1+\epsilon}d_C(1-i/k). \end{array}$$

Proof: These expressions are obtained by the same method that we are now following to obtain a more precise approximation of $N(u)$. For a sample derivation, see Appendix I. ■

The next step is to write a recursion for $N(x)$ which is exact up to terms of the order of $1/k^3$. We then approximate $N(x)$ by $\tilde{N}(x)$ which satisfies the same recursion except that we neglect terms of the order of $1/k^3$:

$$\begin{aligned} \tilde{N}(x) - \tilde{N}(x-1/k) = {} & \left(\frac{2}{kx} - \frac{1}{k^2x^2}\right)\tilde{N}(x) - \frac{1}{k^2}f(x)^2\hat{M}(x) \\ & + \left(-\frac{2}{k}f(x) + \frac{4}{k^2}g(x)\right)\hat{L}(x) + \frac{2}{k}f(x)d_L(x) \\ & + \frac{2}{(1+\epsilon)k^2}f(x)\hat{C}(x) + \left(\frac{2}{(1+\epsilon)k} - \frac{2}{(1+\epsilon)k^2x}\right)\hat{R}(x) \\ & - \frac{2}{(1+\epsilon)k}d_R(x) - \frac{2}{(1+\epsilon)^2k^2}. \end{aligned}$$

Claim 3: For any x on which $N(x)$ is defined, let $\tilde{d}_N(x) = \tilde{N}(x) - N(x)$ denote the discrepancy introduced by approximating $N(x)$ by $\tilde{N}(x)$. Then $\tilde{d}_N(x)$ is of the order of $1/k^2$.

Proof: See Appendix D. ■

We further approximate $\tilde{N}(x)$ by $\hat{N}(x)$ which satisfies the differential equation (16) and is given by expression (17). A more careful analysis of the discrepancy beween $\hat{N}(x)$ and $\tilde{N}(x)$ leads to the following claim:

Claim 4: For any x on which $\tilde{N}(x)$ is defined, $\tilde{N}(x)$ and $\hat{N}(x)$ differ by a term of the order of $1/k$.
More precisely,

$$\hat{N}(x) - \tilde{N}(x) = d_N(x),$$

where

$$\begin{aligned} d_N(x) = \frac{1}{k^2} \sum_{i=0}^{k(1-x)-1} \Bigg[& \hat{N}''(1-i/k) - \frac{1}{(1-i/k)^2}\hat{N}(1-i/k) \\ & - f(1-i/k)^2\hat{M}(1-i/k) + 4g(1-i/k)\hat{L}(1-i/k) \\ & + 2kf(1-i/k)d_L(1-i/k) + \frac{2f(1-i/k)}{(1+\epsilon)}\hat{C}(1-i/k) \\ & - \frac{2}{(1+\epsilon)(1-i/k)}\hat{R}(1-i/k) - \frac{2k}{1+\epsilon}d_R(1-i/k) \\ & - \frac{2}{(1+\epsilon)^2}\Bigg] \cdot \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{2}{k(1-j/k)}\right) + O(1/k^2). \end{aligned} \tag{18}$$

Proof: See Appendix E. ■

Let $d_N(x) = \hat{N}(x) - N(x)$ denote the overall discrepancy introduced by approximating $N(x)$ by $\hat{N}(x)$. Clearly,

$$d_N(x) = \tilde{d}_N(x) + \hat{d}_N(x),$$

where $\tilde{d}_N(x)$ and $\hat{d}_N(x)$ are as defined in Claims 3 and 4, respectively. By these claims, we thus have

$$N(x) = \hat{N}(x) - \hat{d}_N(x) + O(1/k^2),$$

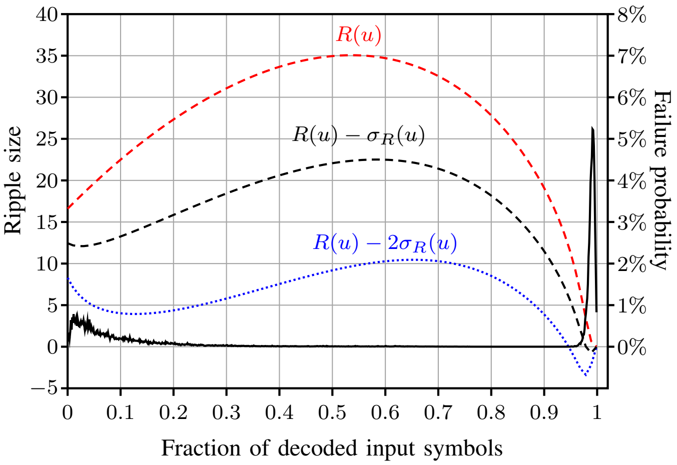

Fig. 1. Ripple size expectation and standard deviation versus the fraction of decoded input symbols. The solid line is the empirical failure probability of the decoder based on 100 million simulations. It confirms that the "problem zones" of the decoder are the ones predicted by the second moment method.

where $\hat{N}(x)$ is given by equation (17) and $\hat{d}_N(x)$ by equation (18). Using the resulting expression for $N(u)$, and the expression for $R(u)$ given by Definition 1, we finally obtain an expression for the variance of the ripple size up to terms of constant order.

Theorem 9: Consider an LT code with parameters $(k, \Omega(x))$ and overhead ϵ and let $\sigma_R^2(u)$ be the variance of the ripple size throughout BP decoding. Then

$$\begin{aligned}\sigma_R^2(u) = (1+\epsilon)u\left(\Omega'(1-u/k) + \frac{1}{1+\epsilon}\ln\frac{u}{k}\right)&\\ \cdot\left(1 + 2(1+\epsilon)k d_R(u/k)\right)&\\ - (1+\epsilon)\frac{u^2}{k}\Omega'(1-u/k)^2&\\ - (1+\epsilon)^2 k^2 d_N(u/k) + O(1).&\end{aligned}$$

Figure 1 shows a plot of the expected ripple size and the functions $h_1(u)$ and $h_2(u)$ given by equation (1), throughout the decoding process, for an LT code with $k = 800$ and $\epsilon = 0.1$, and with the degree distribution

$$\Omega(x) = \frac{1}{\frac{1}{50} + \sum_{i=2}^{50}\frac{1}{i(i-1)}}\left[\frac{1}{50}x + \sum_{i=2}^{50}\frac{1}{i(i-1)}x^i\right],$$

inspired from Luby's Ideal Soliton distribution [1]. The plot also shows the result of real simulations of this code, and confirms that the problem zones of the decoder are those predicted by the functions $h_i(u)$: the closer they are to the x-axis, the more probable it is that the decoder fails. As can be seen, there is a fair chance that the decoder fails when the fraction of decoded input symbols is between 0 and 0.2, and there is a very good chance that the decoder fails when the fraction of decoded input symbols is close to 0.95.

V. Conclusion

We have given an analytic expression for the variance of the ripple size throughout the LT decoding process. This expression is asymptotically of the order of k, and we have expressed it as a function of k as a first step toward finite-length analysis of the LT decoding. The next step is to work around the assumption that u is a "constant fraction" of k. Then we would obtain a guarantee for successful decoding as a function of the LT code parameters and overhead for practical values of k. This would then allow us to solve the corresponding design problem, namely to choose degree distributions that would make the function $h_c(u)$ stay positive for as large a value of c as possible, for a fixed code length k.

Appendix A
Proof of Claim 1

Define the discrepancy function $\tilde{d}_N(x)$ between $N(x)$ and its approximation $\tilde{N}(x)$ as

$$\tilde{d}_N(x) = \tilde{N}(x) - N(x),$$

where x represents the fraction of undecoded symbols at any step in the decoding and hence takes only values that are integer fractions of k. Recall that $N(x)$ satisfies the recursion

$$\begin{aligned}N(x-1/k) = \left(1 - \frac{2}{kx}\right)N(x) + \frac{2}{k}f(x)\hat{L}(x)&\\ - \frac{2}{(1+\epsilon)k}\hat{R}(x) + O(1/k^2)&\end{aligned}$$

and $\tilde{N}(x)$ satisfies the recursion

$$\tilde{N}(x-1/k) = \left(1 - \frac{2}{kx}\right)\tilde{N}(x) + \frac{2}{k}f(x)\hat{L}(x) - \frac{2}{(1+\epsilon)k}\hat{R}(x), \tag{19}$$

with initial condition $\tilde{N}(1) = N(1)$. From the recursions for $N(x)$ and $\tilde{N}(x)$, we get that

$$\tilde{d}_N(x-1/k) = \left(1 - \frac{2}{kx}\right)e_N(x) + O(1/k^2).$$

Using this recursive expression for $\tilde{d}_N(x)$ together with the initial condition $\tilde{d}_N(1) = 0$, and noting that the multiplicative term $(1 - \frac{2}{kx})$ is of constant order, we see that $\tilde{d}_N(x)$ is bounded by a term of the order of $1/k$ for all x on which $N(x)$ is defined. ■

Appendix B
Proof of Claim 2

Here we must bound the discrepancy $\hat{d}_N(x) = \hat{N}(x) - \tilde{N}(x)$ introduced by approximating the discrete function $\tilde{N}(x)$ by the continuous function $\hat{N}(x)$. Recall that $\tilde{N}(x)$ satisfies the recursion given by equation (19), and that $\hat{N}(x)$ satisfies the differential equation

$$\hat{N}'(x) = \frac{2}{x}\hat{N}(x) - 2f(x)\hat{L}(x) + \frac{2}{1+\epsilon}\hat{R}(x) \tag{20}$$

with initial condition $\hat{N}(1) = \tilde{N}(1)$.
Using the Taylor expansion of the function $\hat{N}(x)$ around a given point x given by

$$\hat{N}(x-1/k) = \hat{N}(x) - \frac{1}{k}\hat{N}'(x) + O(1/k^2),$$

in conjunction with equations (19) and (20) gives a recursion for $\hat{d}_N(x)$:

$$\hat{d}_N(x-1/k) = \left(1 - \frac{2}{kx}\right)\hat{d}_N(x) + O(1/k^2).$$

This recursion, together with the initial condition $\hat{d}_N(x) = 0$, implies that $\hat{d}_N(x)$ is bounded by a term of the order of $1/k$ for all x on which $\tilde{N}(x)$ is defined. ■

Appendix C
Proof of Theorem 6

$\hat{N}(x)$ satisfies the differential equation

$$\hat{N}'(x) = \frac{2}{x}\hat{N}(x) - 2f(x)\hat{L}(x) + \frac{2}{1+\epsilon}\hat{R}(x)$$

with initial condition $\hat{N}(1) = \tilde{N}(1)$. The general solution of this differential equation can be easily found by standard methods to be

$$\hat{N}(x) = x^2\Big(m_0\Omega'(1-x)^2 + 2l_0\Omega'(1-x) + \frac{2c_0}{1+\epsilon}\Omega'(1-x)\ln x + \frac{2r_0}{1+\epsilon}\ln x + \frac{1}{(1+\epsilon)^2}(\ln x)^2 + n_0\Big). \tag{21}$$

To find the value of the constant n_0, we use the initial condition $\hat{N}(1) = \tilde{N}(1) = N(1)$. The value of $N(1)$ can be found by looking at the beginning of the decoding process. $N(u)$ was defined as

$$N(u) := \frac{\partial^2 P_u}{\partial y^2}(1,1),$$

so that

$$N(u=k) = \sum_{c\geq 0, r\geq 1}(r-1)(r-2)p_{c,r,u=k}.$$

When there are k undecoded symbols, the coefficients $p_{c,r,k}$ are given by

$$p_{c,r,k} = \begin{cases} \binom{n}{c}\Omega_1^r(1-\Omega_1)^c & \text{if } c+r=n \\ 0 & \text{otherwise.}\end{cases}$$

Then

$$\begin{aligned} N(u=k) &= \sum_{r=1}^{n}(r-1)(r-2)\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} \\ &= \sum_{r=2}^{n} r(r-1)\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} \\ &\quad -2\sum_{r=1}^{n} r\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} \\ &\quad +2\sum_{r=1}^{n}\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r}. \end{aligned}$$

Now

$$\begin{aligned} &\sum_{r=2}^{n} r(r-1)\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} \\ &\quad = n(n-1)\Omega_1^2\sum_{r=0}^{n-2}\binom{n-2}{r}\Omega_1^r(1-\Omega_1)^{n-2-r} \\ &\quad = n(n-1)\Omega_1^2. \end{aligned}$$

Similarly,

$$\begin{aligned} &\sum_{r=1}^{n} r\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} \\ &\quad = n\Omega_1\sum_{r=0}^{n-1}\binom{n-1}{r}\Omega_1^r(1-\Omega_1)^{n-1-r} \\ &\quad = n\Omega_1, \end{aligned}$$

and

$$\sum_{r=1}^{n}\binom{n}{r}\Omega_1^r(1-\Omega_1)^{n-r} = 1-(1-\Omega_1)^n.$$

Normalizing and using the initial condition $\hat{N}(x=1) = N(x=1)$, we get

$$\hat{N}(x=1) = \left(1-\frac{1}{n}\right)\Omega_1^2 - \frac{2}{n}\Omega_1 + \frac{2}{n^2}\left(1-(1-\Omega_1)^n\right).$$

Evaluating the expression for $\hat{N}(x)$ given by equation (21) at $x=1$, and equating it to the above expression, we get

$$m_0\Omega_1^2 + 2l_0\Omega_1 + n_0 = \left(1-\frac{1}{n}\right)\Omega_1^2 - \frac{2}{n}\Omega_1 + \frac{2}{n^2}\left(1-(1-\Omega_1)^n\right),$$

where the values of m_0 and l_0 were already found to be

$$\begin{aligned} m_0 &= \left(1-\frac{1}{n}\right)\left(1-(1-\Omega_1)^{n-2}\right) \\ l_0 &= \frac{-1}{n} + (1-\Omega_1)^{n-2}\left(\Omega_1 + \frac{1-2\Omega_1}{n}\right). \end{aligned}$$

Solving for n_0, we finally obtain

$$n_0 = \frac{2}{n^2}\left(1-(1-\Omega_1)^n\right) - (1-\Omega_1)^{n-2}\left(\Omega_1^2 + \frac{2\Omega_1 - 3\Omega_1^2}{n}\right).$$

■

Appendix D
Proof of Claim 3

As in the proof of Claim 1, we define the discrepancy $\tilde{d}_N(x)$ between $N(x)$ and its approximation $\tilde{N}(x)$ as

$$\tilde{d}_N(x) = \tilde{N}\left(x\right) - N\left(x\right).$$

We would like to get a finer expression for $\tilde{d}_N(x)$. To this end, we use the recursions

$$\begin{aligned} N(x-1/k) = &\left(1-\frac{2}{kx}+\frac{1}{k^2x^2}\right)N(x) + \frac{1}{k^2}f(x)^2\hat{M}(x) \\ &+ \left(\frac{2}{k}f(x) - \frac{4}{k^2}g(x)\right)\hat{L}(x) - \frac{2}{k}f(x)d_L(x) \\ &- \frac{2}{(1+\epsilon)k^2}f(x)\hat{C}(x) - \frac{2}{(1+\epsilon)k}\left(1-\frac{1}{kx}\right)\hat{R}(x) \\ &+ \frac{2}{(1+\epsilon)k}d_R(x) + \frac{2}{(1+\epsilon)^2k^2} + O(1/k^3) \end{aligned}$$

and

$$\begin{aligned} \tilde{N}(x-1/k) = &\left(1-\frac{2}{kx}+\frac{1}{k^2x^2}\right)\tilde{N}(x) + \frac{1}{k^2}f(x)^2\hat{M}(x) \\ &+ \left(\frac{2}{k}f(x) - \frac{4}{k^2}g(x)\right)\hat{L}(x) - \frac{2}{k}f(x)d_L(x) \\ &- \frac{2}{(1+\epsilon)k^2}f(x)\hat{C}(x) - \frac{2}{(1+\epsilon)k}\left(1-\frac{1}{kx}\right)\hat{R}(x) \\ &+ \frac{2}{(1+\epsilon)k}d_R(x) + \frac{2}{(1+\epsilon)^2k^2} \end{aligned} \tag{22}$$

for $N(x)$ and $\tilde{N}(x)$, respectively, to get a recursion for the discrepancy:

$$\tilde{d}_N(x-1/k) = \left(1 - \frac{2}{kx} + \frac{1}{k^2x^2}\right)\tilde{d}_N(x) + O(1/k^3).$$

Together with the initial condition $\hat{N}(1) = N(1)$, which implies $\tilde{d}_N(1) = 0$, we get that $\tilde{d}_N(x)$ is bounded by a term of the order of $1/k^2$ for all the values x takes during the decoding process.

■

Appendix E
Proof of Claim 4

As in the proof of Claim 2, we define $\hat{d}_N(x)$ to be the discrepancy introduced by approximating the discrete function $\tilde{N}(x)$ by the continuous function $\hat{N}(x)$. This time, we would like to obtain a finer expression for $\hat{d}_N(x)$. We will thus consider the recursion for $\tilde{N}(x)$ given by equation (22), and we will write the Taylor expansion of $\hat{N}(x)$ up to terms of the order of $1/k^3$:

$$\hat{N}(x-1/k) = \hat{N}(x) - \frac{1}{k}\hat{N}'(x) + \frac{1}{k^2}\hat{N}''(x) + O(1/k^3). \quad (23)$$

Recall that $\hat{N}(x)$ satisfies the differential equation

$$\hat{N}'(x) = \frac{2}{x}\hat{N}(x) - 2f(x)\hat{L}(x) + \frac{2}{1+\epsilon}\hat{R}(x).$$

Plugging this expression for $\hat{N}'(x)$ into the Taylor expansion of $\hat{N}(x)$ and using the resulting expression in conjunction with the recursion for $\tilde{N}(x)$ of equation (22) gives us a recursion for $\hat{d}_N(x)$:

$$\begin{aligned}\hat{d}_N(x-1/k) = \left(1-\frac{2}{kx}\right)\hat{d}_N(x) + \frac{1}{k^2}\Big[\hat{N}''(x) - \frac{1}{x^2}\hat{N}(x)\\ - f(x)^2\hat{M}(x) + 4g(x)\hat{L}(x) + 2kf(x)d_L(x)\\ + \frac{2f(x)}{(1+\epsilon)}\hat{C}(x) - \frac{2}{(1+\epsilon)x}\hat{R}(x)\\ - \frac{2k}{1+\epsilon}d_R(x) - \frac{2}{(1+\epsilon)^2}\Big] + O(1/k^3).\end{aligned}$$

This recursion can be easily seen to yield the following closed-form expression for $\hat{d}_N(x)$, up to a term of the order of $1/k^2$ representing an accumulation of terms of the order of $1/k^3$:

$$\hat{d}_N(x) = \frac{1}{k^2}\sum_{i=0}^{k(1-x)-1} N_i \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{n_j}{k}\right) + O(1/k^2), \quad (24)$$

where n_j is given by $n_j = \frac{2}{1-j/k}$ and N_i is given by

$$\begin{aligned}N_i = \hat{N}''(1-i/k) - \frac{1}{(1-i/k)^2}\hat{N}(1-i/k)\\ - f(1-i/k)^2\hat{M}(1-i/k) + 4g(1-i/k)\hat{L}(1-i/k)\\ + 2kf(1-i/k)d_L(1-i/k) + \frac{2f(1-i/k)}{(1+\epsilon)}\hat{C}(1-i/k)\\ - \frac{2}{(1+\epsilon)(1-i/k)}\hat{R}(1-i/k) - \frac{2k}{1+\epsilon}d_R(1-i/k)\\ - \frac{2}{(1+\epsilon)^2}.\end{aligned}$$

Note that we have replaced $\tilde{N}(1-i/k)$ by $\hat{N}(1-i/k)$ in the expression for N_i; this introduces an error that is accounted for by the $O(1/k^2)$ term in the expression for $\hat{d}_N(x)$ in equation (24). This is crucial in that it allows us to get an analytic expression for the discrepancy $\hat{d}_N(x)$ between $\tilde{N}(x)$ and $\hat{N}(x)$, and thus for the discrepancy between $N(x)$ and its continuous approximation $\hat{N}(x)$. ■

Appendix F
Proof of Theorem 3

Recall that

$$M(u) = \frac{\partial^2 P_u}{\partial x^2}(1,1).$$

By differentiating both sides of the recursion (2) twice with respect to x and evaluating at $(1,1)$, we get the following recursion for $M(u)$:

$$M(u-1) = (1-p_u)^2 M(u) - (1-p_u)^2\frac{\partial^2 P_u}{\partial x^2}\left(1-p_u, \frac{1}{u}\right).$$

By a similar analysis to that of the proof of Theorem 5 (Appendix H), it can easily be shown that under the assumption $r > 5$, we can bound the dirt term

$$(1-p_u)^2\frac{\partial^2 P_u}{\partial x^2}\left(1-p_u, \frac{1}{u}\right)$$

by a term of constant order, thus obtaining the following difference equation for $M(u)$:

$$M(u) - M(u-1) = (2p_u - p_u^2)M(u) + O(1).$$

Normalizing by n^2, we obtain a difference equation for $M(x)$:

$$M(x) - M(x-1/k) = \frac{2}{k}f(x)M(x) + O(1/k^2).$$

Neglecting lower-order terms, we approximate $M(x)$ by the function $\tilde{M}(x)$ which satisfies

$$\tilde{M}(x) - \tilde{M}(x-1/k) = \frac{2}{k}f(x)\tilde{M}(x),$$

with initial condition $\tilde{M}(1) = M(1)$.
We state the following claim without proof, as its proof is very similar to that of Claim 1.

Claim 5: For any x on which $M(x)$ is defined, $M(x)$ and $\tilde{M}(x)$ differ by a term of the order of $1/k$.

We further approximate $\tilde{M}(x)$ by the continuous function $\hat{M}(x)$ which satisfies the differential equation

$$\hat{M}'(x) = 2f(x)\hat{M}(x)$$

with initial condition $\hat{M}(x) = \tilde{M}(x)$. The general solution of this differential equation is of the form

$$\hat{M}(x) = m_0\left(1 - x\Omega'(1-x) - \Omega(1-x)\right)^2. \quad (25)$$

The value of the constant m_0 can be found from the initial condition $\hat{M}(1) = \tilde{M}(1) = M(1)$. Looking at the beginning of the decoding process, namely at the step $u = k$, we can see that

$$\begin{aligned} M(u=k) &= \sum_{c\geq 0, r\geq 1} c(c-1)p_{c,r,u=k} \\ &= \sum_{c=0}^{n-1} c(c-1)\binom{n}{c}\Omega_1^{n-c}(1-\Omega_1)^c \\ &= n(n-1)\sum_{c=0}^{n-3}\binom{n-2}{c}\Omega_1^{n-2-c}(1-\Omega_1)^{c+2} \\ &= n(n-1)(1-\Omega_1)^2\left(1-(1-\Omega_1)^{n-2}\right). \end{aligned}$$

Normalizing, we get that

$$\hat{M}(1) = \left(1-\frac{1}{n}\right)(1-\Omega_1)^2\left(1-(1-\Omega_1)^{n-2}\right).$$

On the other hand, from equation (25),

$$\hat{M}(1) = m_0(1-\Omega_1)^2.$$

Equating the two expressions, we finally get

$$m_0 = \left(1-\frac{1}{n}\right)\left(1-(1-\Omega_1)^{n-2}\right).$$

The expression obtained for $\hat{M}(x)$ is a good approximation for $\tilde{M}(x)$, as the following claim shows. Again, we state it without proof, as its proof is very similar to that of Claim 2.

Claim 6: For any x on which $\tilde{M}(x)$ is defined, $\tilde{M}(x)$ and $\hat{M}(x)$ differ by a term of the order of $1/k$.

Appendix G
Proof of Theorem 4

Recall that

$$L(u) = \frac{\partial^2 P_u}{\partial x \partial y}(1,1).$$

By differentiating both sides of the recursion (2) and evaluating at $(1,1)$, we obtain a recursion for $L(u)$:

$$\begin{aligned} L(u-1) = p_u(1-p_u)M(u) + \left(1-\frac{1}{u}\right)(1-p_u)L(u) \\ -(1-p_u)C(u) + (1-p_u)\frac{\partial P_u}{\partial x}\left(1-p_u, \frac{1}{u}\right). \end{aligned}$$

Again, by an analysis similar to that of Appendix H nder the assumption $r > 4$, the dirt term

$$(1-p_u)\frac{\partial P_u}{\partial x}\left(1-p_u, \frac{1}{u}\right)$$

is of constant order, so that we have the following difference equation for $L(u)$:

$$\begin{aligned} L(u) - L(u-1) = \left(\frac{1}{u} + p_u - \frac{p_u}{u}\right)L(u) - p_u(1-p_u)M(u) \\ + (1-p_u)C(u) + O(1). \end{aligned}$$

Again, we seek a difference equation for the normalized function $L(x)$. We use the approximations $C(x) = \hat{C}(x) + O(1/k)$, $M(x) = \hat{M}(x) + O(1/k)$, and $p_u = \frac{1}{k}f(x) + O(1/k^2)$ to obtain a difference equation for $L(x)$:

$$\begin{aligned} L(x) - L(x-1/k) = \left(\frac{1}{kx} + \frac{1}{k}f(x)\right)L(x) - \frac{1}{k}f(x)\hat{M}(x) \\ + \frac{1}{k(1+\epsilon)}\hat{C}(x) + O(1/k^2). \end{aligned}$$

Neglecting lower-order terms, we approximate $L(x)$ by the function $\tilde{L}(x)$ which satisfies

$$\begin{aligned} \tilde{L}(x) - \tilde{L}(x-1/k) = \left(\frac{1}{kx} + \frac{1}{k}f(x)\right)\tilde{L}(x) - \frac{1}{k}f(x)\hat{M}(x) \\ + \frac{1}{k(1+\epsilon)}\hat{C}(x) \end{aligned}$$

with initial condition $\tilde{L}(1) = L(1)$.
$\tilde{L}(x)$ is a good approximation for $L(x)$, as the following claim (stated without proof, since its proof is similar to that of Claim 1) shows.

Claim 7: For any x on which $L(x)$ is defined, $L(x)$ and $\tilde{L}(x)$ differ by a term of the order of $1/k$.

We further approximate $\tilde{L}(x)$ by the continuous function $\hat{L}(x)$ which satisfies the differential equation

$$\hat{L}'(x) = \left(f(x) + \frac{1}{x}\right)\hat{L}(x) - f(x)\hat{M}(x) + \frac{1}{1+\epsilon}\hat{C}(x)$$

with initial condition $\hat{L}(x) = \tilde{L}(x)$. The general solution of this differential equation is of the form

$$\begin{aligned} \hat{L}(x) = x\left(1 - x\Omega'(1-x) - \Omega(1-x)\right) \\ \cdot\left(m_0\Omega'(1-x) + \frac{c_0}{1+\epsilon}\ln x + l_0\right) \end{aligned} \quad (26)$$

where the values of c_0 and m_0 are given by equations (5) and (12) respectively. The value of l_0 can be found using the initial condition $\hat{L}(1) = \tilde{L}(1) = L(1)$. Looking at the initial step $u = k$ of the decoding process, we see that

$$\begin{aligned} L(u=k) &= \sum_{c\geq 0, r\geq 1} c(r-1)p_{c,r,k} \\ &= \sum_{c=0}^{n-1} c(n-c-1)\binom{n}{c}\Omega_1^{n-c}(1-\Omega_1)^c. \end{aligned}$$

Note that $c(n-c) = c(n-2) - c(c-1)$, so that $L(u=k)$ becomes

$$\begin{aligned} L(u=k) = (n-2)\sum_{c=0}^{n-1} c\binom{n}{c}\Omega_1^{n-c}(1-\Omega_1)^c \\ -\sum_{c=0}^{n-1} c(c-1)\binom{n}{c}\Omega_1^{n-c}(1-\Omega_1)^c \end{aligned}$$

$$\begin{aligned} &= n(n-2)(1-\Omega_1)\left(1-(1-\Omega_1)^{n-1}\right) \\ &\quad - n(n-1)(1-\Omega_1)^2\left(1-(1-\Omega_1)^{n-2}\right) \end{aligned}$$

$$= n(n-1)(1-\Omega_1)\Omega_1 - n(1-\Omega_1)\left(1-(1-\Omega_1)^{n-1}\right).$$

Normalizing and equating to the evaluation of equation (26) at $x = 1$, we can solve for l_0 as

$$l_0 = \frac{-1}{n} + (1-\Omega_1)^{n-2}\left(\Omega_1 + \frac{1-2\Omega_1}{n}\right).$$

The following claim (whose proof is similar to that of Claim 2) shows that $\hat{L}(x)$ is a good approximation for $\tilde{L}(x.)$

Claim 8: For any x on which $\tilde{L}(x)$ is defined, $\tilde{L}(x)$ and $\hat{L}(x)$ differ by a term of the order of $1/k$.

Appendix H
Proof of Theorem 5

We will prove that $1 - P_u(1,1) + P_u\left(1-p_u, \frac{1}{u}\right)$ can be upper bounded by a term of the order of $1/k$. To see this, note that

$$\begin{aligned}
1 - P_u(1,1) &+ P_u\left(1-p_u, \frac{1}{u}\right) \\
&= \sum_{\substack{c\geq 0\\ r\geq 0}} p_{c,r,u} - \sum_{\substack{c\geq 0\\ r\geq 1}} p_{c,r,u} + \sum_{\substack{c\geq 0\\ r\geq 1}} p_{c,r,u}(1-p_u)^c\left(\frac{1}{u}\right)^{r-1} \\
&= \sum_{\substack{c\geq 0\\ r=0}} p_{c,r,u} + \sum_{\substack{c\geq 0\\ r\geq 1}} p_{c,r,u}(1-p_u)^c\left(\frac{1}{u}\right)^{r-1} \\
&\leq \sum_{\substack{c\geq 0\\ r=0}} p_{c,r,u} + \sum_{\substack{c\geq 0\\ r\geq 1}} \left(\frac{1}{u}\right)^{r-1} \\
&= \sum_{\substack{c\geq 0\\ r=0}} p_{c,r,u} + \sum_{\substack{c\geq 0\\ r=1,\ldots,3}} \left(\frac{1}{u}\right)^{r-1} + \sum_{\substack{c\geq 0\\ r\geq 4}} \left(\frac{1}{u}\right)^{3}\left(\frac{1}{u}\right)^{r-5}.
\end{aligned}$$

For $r \geq 4$, the first two summands vanish. As for the third summand, it is the sum of $O(k^2)$ terms, each of them being $O(1/k^3)$, so that their sum is $O(1/k)$.

Thus

$$2\left(1 - P_u(1,1) + P_u\left(1-p_u, \frac{1}{u}\right)\right) = 2 + O(1/k) \text{ for } r \geq 4.$$

Appendix I
Calculation of the Discrepancy Terms of Theorem 8

Here we show sample derivations of some of the discrepancy terms of Theorem 8. We will present detailed calculations of the discrepancy terms $d_C(x)$ and $d_R(x)$; very similar calculations can be carried out for $d_M(x)$ and $d_L(x)$.

A. An Expression for $d_C(x)$

We start by the derivation of the discrepancy term $d_C(x)$ introduced by approximating $C(x)$ by an analytic function. For convenience, we restate the expression we wish to obtain in the following lemma.

Lemma 1: The discrepancy term $d_C(x)$ defined as $d_C(x) = \hat{C}(x) - C(x)$ is given by the expression

$$d_C(x) = \frac{1}{k^2}\sum_{i=0}^{k(1-x)-1} C_i \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{c_j}{k}\right) + O(1/k^2),$$

where

$$C_i = \hat{C}''(1-i/k) - g(1-i/k)\hat{C}(1-i/k)$$

and

$$c_j = f(1-j/k).$$

Before we prove Lemma 1, we follow a procedure similar to that of Section IV and Appendices F and G to derive a closed-form expression for $C(x)$ up to lower-order terms. Through this analysis, an expression for the discrepancy term $d_C(x)$ will naturally arise.

Recall that

$$C(u) = \sum_{c\geq 0, r\geq 1} c p_{c,r,u} = \frac{\partial P_u}{\partial x}(1,1).$$

Differentiating both sides of the recursion (2) with respect to x and evaluating at $(1,1)$, we obtain a recursion for $C(u)$, given by

$$C(u-1) = (1-p_u)C(u) - (1-p_u)\frac{\partial P_u}{\partial x}\left(1-p_u, \frac{1}{u}\right).$$

A simple analysis, similar to that of Appendix H, shows that the dirt term

$$(1-p_u)\frac{\partial P_u}{\partial x}\left(1-p_u, \frac{1}{u}\right)$$

can be bounded by a term of the order of $1/k^2$, for $r \geq 6$. In what follows, we thus make the assumption that the ripple size does not go below 6, so that we can write

$$C(u-1) = (1-p_u)C(u) + O(1/k^2).$$

We normalize $C(u)$ by n in order to work with expressions of constant order, and obtain the following difference equation for $C(x)$:

$$C(x) - C(x-1/k) = \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)C(x) + O(1/k^3),$$

where expressions for the functions f and g are given by equations (3) and (4), respectively.
As a first step, we approximate $C(x)$ by the function $\tilde{C}(x)$, which satisfies

$$\tilde{C}(x) - \tilde{C}(x-1/k) = \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)\tilde{C}(x),$$

with initial condition $\tilde{C}(1) = C(1)$. We further approximate $\tilde{C}(x)$ by the continuous function $\hat{C}(x)$, which satisfies the differential equation

$$\hat{C}'(x) = f(x)\hat{C}(x)$$

with initial condition $\hat{C}(1) = \tilde{C}(1) = C(1)$. The solution of this differential equation is readily seen to be of the form

$$\hat{C}(x) = c_0\left(1 - x\Omega'(1-x) - \Omega(1-x)\right),$$

where the value of the constant c_0 is to be determined by the initial condition $c_0(1-\Omega_1) = C(1)$. To obtain an expression for $C(1)$, we look at the beginning of the decoding process and note that

$$\begin{aligned} C(u=k) &= \sum_{c\geq 1, r\geq 1} p_{c,r,k} \\ &= \sum_{c=1}^{n-1} c\binom{n}{c}\Omega_1^{n-c}(1-\Omega_1)^c \\ &= n(1-\Omega_1)\sum_{c=0}^{n-2}\binom{n-1}{c}\Omega_1^{n-1-c}(1-\Omega_1)^c \\ &= n(1-\Omega_1)\left(1-(1-\Omega_1)^{n-1}\right), \end{aligned}$$

so that

$$C(x=1) = (1-\Omega_1)\left(1-(1-\Omega_1)^{n-1}\right).$$

This gives us

$$c_0 = 1-(1-\Omega_1)^{n-1}.$$

We thus obtain an analytic approximation $\hat{C}(x)$ of $C(x)$. We now turn to obtaining and bounding an expression for the error introduced by this approximation.

Proof of Lemma 1: We start by bounding the discrepancy $\tilde{d}_C(x) = \tilde{C}(x) - C(x)$ introduced by approximating $C(x)$ by $\tilde{C}(x)$. $C(x)$ satisfies the recursion

$$C(x-1/k) = \left(1 - \frac{1}{k}f(x) + \frac{1}{k^2}g(x)\right)C(x) + O(1/k^3),$$

whereas $\tilde{C}(x)$ satisfies the recursion

$$\tilde{C}(x-1/k) = \left(1 - \frac{1}{k}f(x) + \frac{1}{k^2}g(x)\right)\tilde{C}(x), \quad (27)$$

with initial condition $\tilde{C}(1) = C(1)$. This gives a recursion for $\tilde{d}_C(x)$, as follows:

$$\tilde{d}_C(x-1/k) = \left(1 - \frac{1}{k}f(x) + \frac{1}{k^2}g(x)\right)\tilde{d}_C(x) + O(1/k^3),$$

with initial condition $\tilde{d}_C(1) = 0$. Noting that the multiplicative term $\left(1 - \frac{1}{k}f(x) + \frac{1}{k^2}g(x)\right)$ is of constant order, we can thus write, for any x on which $\tilde{d}_C(x)$ is defined,

$$\tilde{d}_C(x) = O(1/k^2). \quad (28)$$

We now seek an expression for the discrepancy $\hat{d}_C(x) = \hat{C}(x) - \tilde{C}(x)$ introduced by approximating $\tilde{C}(x)$ by the continuous function $\hat{C}(x)$. $\tilde{C}(x)$ satisfies the recursion given by equation (27), and $\hat{C}(x)$ satisfies the differential equation

$$\hat{C}'(x) = f(x)\hat{C}(x). \quad (29)$$

We write the Taylor expansion of $\hat{C}(x)$ up to terms of the order of $1/k^3$ as

$$\hat{C}(x-1/k) = \hat{C}(x) - \frac{1}{k}\hat{C}'(x) + \frac{1}{k^2}\hat{C}''(x) + O(1/k^3)$$

and plug the expression for $\hat{C}'(x)$ given by the differential equation (29) into this expansion; together with the recursion for $\tilde{C}(x)$, this gives us the following recursion for $\hat{d}_C(x)$:

$$\begin{aligned} \hat{d}_C(x-1/k) = &\left(1 - \frac{1}{k}f(x)\right)\hat{d}_C(x) \\ &+ \frac{1}{k^2}\hat{C}''(x) - \frac{1}{k^2}g(x)\tilde{C}(x) + O(1/k^3). \end{aligned}$$

This recursion is clearly seen to yield the closed-form expression

$$\hat{d}_C(x) = \frac{1}{k^2}\sum_{i=0}^{k(1-x)-1} C_i \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{c_j}{k}\right) + O(1/k^2), \quad (30)$$

with C_i and c_j as defined in the statement of Lemma 1. Note that the expression for C_i is a function of $\hat{C}(1-i/k)$ instead of $\tilde{C}(1-i/k)$. This introduces an error accounted for by the $O(1/k^2)$ term in equation (30).

Putting together the expressions for the discrepancies $\tilde{d}_C(x)$ and $\hat{d}_C(x)$ respectively given by equations (28) and (30), we finally obtain the closed-form expression given by Lemma 1 for $d_C(x) = \tilde{d}_C(x) + \hat{d}_C(x)$. ■

B. An Expression for $d_R(x)$

We follow a similar procedure to that of the previous section to derive an analytic expression for the discrepancy $d_R(x)$ introduced by approximating $R(x)$ by the continuous function $\hat{R}(x)$. We want to prove the following lemma.

Lemma 2: The discrepancy term $d_R(x)$ defined as $d_R(x) = \hat{R}(x) - R(x)$ is given by the expression

$$d_R(x) = \frac{1}{k^2}\sum_{i=0}^{k(1-x)-1} R_i \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{r_j}{k}\right) + O(1/k^2),$$

where

$$\begin{aligned} R_i = &\hat{R}''(1-i/k) + g(1-i/k)\hat{C}(1-i/k) \\ &+ kf(1-i/k)d_C(1-i/k) \end{aligned}$$

and

$$r_j = \frac{1}{1-j/k}.$$

Again, before we prove the lemma, we derive a closed-form expression for $R(x)$ up to lower-order terms.

By definition,

$$R(u) = \sum_{c\geq 0, r\geq 1}(r-1)p_{c,r,u} = \frac{\partial P_u}{\partial y}(1,1).$$

Differentiating both sides of the recursion (2) with respect to y and evaluating at $(1,1)$, we obtain the recursion

$$R(u-1) = \left(1 - \frac{1}{u}\right)R(u) + p_u C(u) - P_u(1,1) + P_u\left(1 - p_u, \frac{1}{u}\right).$$

A similar analysis to that of Appendix H shows that the dirt term

$$-P_u(1,1) + P_u\left(1 - p_u, \frac{1}{u}\right)$$

can be approximated by $-1 + 1/k^2$, for $r \geq 5$. We make this assumption in what follows and can thus work with the following difference equation for $R(u)$:

$$R(u) - R(u-1) = \frac{1}{u}R(u) - p_u C(u) + 1 + O(1/k^2).$$

Normalizing $R(u)$ by n, we obtain a difference equation for $R(x)$:

$$R(x) - R(x-1/k) = \frac{1}{kx}R(x) - \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)\hat{C}(x) + \frac{1}{k}f(x)d_C(x) + \frac{1}{k(1+\epsilon)} + O(1/k^3),$$

where f is as given by equation (3). Note that we replaced $C(x)$ by its approximation $\hat{C}(x)$ and accounted for the resulting error.

We approximate $R(x)$ by $\tilde{R}(x)$, which satisfies

$$R(x) - R(x-1/k) = \frac{1}{kx}R(x) - \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)\hat{C}(x) + \frac{1}{k}f(x)d_C(x) + \frac{1}{k(1+\epsilon)},$$

with initial condition $\tilde{R}(1) = R(1)$, and further approximate $\tilde{R}(x)$ by the continuous function $\hat{R}(x)$, which satisfies the differential equation

$$\hat{R}'(x) = \frac{\hat{R}(x)}{x} - f(x)\hat{C}(x) + \frac{1}{k(1+\epsilon)}$$

with initial condition $\hat{R}(1) = \tilde{R}(1) = R(1)$. The general solution of this differential equation can be easily found by standard techniques to be

$$\hat{R}(x) = x\left(c_0\Omega'(1-x) + \frac{1}{1+\epsilon}\ln x + r_0\right),$$

where c_0 is given by equation (5) and the value of r_0 can be determined by the initial condition $c_0\Omega_1 + r_0 = R(1)$. For the value of $R(1)$, we look at the beginning of the decoding process, as we did previously in order to derive an expression for $C(1)$. By the same method we obtain

$$R(u=k) = n\Omega_1 - 1 + (1-\Omega_1)^n,$$

so that

$$r_0 = \Omega_1(1-\Omega_1)^{n-1} - \frac{1-(1-\Omega_1)^n}{n}.$$

We can now obtain an expression for the error introduced by approximating $R(x)$ by $\hat{R}(x)$.

Proof of Lemma 2: We first bound the discrepancy $\tilde{d}_R(x) = \tilde{R}(x) - R(x)$ introduced by approximating $R(x)$ by $\tilde{R}(x)$. $R(x)$ satisfies the recursion

$$R(x-1/k) = \left(1 - \frac{1}{kx}\right)R(x) + \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)\hat{C}(x) - \frac{1}{k}f(x)d_C(x) - \frac{1}{k(1+\epsilon)} + O(1/k^3),$$

whereas $\tilde{R}(x)$ satisfies the recursion

$$\tilde{R}(x-1/k) = \left(1 - \frac{1}{kx}\right)\tilde{R}(x) + \left(\frac{1}{k}f(x) - \frac{1}{k^2}g(x)\right)\hat{C}(x) - \frac{1}{k}f(x)d_C(x) - \frac{1}{k(1+\epsilon)} \qquad (31)$$

with initial condition $\tilde{R}(1) = R(1)$. This gives the following recursion for $\tilde{d}_R(x)$:

$$\tilde{d}_R(x-1/k) = \left(1 - \frac{1}{kx}\right)\tilde{d}_R(x) + O(1/k^3),$$

with initial condition $\tilde{d}_R(1) = 0$. Since the multiplicative term $\left(1 - \frac{1}{kx}\right)$ is of constant order, we can write, for any x on which $\tilde{d}_R(x)$ is defined,

$$\tilde{d}_R(x) = O(1/k^2). \qquad (32)$$

We now turn to finding an expression for the discrepancy $\hat{d}_R(x) = \hat{R}(x) - \tilde{R}(x)$ introduced by approximating $\tilde{R}(x)$ by the continuous function $\hat{R}(x)$. The Taylor expansion of $\hat{R}(x)$ up to terms of the order of $1/k^3$ is given by

$$\hat{R}(x-1/k) = \hat{R}(x) - \frac{1}{k}\hat{R}'(x) + \frac{1}{k^2}\hat{R}''(x) + O(1/k^3),$$

where we can replace $\hat{R}'(x)$ by its expression given by the differential equation

$$\hat{R}'(x) = \frac{\hat{R}(x)}{x} - f(x)\hat{C}(x) + \frac{1}{k(1+\epsilon)}. \qquad (33)$$

Using the resulting expression for $\hat{R}(x-1/k)$, together with the expression for $\tilde{R}(x-1/k)$ given by the recursion (31), we obtain a recursion for $\hat{d}_R(x)$:

$$\hat{d}_R(x-1/k) = \left(1 - \frac{1}{kx}\right)\hat{d}_R(x) + \frac{1}{k^2}\left(\hat{R}''(x) + g(x)\hat{C}(x)kf(x)d_C(x)\right) + O(1/k^3).$$

This recursion is easily seen to yield the closed-form expression

$$\hat{d}_R(x) = \frac{1}{k^2}\sum_{i=0}^{k(1-x)-1} R_i \prod_{j=i+1}^{k(1-x)-1}\left(1 - \frac{r_j}{k}\right) + O(1/k^2), \qquad (34)$$

where R_i and r_j are as defined in the statement of Lemma 2.

Together, equations (32) and (34) give us the closed-form expression for $d_R(x) = \tilde{d}_R(x) + \hat{d}_R(x)$ given by Lemma 2. ■

References

[1] M. Luby, "LT Codes," in *Proceedings of the ACM Symposium on Foundations of Computer Science (FOCS)*, 2002.

[2] A. Shokrollahi, "Raptor Codes," in *IEEE/ACM Trans. Netw.*, vol. 14, pp. 2551–2567, 2006.

[3] R. Karp, M. Luby, and A. Shokrollahi, "Finite Length Analysis of LT Codes," in *Proceedings of the International Symposium on Information Theory (ISIT),* 2004.