Rateless Distributed Source Code Design

Dino Sejdinović, Robert J. Piechocki and Angela Doufexi
Centre for Communications Research
Department of Electrical & Electronic Engineering
University of Bristol, Bristol, UK, BS8 1UB
{d.sejdinovic, r.j.piechocki, a.doufexi}@bristol.ac.uk

ABSTRACT

Over the past decade, rateless codes, i.e., digital fountain codes, have emerged as an efficient and robust solution for reliable data transmission over packet erasure networks and a particularly suitable one for multicasting and broadcasting applications where users may experience variable channel conditions and packet loss rates, such as mobile environments. Luby Transform (LT) and Raptor codes are practical fountain codes with a capacity approaching performance and a low computational cost. In addition to their channel coding applications, the use of fountain codes for various kinds of distributed source compression and distributed joint-source channel coding has been extensively studied lately, and with promising results. However, a systematic treatise of the code design and optimization considerations for such non-standard applications of fountain codes is still absent. In this contribution, we overview the main results concerned with rateless codes for distributed source coding and outline several examples of data dissemination protocols where carefully designed fountain codes can provide strikingly simple, yet robust solutions yielding both distributed source coding and channel coding gains.

1. INTRODUCTION

1.1 Fountain coding

Unlike traditional coding schemes, fountain codes are able to adapt their rate on-the-fly - they are rateless in the sense that a potentially limitless number of encoding symbols can be generated from the source data and the original message can be recovered from any sufficiently large set of encoding symbols [1]. The first practical fountain codes were Luby-Transform (LT) codes [2], whereas their extension, Raptor codes [3], are a state-of-the-art application layer forward error correction scheme with excellent performance and linear encoding/decoding complexity. Raptor codes have been standardized in the Third Generation mobile cellular wireless multimedia broadcast and multicast [4] and digital video

broadcasting to handheld devices [5]. It is possible to adapt LT and Raptor codes for transmission over noisy channels [6, 7, 8, 9], as they are more amenable to soft-decision decoding than classical erasure codes, such as Reed-Solomon codes. Even though it has been demonstrated that the universal code parameters generally do not exist for noisy channel models [9], fountain codes have shown promising performance over channels such as binary input additive white Gaussian noise channel (BIAWGNC) [9] and block-fading channels [8]. In addition, fountain codes have been used in the context of lossless data compression [10], distributed source coding [11, 12, 13, 14, 15], distributed joint source-channel coding [16, 17, 18], transmission of scalable video [19, 20, 21, 22] and transmission of multiple independent sources in relay networks [23].

1.2 Distributed source coding problem

Distributed source coding (DSC) problem deals with the compression of correlated sources. From the seminal Shannon's source coding theorem, we know that the lower bound on the total compression rate of two correlated sources X and Y is their joint entropy $H(X, Y)$. However, another celebrated result, by Slepian and Wolf [24], asserts that, remarkably, the separate compression (Slepian-Wolf Coding - SWC) suffers no rate loss compared to the case of joint compression. Namely, the region of the achievable compression rates is given by

$$
\begin{aligned}
R_X &\geq H(X|Y) \\
R_Y &\geq H(Y|X) \\
R_X + R_Y &\geq H(X, Y),
\end{aligned}
\tag{1}
$$

where R_X and R_Y are respectively the compression rates corresponding to sources X and Y.

The scenario in which source Y is fully known at the decoder, i.e., it is independently compressed using an entropy coder with rate $H(Y)$, while the source X is compressed using the estimate of the correlation between X and Y (with rate ideally close to $H(X|Y)$) is one of the simplest and the most widely studied cases of DSC problem. It is commonly referred to as the asymmetric SWC or coding with decoder side information.

Robust rateless distributed source code design in the asymmetric SWC would find applications in data synchronization scenarios [25], where each receiver typically has a possibly outdated version of some common database, as well as in cross-layer design of scalable video transmission [17]. A clear benefit of a rateless code construction for asymmetric SWC is that it simultaneously represents a distributed joint

source-channel coding scheme. Namely, rateless scheme would yield both the distributed source coding gains, by utilizing the presence of side information, and the channel coding gains, as it enables reliable transmission over lossy links.

Correlation between sources X and Y can take different forms. Indeed, a common way to model correlation is to view X and Y as, respectively, the input and the output of a certain communication channel, referred to as correlation channel or virtual channel. One of the simplest cases is the one where Y is a partial information about X, i.e., the output of a binary erasure channel (BEC) when X is the input.

In section 2, main relevant results on the design and analysis of fountain codes for channel coding are reviewed. In section 3, the use of rateless codes in coding with partial information is studied. Section 4 considers the problem of symmetric Slepian-Wolf coding where two transmitters have a number of data packets in common and section 5 studies the possibilities of soft decoding of fountain codes for use in more complicated correlation channel models. Section 6 concludes the paper.

2. FOUNTAIN CODES FOR CHANNEL CODING

Practical fountain codes, such as LT (Luby Transform) codes [2], are binary linear sparse graph codes [26], decoded by a message-passing belief propagation (BP) algorithm [27]. The message block $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ consists of k input symbols, i.e., information packets[1] $x_i \in \mathbb{F}_2^b$, $i \in N_k$, $b \in \mathbb{N}$, and the encoding packets are produced by bitwise XOR operation. The only two parameters of an LT code ensemble $LT(k, \Omega(x))$ are the block length k and a certain probability mass function Ω on set N_k, called output degree distribution [2]. Output degree distribution is typically identified with its generating polynomial $\Omega(x) = \sum_{d=1}^{k} \Omega_d x^d$, where Ω_d is the probability of degree $d \in N_k$. The generation of a single LT encoding packet consists of two simple steps which can be performed as many times as necessary in order to produce enough encoding packets for successful decoding: (a) Sample an output degree d with probability Ω_d, (b) Sample d distinct packets from the message block uniformly at random and XOR them. Good LT codes require the average degree $\mu = \Omega'(1)$ of degree distribution $\Omega(x)$ to grow at least as $\mathcal{O}(\log k)$ (otherwise large error floors occur in the waterfall region of the decoder) and their computational complexity is $\mathcal{O}(k \log k)$. The performance of a fountain code ensemble is typically measured in terms of the average packet error rate, i.e., probability that an information packet in the message block cannot be recovered by the decoder, as the function of *code overhead* $\varepsilon = n/k - 1$, where n is the number of encoding packets observed at the receiver.

Raptor codes [3] are a modification of LT codes obtained by precoding the message block by a high rate LDPC code. Raptor codes use a constant average output degree distribution capped at some maximum degree d_{max}, , i.e., a *light*

---

[1]For any natural number m, we denote by N_m the set $\{1, 2, \ldots, m\}$

[2]Performance of sparse graph codes is dependant on the distribution of node degrees in the decoding graph, cf. [26]. In this paper, we reserve the symbols $\Omega(x), \Phi(x)$ and $\omega(x), \phi(x)$ for the *node perspective* and their corresponding *edge perspective* output node degree distributions respectively, $\omega(x) = \Omega'(x)/\Omega'(1)$, $\phi(x) = \Phi'(x)/\Phi'(1)$

degree distribution. Raptor codes were shown to have excellent performance and linear encoding and decoding complexity, as the number of edges in the decoding graph is $\mathcal{O}(k)$. Lower encoding and decoding complexity is possible by relaxing the condition that *all* the input packets in the message block need to be decoded at the receiver end. Instead, only a certain sufficiently large *fraction* of the input symbols is required to be decoded and precoding provides the additional redundancy within the message block, which should be sufficient to complete the decoding and thus remove the error floor. Further elaboration of these ideas can be found in [28] where the intermediate performance of LT codes is studied.

The asymptotic performance of an LT code ensemble with a light degree distribution is captured by the following version of And-Or lemma [29], based on the Poisson approximation $I(x) = e^{\alpha(x-1)}$ of the input degree distribution (equal to Binomial($\frac{1}{k}, \alpha k$) by the code design).

LEMMA 1. *At overhead ε, the packet error rate of an ensemble* $LT(k, \Omega(x))$ *converges to* $y = \lim_{l \to \infty} y_l$, *as* $k \to \infty$, *where:*

$$
\begin{aligned}
y_0 &= 1, \\
y_l &= \exp\left(-(1+\varepsilon)\Omega'(1-y_{l-1})\right), \quad (2)
\end{aligned}
$$

or equivalently:

$$
\begin{aligned}
y_0 &= 1, \\
y_l &= \exp\left(-\alpha\omega(1-y_{l-1})\right), \quad (3)
\end{aligned}
$$

where α is the average input degree and $\omega(x)$ is the edge perspective output degree distribution.

Above lemma can be transformed into the linear programming routine for the calculation of asymptotically good output degree distributions. We fix the average input node degree $\alpha = \Omega'(1)(1+\varepsilon)$ and minimize the code overhead such that the desired packet error rate δ is achieved. The code overhead can be expressed in terms of the edge perspective distribution $\omega(x)$ as $1 + \varepsilon = \alpha \sum_{d=1}^{d_{max}} \frac{\omega_d}{d}$. The generic linear program is given by:

$$
\text{LP}: \qquad \min \alpha \sum_{d}^{d_{max}} \frac{\omega_d}{d} \qquad (4)
$$

$$
\alpha \sum_{d=1}^{d_{max}} \omega_d z_i^{d-1} \geq -\ln(1-z_i), \ \ i \in N_m,
$$

$$
\sum_{d=1}^{d_{max}} \omega_d = 1, \ \ \omega_d \geq 0, \ d \in N_{d_{max}},
$$

where $0 = z_1 < z_2 < \cdots < z_m = 1 - \delta$ are m equidistant points on $[0, 1 - \delta]$, δ is the desired error rate, and d_{max} is the maximum degree of $\omega(x)$. The degree distribution $\Omega(x)$ used in actual LT encoder can be determined from $\omega(x)$ as $\Omega(x) = \frac{\int_0^x \omega(z)dz}{\int_0^1 \omega(z)dz}$. Note that the solution of (4) does not depend on the average input degree α. Namely, we can allow variables ω_d to sum to an undetermined α and obtain an unnormalized edge perspective degree distribution. In further, we will omit α from the linear programs.

Classical fountain codes are inherently non-systematic as their encoding symbols are random functions of the input symbols. Nonetheless, systematic fountain codes may be designed at the expense of higher encoding/decoding complexity. Namely, it is necessary to explicitly calculate the set

of *intermediate symbols* $\bar{\mathbf{x}}$, from the message block $\mathbf{x}$, such that the input message will be replicated in the fountain bitstream. The encoder solves the equation

$$\mathbf{G}_{LT}^{\{1:k\}} \bar{\mathbf{x}}^{\mathbf{T}} = \mathbf{x}^{\mathbf{T}} \qquad (5)$$

where $\mathbf{G}_{LT}^{\{1:k\}}$ is an invertible $k \times k$ matrix formed by the first k rows of the LT generator matrix (preprocessing is required to ensure $\mathbf{G}_{LT}^{\{1:k\}}$ is invertible). In systematic Raptor codes, encoder and decoder both perform two basic operations [30]: code constraints processing, which solves a set of constraint equations (5) (at the encoder) or decodes the intermediate symbols from the observed encoding symbols (at the decoder), and LT encoding, which generates the actual output stream (at the encoder) or calculates the input vector based on the intermediate symbols (at the decoder). The computation of the intermediate symbols with Gaussian elimination is generally quadratic in k, unless a special structure of matrix $\mathbf{G}_{LT}^{\{1:k\}}$ is imposed such that the linear system (5) can be solved with the direct elimination of one unknown at a time.

3. CODING WITH PARTIAL INFORMATION

To demonstrate the utility of a rateless code solution for coding with partial information, let us consider the following example.

EXAMPLE 1. *A source node contains a large number k of information packets to be disseminated to a large number of receivers over lossy links. However, each receiver already knows a subset of the data packets, i.e., approximately $(1-p)k$ packets for $0 < p < 1$. Different receivers can possibly have knowledge of different packets. This could have arisen, e.g., as a result of transmission from other sources. Now, since the transmitter has no knowledge of which packets are available at which receivers, he is bound to encode over all its packets for multicast transmission.*

Ideally, a rateless code is the solution sought after for the setting outlined in the example, as it would be able to naturally adapt its rate to different or variable packet loss rates across the set of receivers. Alternatively, some kind of Hybrid-ARQ scheme could be employed, but due to a large number of receivers, feedback resources may be severely limited and this solution may lead to feedback implosion[3]. Clearly, each receiver must receive at least pk encoding packets to successfully recover the unknown part of the message. But how close can we get to this lower bound? As shown in [13, 14], standard LT and Raptor codes are not able to make any use of the partial information, and would require that each receiver receives slightly more than k encoding packets, to correctly recover entire message, which is very inefficient, especially for small p. However, slight modification of LT and Raptor codes is able to perform much better, as we demonstrate in the rest of this section.

3.1 Systematic Raptor-based solution

The setting of Example 1 can be viewed as an asymmetric SWC problem where correlation channel is a binary erasure channel (BEC) with erasure probability p. In addition, we

---

[3]Feedback implosion is a common problem in broadcasting arising as different receivers may request retransmission of different data packets, which in turn may lead to the multiple transmission of all data packets.
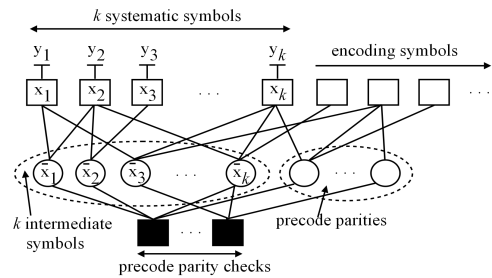

Figure 1: Systematic Raptor design for coding with partial information.

assume that the transmission occurs over another BEC. By utilizing the systematic Raptor design, this problem is reduced to a channel coding problem. Namely, we assume that the transmission of the systematic symbols has already occurred and resulted in the decoder side information. Now, systematic Raptor encoder needs to produce and transmit only the non-systematic encoding symbols. Hence, a universal systematic rateless code for transmission over an erasure channel would be sufficient to optimally solve the proposed problem. The application of systematic Raptor design in this setting is illustrated in Fig. 1. The encoder calculates intermediate symbols $\bar{\mathbf{x}}$ for each information sequence via Gaussian elimination, and then proceeds with the encoding from the $(k+1)$-th row of the generator matrix - as the strategy consists of transmitting only the non-systematic encoding symbols. The decoder directly embeds the decoder side information $\mathbf{y}$ in the decoding graph. The erased symbols in $\mathbf{y}$ are simply ignored, whereas nonerased symbols are used as the output corresponding to the systematic symbols, i.e., the first k symbols of the fountain encoding stream. Upon recovering the intermediate symbols, an additional *encoding* step is performed in order to calculate the (unknown part of the) actual message $\mathbf{x}$ by multiplying intermediate symbols $\bar{\mathbf{x}}$ with the first k rows of the LT generator matrix. The universality of Raptor codes for erasure channels implies that application of the systematic Raptor codes to the proposed problem will result in a nearly optimal design.

3.2 Non-systematic LT codes with modified degree distribution

An arguably simpler way is to avoid systematic Raptor construction altogether and to modify degree distribution of an LT code in such a way as to make use of partial information about the source. This was attempted in [13] by the introduction of a *shifted robust soliton distribution* with superior performance over the standard output degree distributions in the side information setting. The rationale behind the shifted robust soliton distribution is simple. If the original information sequence $\mathbf{x}$ contains k symbols and τ of those symbols are already known at the decoder, each output node in the decoding graph will have (τ/k)-fraction of edges removed prior to the execution of the decoding algorithm. Thus, shifted robust soliton distribution is designed such that the output degree distribution obtained after this removal of edges resembles robust soliton distributions from [2].

A different approach to the design of output degree distributions for fountain codes with decoder side information

was developed in [14, 15]. This approach uses And-Or performance analysis analogous to that of Lemma 1 and characterizes the role that the decoder side information plays in the structure of the decoding graph. For the case when correlation channel is a BEC with probability p, this role is described by a simple relationship between the degree distribution used to generate encoding symbols and the degree distribution on the resulting decoding graph, given by the following Lemma.

LEMMA 2. *Let* $\Phi(x) = \sum_{d=1}^{k} \Phi_d x^d$ *and* $\Omega(x) = \sum_{d=1}^{k} \Omega_d x^d$ *be respectively the the output degree distribution used at the encoder (incoming degree distribution) and the output degree distribution after removal of the known source nodes from the decoding graph (resulting degree distribution), then:*

$$\Omega(x) = \Phi(1 - p + px), \qquad (6)$$

PROOF. The probability that an arbitrary output node has degree i after removal of the known source nodes conditioned on its degree before removal being $j \geq i$ is given by $\binom{j}{i}(1-p)^{j-i}p^i$. Thus, the relation between the distributions Φ and Ω is given by

$$\Omega_i = \sum_{j=i}^{d_{max}} \Phi_j p^i (1-p)^{j-i}, \ \ i \in N_{d_{max}}, \qquad (7)$$

which is equivalent to (6). □

From previous lemma, it follows that $\omega(x) = \phi(1 - p + px)$, where $\phi(x)$ is the incoming edge perspective output degree distribution, whereby a linear program for optimization of the incoming degree distributions can be obtained:

$$\text{LP}: \qquad \min \frac{1}{p} \sum_{d=1}^{d_{max}} \frac{\omega_d}{d}$$

$$\sum_{d=1}^{d_{max}} \omega_d z_i^{d-1} \geq -\ln(\frac{1-z_i}{p}), \ i \in N_m \qquad (8)$$

$$\omega_d \geq 0, \ d \in N_{d_{max}}.$$

where $1-p = z_1 < z_2 < \cdots < z_m = 1-\delta p$ are equidistant points on $[1-p, 1-\delta p]$.

Similarly as in [28], we study the dual of the above (undiscretized) program given by:

$$\max_{p_Z} \mathbb{E}[-\ln(\frac{1-Z}{p})] \qquad (9)$$

$$\mathbb{E}[Z^{d-1}] \leq \frac{1}{pd}, \ d \in N_{d_{max}}$$

$$Z \in [1-p, 1-\delta p],$$

The dual program enables us to lower-bound the necessary code overhead ε with respect to the minimum value pk of received encoding packets, i.e., $n = (1+\varepsilon)pk$ encoding packets should suffice for the successful decoding. Indeed, for any feasible solution Z of the dual, the value of the objective function of the dual ψ_{dual} is less than the optimal value of the objective function of the primal $\psi_{primal}^* = 1+\varepsilon^*$. We can obtain a feasible solution of the dual by restricting the search to the discrete distributions with finite support (e.g., on a uniform grid), and thus calculate the lower bound on the code overhead penalty arising in the non-systematic fountain codes for coding with partial information. The lower bounds obtained this way, for $\delta = 0.01$ and uniform grid of
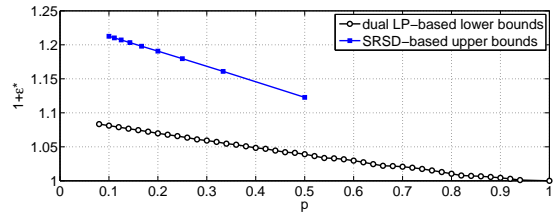

Figure 2: Lower and upper bounds for the overhead penalty in coding with partial information.
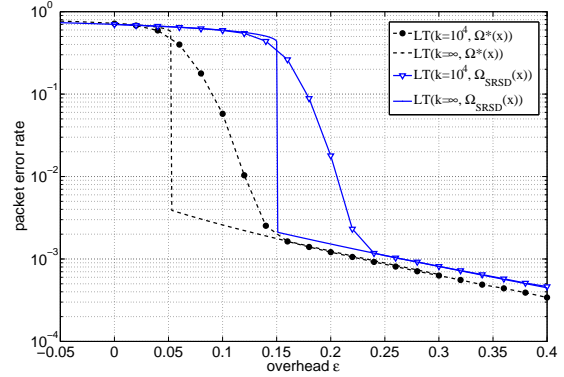

Figure 3: Non-systematic LT codes for coding with partial information.

granularity 0.001 are shown in Fig. 2 and compared with upper bounds arising from the shifted robust soliton distribution from [13] derived in [15]. Note that for case $p = 1$, or the classical channel coding problem, there is no penalty - namely, an ensemble of LT codes whose degree distributions converge pointwise to a limiting soliton distribution [28] achieves vanishing error rate at an arbitrarily small overhead as $k \to \infty$, whereas there is a clear gap for $p < 1$.

Even though the overhead penalty exists in the use of non-systematic fountain codes for coding with partial information, our results indicate that their performance does not fall too far from the optimal, and may be an interesting solution afterall. In Fig. 3, we compared the asymptotic performance and finite length ($k = 10^4$) performance of a custom degree distribution $\Omega^*(x) = 0.4816x^5 + 0.3916x^6 + 0.0792x^{29} + 0.0051^{30} + 0.0425x^{100}$ calculated with linear program (8), for $p = 0.5$, $d_{\max} = 100$ and $\delta = 0.004$, with that of truncated shifted robust soliton distribution with the same maximum degree. Packet error rates are plotted as functions of code overhead $\varepsilon = \frac{n}{pk} - 1$, where n is the number of received encoding packets. The results clearly show that linear programming optimization with (8) yields degree distributions which could be applied to practical coding with partial information. Note that the error floors in Fig. 3 can be removed by precoding, resulting in the non-systematic Raptor codes for coding with partial information.

4. SYMMETRIC SWC WITH LT CODES

In this section, we will present the case study of a simple symmetric DSC scenario using carefully designed LT codes. Consider the following example:

EXAMPLE 2. *Assume that two source nodes S_1 and S_2 are trying to disseminate a message of k packets to a large number of receivers. However, both S_1 and S_2 actually con-*

Figure 4: Decentralized distributed fountain coding.

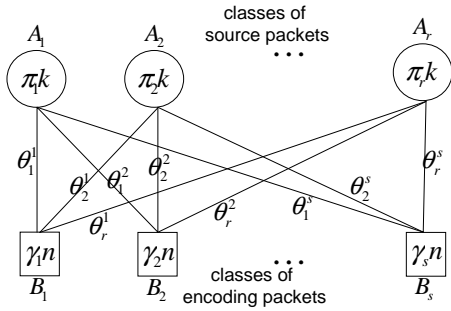**Figure 5: Symmetric Slepian Wolf Coding in Example 2 as an instance of decentralized distributed fountain coding.**

tain more than a half of all the packets (but the number of packets at S_1 and S_2 is the same). In other words, a certain portion $p < 1/2$ of the packets is common to S_1 and S_2. The source nodes do not communicate and they are oblivious of which packets are available at other source node.

Ideally, receivers would need just slightly more than k encoding packets to recover the entire message block. But how does the fact that the data available at source nodes is correlated influence the decoding performance? Note that there are three different classes of packets: (1) $\frac{1-p}{2}k$ packets available only at node S_1, (2) pk packets available at both nodes, and (3) $\frac{1-p}{2}k$ packets available only at node S_2. Let us further assume that each source node will transmit $\frac{n}{2}$ encoding packets produced with a $LT((\frac{1-p}{2}+p)k, \Omega(x))$ code.

The above setting can be viewed in light of the decentralized distributed fountain coding framework developed in [31]. This generic setting describes the asymptotic performance of dissemination of data dispersed across a set of the nodes in the network with s multiple independent LT encoders, such that the j-th encoder, $j \in N_s$, uses the degree distribution $\Omega_j(x)$. The decentralised generation of the encoding packets is described by a weighted bipartite graph $\mathcal{G} = (\mathcal{A}, \mathcal{B}, \Theta)$, illustrated in Fig. 4. In $\mathcal{G}$, nodes $\mathcal{A} = \{A_1, A_2, \ldots A_r\}$ represent a disjoint partition of N_k, such that $\forall i \in N_r$, $|A_i| = \pi_i k$, for some $\pi_i \in [0, 1]$, and nodes $\mathcal{B} = \{B_1, B_2, \ldots B_s\}$ represent a disjoint partition of N_n, such that $\forall j \in N_s$, $|Y_j| = \gamma_j n$, for some $\gamma_j \in [0, 1]$, and $\Theta = (\theta_i^j)$ is an $k \times n$ matrix, such that θ_i^j is the weight associated with the edge $A_i B_j$. The weights are normalized such that $\forall j \in N_s$, $\sum_{i \in N_r} \theta_i^j = 1$. It is useful to think of A_i, $i \in N_r$, and B_j, $j \in N_s$ as determining divison of raw data packets and encoding data packets, respectively, into classes. Thus, graph $\mathcal{G}$ characterises: (1) availability of data at source nodes: node j has access to the i-th class of packets if $\theta_i^j \neq 0$ (2) rate of production of encoding symbols at each of the source nodes: j-th encoder produces the j-th class of encoding packets, and (3) *bias* introduced towards certain portions of data in formation of encoding packets: at j-th encoder, during the generation of each encoding packet, packets from the i-th class are sampled with probability θ_i^j. If there is no bias, as would be the case in the setting of Example 2 as the source nodes are oblivious of the division of the input packets into classes, non-zero θ_i^j are proportional to the sizes of A_i. Graph $\mathcal{G}$ and the set of degree distributions $\{\Omega_j(x)\}_{j \in N_s}$ describe a code ensemble $DDLT(k, \mathcal{G}, \{\Omega_j(x)\}_{j \in N_s})$.

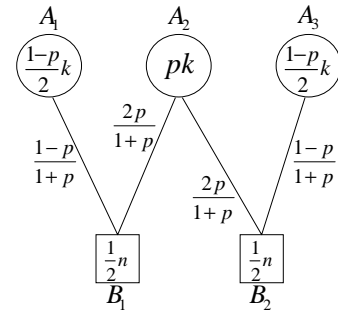A generalization of the classical And-Or tree argument yields the following result [31]:

THEOREM 1. *At overhead ε, packet error rate of an ensemble $DDLT(k, \mathcal{G}, \{\Omega_j(x)\}_{j \in N_s})$ within the i-th class of input packets converges to $y_{i,\infty} = \lim_{l \to \infty} y_{i,l}$, as $k \to \infty$, where:*

$$y_{i,0} = 1, \tag{10}$$

$$y_{i,l+1} = \exp\left[-(1+\varepsilon)\sum_{j=1}^{s} \theta_i^j \frac{\gamma_j}{\pi_i} \Omega_j'(1 - \sum_{m=1}^{r} \theta_m^j y_{m,l})\right].$$

The setting described in Example 2 can be cast as an instance of the generic decentralized distributed fountain coding, and this is illustrated in Fig. 5. Therefore, the recursive equation describing the asymptotic packet error rate in, for example, class 1 of input packets is given by:

$$y_{1,l+1} = \exp\left[-\frac{1+\varepsilon}{1+p} \cdot \right. \tag{11}$$
$$\left. \Omega'\left(\frac{(1-p)(1-y_{1,l}) + 2p(1-y_{2,l})}{1+p}\right)\right],$$

and similar equations are valid for $y_{2,l+1}$ and $y_{3,l+1}$. It is easily checked that $y_{3,l} = y_{1,l}$ and that $y_{2,l} = y_{1,l}^2$, $\forall l \geq 0$. Thus, one can trace the asymptotic behaviour of all three packet error rates with a single parameter, which allows simple transformation of the above recursive equations into a linear program optimization procedure.

Now, let us for the sake of simplicity assume $p = 1/3$, i.e., a third of all the packets are available at both sources. Simple transformations yield the following linear program:

$$\text{LP}: \quad \min \sum_{d=1}^{d_{max}} \frac{\omega_d}{d} \tag{12}$$

$$\frac{3}{4} \sum_{d=1}^{d_{max}} \omega_d z_i^{d-1} \geq -\ln\left(\frac{\sqrt{9 - 8z_i} - 1}{2}\right), \ i \in N_m$$

$$\omega_d \geq 0, \ d \in N_{d_{max}}.$$

where $0 = z_1 < z_2 < \cdots < z_m = 1 - \frac{\delta}{2} - \frac{\delta^2}{2}$ are equidistant points on $[0, 1 - \frac{\delta}{2} - \frac{\delta^2}{2}]$. The solution of this linear program is an edge perspective degree distribution which reaches the packet error rate of δ within classes 1 and 3 (and δ^2 within class 2) at the minimum overhead.

The degree distribution we obtained using a linear program in (13), by setting $\delta = 0.01$ and $d_{max} = 100$ is given by $\Omega^*(x) = 0.0020x + 0.4305x^2 + 0.2205x^3 + 0.0793x^5 + 0.1097x^6 + 0.0508x^{12} + 0.0409x^{13} + 0.0343x^{30} + 0.0106x^{32} +$
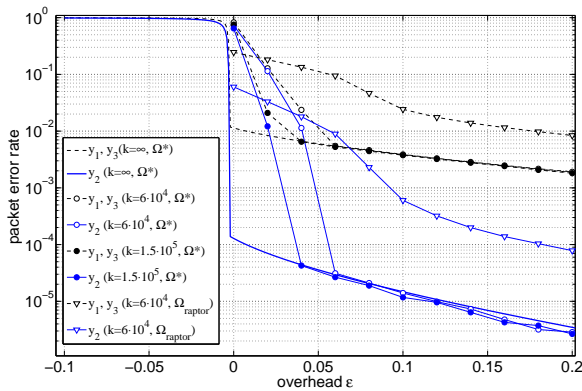
Figure 6: Symmetric Slepian Wolf Coding with LT codes.

Figure 7: The comparison of asymmetric DSC with a BIAWGN correlation channel with various kinds of Raptor codes.

$0.0215x^{100}$. Our numerical results indicate that the designed degree distribution successfully takes advantage of the fact that two sources contain correlated information. Simulation results for large blocklengths, $k = 6 \cdot 10^4$ and $k = 1.5 \cdot 10^5$ are consistent with our asymptotic analysis, as demonstrated in Fig. 6. For comparison, we included results for typical Soliton-like degree distribution $\Omega_{raptor}(x)$ used in Raptor codes [3, 12], which is clearly penalized by much higher error floors in this setting.

The symmetric Slepian Wolf coding is particularly the case for which non-systematic LT design with its simplicity and lower computational complexity has an advantage over the systematic raptor design, as it is not clear if and how the described scenario can take advantage of the systematic raptor design.

5. SOFT-DECISION DECODING FOR THE SIDE INFORMATION PROBLEM

In this section, by using the ideas of fountain coding on noisy channels, we review the rateless code design for more realistic versions of distributed source coding, i.e., when correlation channel describing the relationship between sources X and Y is noisy, as in the following example.

EXAMPLE 3. *Binary information source X over the alphabet $\{-1, 1\}$ is to be multicast to a large number of receivers over possibly noisy links. Each receiver j has an apriori access to a distorted version of the source, e.g., a correlated source Y_j such that $\mathbb{P}[X \neq Y_j] = p$, or alternatively a source impaired by a zero mean Gaussian noise with variance σ^2, i.e., $Y_j = X + \mathcal{N}(0, \sigma^2)$.*

A basic theoretical tool for analysis of BP decoder is density evolution (DE) [32], and this tool heavily relies on Gaussian approximation [33] of the messages passed during the decoding algorithm. In fountain code design for noisy channels [9], more accurate analysis is possible [34]. In [15], these techniques were used to provide an optimization framework for asymmetric SWC, where correlation channel was assumed to be a BIAWGN channel, but the techniques also apply to other correlation channels, e.g., when correlation channel is a binary symmetric channel (BSC). In [12], systematic raptor codes were adopted for asymmetric SWC
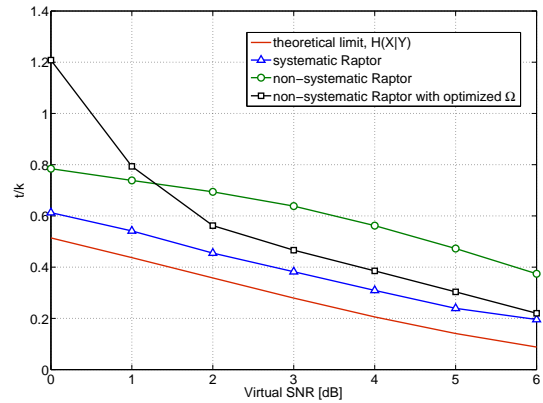
in an analogous manner to that presented in Section 3.1, and it was demonstrated how systematic raptor-based solution is a rather natural approach, which effectively reduces the problem to that of channel coding. Proposed approach was benchmarked and compared to that of Turbo codes and LDPC codes. However, it was argued that non-systematic Raptor codes are not applicable to DSC problem. Nonetheless, similarly as in the case of partial information, a code design problem for non-systematic fountain codes can be formulated as a natural modification of density evolution formulae, which would take into account the soft information already available about the source at the decoder side. Albeit being suboptimal, this strategy bears striking simplicity and lower complexity, as the size of the decoding graph is much smaller. For the actual code design techniques we refer interested reader to [15], where we derived the code optimization for the case were correlation channel is a BIAWGN channel for modified non-systematic Raptor codes. It has been shown that a non-systematic Raptor code with optimized $\Omega(x)$ does approach the performance of the systematic Raptor code scheme in the area of Signal-to-Noise Ratio (SNR) of the correlation channel for which it was optimized, but exhibits high sensitivity to overestimates of correlation SNR. This is illustrated by simulation results in Fig. 7, where we compare the performance of the non-systematic Raptor code with $\Omega(x)$ optimized for correlation channel with $SNR = 3$dB with the performance of a systematic and non-systematic Raptor code with classical fountain coding degree distribution $\Omega_{raptor}(x)$ from [3]. The vertical axis represents the average compression rate necessary for succesful decoding, i.e., t/k, where t is the average number of encoding symbols received at the decoder, and $k = 3140$ was the block length.

6. CONCLUSIONS

We have presented an overview of uses of fountain codes for distributed source coding. As fountain codes are essentially erasure correcting codes, they find the best use in distributed source coding problems where correlation channel is an erasure channel, and here we presented the code design for both symmetric and asymmetric Slepian-Wolf coding for such cases. Systematic raptor code solution achieves optimal

performance in asymmetric case, but at the cost of increased computational complexity, whereas in symmetric case, carefully designed non-systematic fountain codes exhibit very good performance. We have also briefly overviewed the use of fountain codes in asymmetric Slepian-Wolf coding with a noisy correlation channel. Our results indicate that fountain codes are a natural practical coding scheme for DSC scenarios and that their design may be tuned to perform close to the optimal achievable rates with low computational cost.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] J. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," *Proc. ACM SIGCOMM*, Vancouver, Canada, Sept. 1998.

[2] M. Luby, "LT codes," *Proc. IEEE Symp. Foundations of Computer Science (FOCS)*, Vancouver, Canada, Nov. 2002.

[3] A. Shokrollahi, "Raptor codes," *IEEE Trans. Info. Theory*, 52(6), 2551-2567, June 2006.

[4] 3GPP, "3GPP TS 26.346 V7.0.0, Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Protocols and Codecs", Sept. 2007.

[5] "Digital Video Broadcasting (DVB): Transmission System for Handheld Terminals (DVB-H)", ETSI EN 302 304 V1.1.1, 2004.

[6] R. Palanki, J. Yedidia, "Rateless codes on noisy channels," in *Proc. IEEE Int. Symp. on Info. Theory*, Chicago, USA, July 2004.

[7] H. Jenkač, T. Mayer, T. Stockhammer, W. Xu., "Soft decoding of LT codes for wireless broadcast," *Proc. IST Summit 2005*, Dresden, Germany, June 2005.

[8] J. Castura, Y. Mao. "Rateless coding over fading channels," *IEEE Commun. Letters*, 10(1), 46-48, Jan. 2006.

[9] O. Etesami, A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Info. Theory*, 52(5), 2033-2051, May 2006.

[10] G. Caire, S. Shamai, A. Shokrollahi, S. Verdu, "Fountain codes for lossless data compression", *Algebraic Coding Theory and Information Theory*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 68, 1-20, AMS, 2006.

[11] B.N. Ndzana, A. Shokrollahi, J. Abel, "Fountain codes for the Slepian-Wolf problem", *Proc. Allerton Conf.*, Allerton, USA, 2006.

[12] M. Fresia, L. Vandendorpe, "Distributed source coding using Raptor codes", *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Washington D.C., USA, Nov. 2007.

[13] S. Agarwal, A. Hagedorn, A. Trachtenberg, "Adaptive rateless coding under partial information", *Proc. Info. Theory and Applications (ITA) Workshop*, San Diego, USA, Feb. 2008.

[14] D. Sejdinović, R. Piechocki, A. Doufexi, M. Ismail, "Fountain coding with decoder side information," in *Proc. IEEE Int. Conf. on Communications*, Beijing, China, May 19-23, 2008.

[15] D. Sejdinović, R. Piechocki, A. Doufexi, M. Ismail, "Fountain code design for data multicast with side information," *IEEE Trans. Wireless Commun.*, to appear.

[16] N. Dütsch, H. Jenkač, T. Mayer, J. Hagenauer, "Joint source-channel-fountain coding for asynchronous broadcast," *Proc. IST Summit 2005*, Dresden, Germany, June 2005.

[17] Q. Xu, V. Stanković, Z. Xiong, "Distributed joint source-channel coding of video using Raptor codes," *IEEE J. Sel. Areas Commun.*, 25, 851-861, May 2007.

[18] O.Y. Bursalioglu, M. Fresia, G. Caire, H.V. Poor, "Lossy joint source-channel coding using Raptor codes," *Int. Journal of Digital Multimedia Broadcasting*, 2008.

[19] J.P. Wagner, J. Chakareski, P. Frossard, "Streaming of scalable video from multiple sources using rateless codes," *Proc. Int. Conf. Multimedia & Expo (ICME)*, Toronto, Canada, July 2006.

[20] T. Stockhammer, T. Gasiba, W.A. Samad, T. Schierl, H. Jenkač, T. Wiegand, W. Xu, "Nested harmonic broadcasting for scalable video over mobile datacast channels," *Wireless Communications. and Mobile Computing (Wiley)*, 7(2), 235-256, Feb. 2007.

[21] Q. Xu, V. Stanković, Z. Xiong, "Wyner-Ziv video compression and fountain codes for receiver-driven layered multicast," *IEEE Trans. Circ. and Syst. for Video Tech.*, 17, 901-906, July 2007.

[22] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, Z. Xiong, "Expanding window fountain codes for scalable video multicast", *IEEE Trans. Multimedia*, 11(6), 2009.

[23] S. Puducheri, J. Kliewer, T. Fuja, "The design and performance of distributed LT codes", *IEEE Trans. Info. Theory*, 53(10), 3740-3754, Oct. 2007.

[24] D. Slepian, J.K. Wolf. "Noiseless coding of correlated information sources", *IEEE Trans. Info. Theory*, IT-19, 471-480, July 1973.

[25] S. Agarwal, A. Hagedorn, A. Trachtenberg, "Near optimal update-broadcast of datasets", *Proc. Int. Workshop on Data Intensive Sensor Networks (DISN)*, Mannheim, Germany, May 2007.

[26] T. Richardson, R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[27] F. Kschischang, B.J. Frey and H.-A. Loeliger, "Factor graphs and the sum-product algorithm", *IEEE Trans. Info. Theory*, 47(2), 498-519, 2001.

[28] S. Sanghavi, "Intermediate performance of rateless codes", *IEEE Info. Theory Workshop 2007*, Tahoe, USA, Sept. 2007.

[29] M. Luby, M. Mitzenmacher and A. Shokrollahi, "Analysis of random processes via And-Or tree evaluation," *Proc. SIAM Symp. on Discrete Algorithms (SODA)*, San Francisco, USA, Jan. 1998.

[30] M. Luby, M. Watson, T. Gasiba, T. Stockhammer and W. Xu, "Reliable multimedia download delivery in cellular broadcast networks", *IEEE Trans. Broadcasting*, 53(1/2), 235-246, Mar. 2007.

[31] D. Sejdinović, R. Piechocki, A. Doufexi, M. Ismail, "Decentralised distributed fountain coding: asymptotic analysis and design", submitted to *IEEE Commun. Letters*.

[32] T. Richardson, R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding", *IEEE Trans. Info. Theory*, 47(2), 599-618, 2001.

[33] S.-Y. Chung, T. Richardson, R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation", *IEEE Trans. Info. Theory*, 47(2), 657-670, Feb. 2001.

[34] M. Ardakani, F.R. Kschischang, "A more accurate one-dimensional analysis and design of irregular LDPC codes", *IEEE Trans. Commun.*, 52(12), 2106-2114, Dec. 2004.