A Latency Outer Bound for Binary Source Broadcasting under Hamming Distortion

Louis Tan and Ashish Khisti
Dept. of Electrical & Computer Engineering
University of Toronto
Toronto, ON M5S 3G4 Canada
{ltan, akhisti}@comm.utoronto.ca

Emina Soljanin
Bell Labs, Alcatel-Lucent
Murray Hill NJ 07974, USA
emina@alcatel-lucent.com

***Abstract*—We study the problem of sending a binary source over a two-receiver erasure broadcast channel with source-channel bandwidth mismatch, under a Hamming distortion measure. Given the distortion constraints of both users, our main focus is to derive a lower bound for the minimum latency required before both users will be able to reconstruct the source subject to their distortion constraints. Our derivation involves adapting an outer bound given by Reznic et al. (2006) for the related quadratic Gaussian source broadcast problem. This outer bound involves the introduction of an auxiliary random variable and the entropy power inequality, which is replaced with analogous inequalities in the problem we consider. We numerically evaluate our lower bound and find that it can be tighter than the one given by the source-channel separation theorem.**

I. Introduction

We consider the minimum latency that must be incurred before two users are able to reconstruct a binary source that is sent to them over an erasure broadcast channel. In this problem, the source is equiprobable and the two users wish to individually reconstruct it only to within some degree of prespecified Hamming distortion. The goal then, is to determine a lower bound for the latency required before both users will be able to do so. The latency is measured in terms of the number of channel symbols per source symbol that must be sent over the broadcast channel before both users can reconstruct the source. It also represents the source-channel bandwidth mismatch.

Our problem is motivated by the single-server streaming model that has been introduced in [1]. Here, a server wishes to communicate a source sequence to a group of heterogeneous users over a broadcast channel. The underlying channels are modelled as erasure channels and each user's channel has a certain loss probability. Furthermore each user is only interested in retrieving a certain fraction of source packets. The transmitter continuously broadcasts coded packets and each receiver waits until it is able to retrieve sufficiently many packets of the underlying source.

The focus in [1] is to optimize the degree distribution of a rateless codes for this application. In this correspondence, we take a complementary view that formulates this problem in terms of the joint source-channel coding of a binary source that is to be sent over a binary erasure broadcast channel subject to distortion constraints. In particular, we will be interested in studying the additional latency experienced by a user due to the presence of an additional user in the network, who the transmitter must simultaneously serve.

In the absence of this additional user, the source-channel separation theorem gives an achievable lower bound for the latency incurred by a user, given his distortion constraint and channel noise. When applied to a broadcast network, despite the fact that this bound does not consider the tensions involved in the addition of another user, it was nevertheless shown in [2] that this lower bound could still be simultaneously achieved by both users over a range of distortion values in the related quadratic Gaussian source broadcast problem. In effect, both users operate as if the other one is not present and we say that both users simultaneously achieve their *individual* point-to-point optimal latencies.

On the other hand, if we only consider the *overall* or *maximum* network latency, it was shown in [3] that in general, there is a conflict between the needs of both users. Indeed, it may be necessary for the overall latency to be strictly larger than the point-to-point optimal latencies of either of the two users given their collective distortion constraints. We now show a similar result for our current problem involving binary source broadcasting.

II. Problem Formulation

The problem is illustrated in Fig. 1. We consider a binary memoryless stationary source $\{S(i)\}_{i=1,2,\dots}$ producing equiprobable symbols in the alphabet $\mathcal{S} = \{0, 1\}$, which we wish to communicate to two users over an erasure broadcast channel. Let $\mathcal{S}^k$ be the set of all k-vectors with components in $\mathcal{S}$, and denote the vector $(S(1), S(2), \dots, S(k))$ as S^k.

The source is communicated by a block encoding function that maps the length k source sequence S^k, to a length n channel input sequence $X^n = (X(1), X(2), \dots, X(n))$ where $X(l)$ denotes the l^{th} channel input taken from the alphabet $\mathcal{X} = \{-1, +1\}$.

Let $Y_i(l)$ be the channel output observed by user i on the l^{th} channel use for $i \in \{1, 2\}$ and $l = 1, 2, \dots, n$. The channel model is given by

$$Y_i(l) = X(l) \cdot N_i(l), \tag{1}$$

where $N_i(l)$ is a $\mathrm{Bern}(1-\epsilon_i)$ random variable representing the noise at user i's l^{th} channel output. The channel is memoryless in the sense that $N_i(l)$ is drawn i.i.d. from a $\mathrm{Bern}(1-\epsilon_i)$ distribution. Specifically, the statistics of $N_i(l)$ are such that

$$N_i(l) = \begin{cases} 0 & \text{with probability } \epsilon_i, \\ 1 & \text{with probability } 1 - \epsilon_i. \end{cases} \tag{2}$$

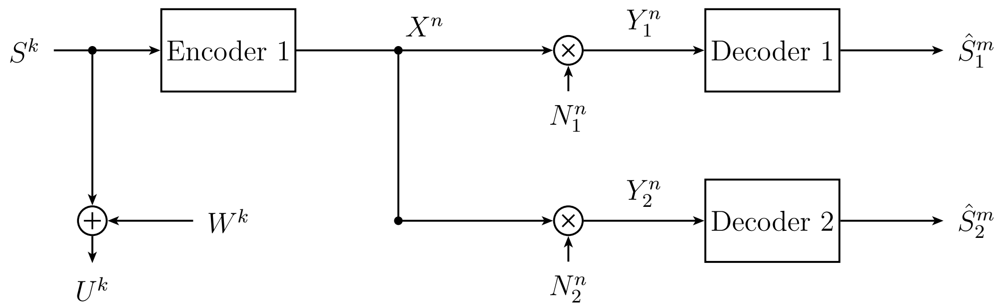

Fig. 1. Broadcasting an equiprobable binary source over an erasure broadcast channel with bandwidth mismatch. The output U^k, of the test channel of Section IV, is also shown.

Thus, $Y_i(l)$ takes on values in the alphabet $\mathcal{Y} = \{-1, 0, +1\}$. If we define '0' as the erasure symbol, we can interpret $Y_i(l)$ as the output of an erasure channel, with erasure probability ϵ_i, when $X(l)$ is the channel input. We will assume that $\epsilon_1 < \epsilon_2$ so that user 1 is in some sense, the "stronger" user.

Having observed the channel output, user i then uses it to reconstruct the source as $\hat{S}_i^k$, a length k sequence with components in $\hat{\mathcal{S}} = \{0, 1\}$. On a symbol-by-symbol basis, we measure the reconstruction's fidelity with the Hamming distortion $d : \mathcal{S} \times \hat{\mathcal{S}} \to \{0, 1\}$ given by

$$d(s, \hat{s}) = s \oplus \hat{s}. \tag{3}$$

The per-letter distortion of a vector is then defined as

$$d(s^k, \hat{s}^k) = \frac{1}{k} \sum_{i=1}^{k} d(s_i, \hat{s}_i). \tag{4}$$

We now define the components of our problem.

Definition 1. *A (k, n, D_1, D_2) source-channel code for source S on the erasure broadcast channel consists of*

1) *An encoding function $f_k : \mathcal{S}^k \to \mathcal{X}^n$ such that $X^n = f_k(S^k)$*
2) *Two decoding functions $g_i : \mathcal{Y}^n \to \hat{\mathcal{S}}^k$ such that $\mathbb{E}d(S^k, g_i(X^n \cdot N_i^n)) \le D_i$ holds for $i \in \{1, 2\}$.*

where $\mathbb{E}(\cdot)$ is the expectation operation.

A point is now made about the modelling of latencies in our problem. We define the *latency* or *bandwidth expansion factor* $b \in [0, \infty)$, as the number of channel uses per source symbol that are delivered over the broadcast channel, i.e., $b \triangleq n/k$. This is to say that $b \cdot k$ channel uses are required before both users can reconstruct S^k subject to their distortion constraints. Our problem is now defined as characterizing the achievable latency region under a given pair of distortion constraints as per the next definition.

Definition 2. *A latency b, is said to be (D_1, D_2)-achievable over the erasure broadcast channel if for every $\delta > 0$, there exists for sufficiently large k, a $(k, b \cdot k, d_1, d_2)$ source-channel code such that*

$$D_i + \delta \ge d_i, \quad i \in \{1, 2\}. \tag{5}$$

The achievable latency region is the set of all achievable latencies under the prescribed distortion pair.

III. Main Result

In this section, we present our main result, which is an outer bound for the achievable latency region. Before presenting the outer bound however, we will first derive a simple one based on the source-channel separation theorem. Consider an equiprobable binary source being sent over an erasure channel, with erasure probability ϵ. We wish to reconstruct the source subject to a Hamming distortion of D. From the source-channel separation theorem it is not hard to see that the value b^* is an outer bound for the latency where

$$b^* = \frac{1 - H(D)}{1 - \epsilon}. \tag{6}$$

In a broadcast setting, where there are two users that have to simultaneously satisfy the distortion constraints D_1 and D_2, over an erasure broadcast channel with erasure probabilities ϵ_1 and ϵ_2, it should therefore be clear that a lower bound for the latency in our problem is the value B^* where

$$B^* = \max\left(\frac{1 - H(D_1)}{1 - \epsilon_1}, \frac{1 - H(D_2)}{1 - \epsilon_2}\right). \tag{7}$$

We now present our new outer bound, which can be tigther than the one given above. The derivation will be presented in the following section.

Theorem 1. *Define the function $f_0(\kappa, \theta)$ so that*

$$\begin{aligned} f_0(\kappa, \theta) \triangleq \frac{1 - H(\kappa * D_2)}{1 - \epsilon_2} \\ + \frac{\theta - 1 + H\left(H^{-1}(1 - \theta) * \kappa\right)}{1 - \epsilon_1}, \end{aligned} \tag{8}$$

Let (κ^*, θ^*) be the solution to the optimization problem given by

$$\begin{aligned} \max_{\kappa} \min_{\theta} \quad & f_0(\kappa, \theta) \\ \text{subject to} \quad & 0 \le \kappa \le 1/2 \\ & 1 - H(D_1) \le \theta \le 1. \end{aligned} \tag{9}$$

If the latency b, is (D_1, D_2)-achievable over the erasure broadcast channel, then

$$b \ge f_0(\kappa^*, \theta^*). \tag{10}$$

We numerically evaluate the optimization problem in Theorem 1 and illustrate the outer bound in Figure 2. In this plot, the channel erasure probabilities are set at $\epsilon_1 = 0.1$, $\epsilon_2 = 0.5$. We fix a value of $D_1 = 1/32$ and vary D_2 in

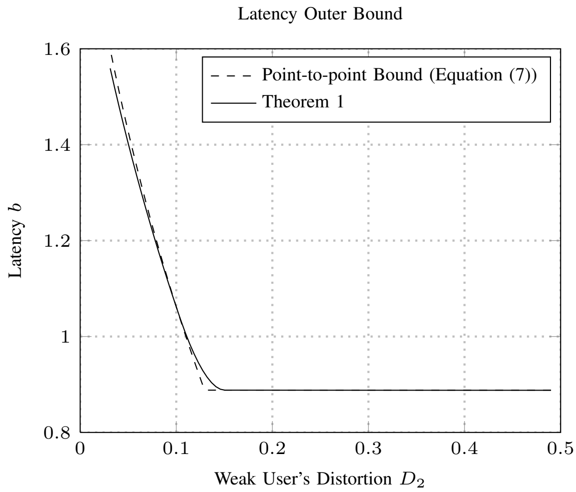

Fig. 2. A numerical plot of the latency outer bound of Theorem 1. We fix $\epsilon_1 = 0.1$, $\epsilon_2 = 0.5$ and $D_1 = 1/32$ and vary D_2 in the range $[1/32, 1/2]$ on the x-axis while plotting the outer bound for b on the y-axis. The point-to-point outer bound from the source-channel separation theorem is also shown.

the range $[1/32, 1/2]$ on the x-axis while plotting the outer bound for b on the y-axis. As can be seen, the outer bound of Theorem 1 can be tighter than the trivial point-to-point outer bound of Equation (7).

IV. Derivation

We will now derive the outer bound for the achievable latency region of Theorem 1 by adapting an outer bound for the quadratic Gaussian source broadcast problem given by Reznic et al [3]. The outer bound involves the introduction of an auxiliary random variable and the entropy power inequality, which is replaced with analogous inequalities in the problem we consider.

Let $\underline{N}$ be the collection of both users' noise processes. That is, we define $\underline{N} = (N_1(1), N_1(2), \ldots, N_1(n), N_2(1), N_2(2), \ldots, N_2(n))$, where we again remind the reader that $N_i(l) \overset{\text{i.i.d.}}{\sim} \text{Bern}(1-\epsilon_i)$. Let us first consider an achievable distortion D_1, that user 1 achieves. From rate-distortion theory, we can begin by writing

$$
\begin{aligned}
kR(D_1) &\leq I(S^k; \hat{S}_1^k) && (11)\\
&\overset{(a)}{\leq} I(X^n; Y_1^n) && (12)\\
&\leq I(X^n; Y_1^n, \underline{N}) && (13)\\
&\overset{(b)}{=} I(X^n; Y_1^n | \underline{N}) && (14)
\end{aligned}
$$

where

(a) follows from the Markov chain $S^k - X^n - Y_1^n - \hat{S}_1^k$, and the data processing inequality

(b) follows from the fact that the noise processes are independent of the channel input.

To capture the tension between the simultaneously achievable distortions, our goal is to now upper bound the right hand side of Equation (14) with a function of user 2's distortion in order to give a bound on user 1's distortion, which appears on the left hand side of Equation (14). To do this, we follow the procedure of Reznic, Feder and Zamir [3], and first introduce an auxiliary random variable $U^k = (U(1), U(2), \ldots, U(k))$, whose l^{th} element is obtained by passing $S(l)$ through a memoryless binary symmetric channel with crossover probability κ (see Fig. 1).

In other words, we first define the random vector $W^k = (W(1), W(2), \ldots, W(k))$, whose entries are sampled i.i.d. from a Bernoulli distribution such that

$$
W(l) = \begin{cases} 0 & \text{with probability } 1-\kappa, \\ 1 & \text{with probability } \kappa \end{cases} \tag{15}
$$

for $l = 1, 2, \ldots, k$ and some $\kappa \in [0, \frac{1}{2}]$. We then define the vector U^k such that it's l^{th} element is given as

$$
U(l) = S(l) \oplus W(l), \tag{16}
$$

where $\oplus$ is the XOR operation.

We now incorporate the auxiliary random variable by continuing from Equation (14) and expanding the right hand side of it as

$$
\begin{aligned}
I(X^n; Y_1^n | \underline{N}) &= I(U^k, X^n; Y_1^n | \underline{N}) - I(U^k; Y_1^n | X^n, \underline{N}) && (17)\\
&\overset{(a)}{=} I(U^k, X^n; Y_1^n | \underline{N}) && (18)\\
&= I(U^k; Y_1^n | \underline{N}) + I(X^n; Y_1^n | U^k, \underline{N}) && (19)\\
&\overset{(b)}{=} H(U^k) - H(U^k | Y_1^n, \underline{N}) && (20)\\
&\qquad + I(X^n; Y_1^n | U^k, \underline{N})\\
&\overset{(c)}{=} k - H(U^k | Y_1^n, \underline{N}) && (21)\\
&\qquad + I(X^n; Y_1^n | U^k, \underline{N})\\
&\overset{(d)}{=} k - H(U^k | Y_1^n, \underline{N}) + H(Y_1^n | U^k, \underline{N}) && (22)
\end{aligned}
$$

where

(a) follows from the fact that Y_1^n is a function of X^n and $\underline{N}$

(b) follows from the fact that U^k is independent of $\underline{N}$.

(c) follows from the fact that U^k is memoryless and $S(l)$ in Equation (16) is equiprobable

(d) follows from the fact that Y_1^n is a function of X^n and $\underline{N}$.

The next two subsections are dedicated to individually bounding the last two terms on the right hand side of Equation (22). We then combine all the bounds to formulate the outer bound as an optimization problem in Section IV-C.

A. $H(Y_1^n | U^k, \underline{N})$ Upper Bound

We begin this subsection by first restating a Lemma that relates the entropies of two random vectors that are each a separately erased version of a common random vector. The lemma first appeared in a related form in a work by Zhu and Guo while studying the layered erasure one-sided interference channel [4]. We refer the interested reader to the appendix of [5] for a detailed proof of the following lemma.

Lemma 1. *If $\hat{T}$ is a collection of arbitrary random variables independent of $\underline{N}$, then*

$$
\frac{1}{1-\epsilon_1} H(Y_1^n | \underline{\hat{T}}, \underline{N}) \leq \frac{1}{1-\epsilon_2} H(Y_2^n | \underline{\hat{T}}, \underline{N}). \tag{23}
$$

We use Lemma 1 to write

$$\begin{aligned}
H(Y_1^n|U^k,\underline{N}) &\le \frac{1-\epsilon_1}{1-\epsilon_2}H(Y_2^n|U^k,\underline{N}) \\
&= \frac{1-\epsilon_1}{1-\epsilon_2}\left(H(Y_2^n|\underline{N}) - I(U^k;Y_2^n|\underline{N})\right) \\
&\overset{(a)}{\le} \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - I(U^k;Y_2^n|\underline{N})\right) \\
&\overset{(b)}{=} \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - I(U^k;Y_2^n,\underline{N})\right) \\
&\le \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - I(U^k;Y_2^n)\right) \\
&\overset{(c)}{\le} \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - I(U^k;\hat{S}_2^k)\right) \quad (24)
\end{aligned}$$

where

(a) follows from the fact that the entropy of a memoryless erasure channel output is maximized when the input is i.i.d. and equiprobable
(b) follows from the fact that U^k is independent of $\underline{N}$
(c) follows from the Markov chain $U^k - Y_2^n - \hat{S}_2^k$ and the data processing inequality.

We are now in a position to incorporate D_2, user 2's distortion, into the outer bound. We do this by first computing the expected distortion between the auxiliary random vector U^k and user 2's reconstruction $\hat{S}_2^k$, under a Hamming distortion measure. Before doing this, for $i = 1, 2, \ldots, k$, it will be convenient to define the indicator variable E_i such that

$$E_i = \begin{cases} 0 & \text{if } \hat{S}_2(i) = S(i), \\ 1 & \text{if } \hat{S}_2(i) \ne S(i). \end{cases} \quad (25)$$

It is worth noting that E_i is a random variable with an expected value $\mathbb{E}E_i$ satisfying

$$\frac{1}{k}\sum_{i=1}^{k} \mathbb{E}E_i = D_2. \quad (26)$$

We now calculate

$$\begin{aligned}
\mathbb{E}d_H(U^k,\hat{S}_2^k) &= \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^{k} d_H(U(i),\hat{S}_2(i))\right) \\
&= \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^{k} U(i)\oplus\hat{S}_2(i)\right) \\
&\overset{(a)}{=} \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^{k} \{S(i)\oplus W(i)\}\oplus\{S(i)\oplus E_i\}\right) \\
&= \frac{1}{k}\sum_{i=1}^{k} \mathbb{E}\left(W(i)\oplus E_i\right) \\
&\overset{(b)}{=} \frac{1}{k}\sum_{i=1}^{k} \kappa * \mathbb{E}E_i \\
&\overset{(c)}{=} \kappa * D_2 \quad (27)
\end{aligned}$$

where

(a) follows from Equations (16) and (25)
(b) follows from the fact that E_i is independent of $W(i)$ and the fact that the indicator variable takes on the value '1' with a probability equal to its expected value
(c) follows from Equation (26).

Continuing from Equation (24), we have that

$$\begin{aligned}
H(Y_1^n|U^k,\underline{N}) &\le \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - I(U^k;\hat{S}_2^k)\right) \\
&\overset{(a)}{\le} \frac{1-\epsilon_1}{1-\epsilon_2}\big(n(1-\epsilon_2) \\
&\qquad -kR_H(\mathbb{E}d_H(U^k,\hat{S}_2^k))\big) \\
&\overset{(b)}{=} \frac{1-\epsilon_1}{1-\epsilon_2}\left(n(1-\epsilon_2) - kR_H(\kappa * D_2)\right) \\
&\overset{(c)}{=} \frac{1-\epsilon_1}{1-\epsilon_2}\big(n(1-\epsilon_2) \\
&\qquad - k(1 - H(\kappa * D_2))\big) \quad (28)
\end{aligned}$$

where

(a) follows from the rate-distortion theorem
(b) follows from Equation (27)
(c) follows from the rate distortion function for an equiprobable source with a Hamming distortion measure.

B. $H(U^k|Y_1^n,\underline{N})$ *Lower Bound*

In the last subsection, we used Lemma 1 in a way that was analogous to the entropy power inequality, to relate the entropies of two random vectors that were each an erased version of a common random vector. In this subsection we will use Mrs. Gerber's Lemma to bound the entropies of two binary random vectors where one is the result of passing the other through a memoryless binary symmetric channel. We first restate the Lemma [6].

Lemma 2 (Mrs. Gerber's Lemma)**.** *Let* $H^{-1} : [0,1] \to [0,1/2]$ *be the inverse of the binary entropy function, i.e.,* $H(H^{-1}(v)) = v$. *Let* S^k *be a binary-valued random vector and* T *be an arbitrary random variable. If* W^k *is a vector of independent and identically distributed* $\mathrm{Bern}(\kappa)$ *random variables independent of* (S^k, T) *and* $U^k = S^k \oplus W^k$, *then*

$$\frac{H(U^k|T)}{k} \ge H\left(H^{-1}\left(\frac{H(S^k|T)}{k}\right) * \kappa\right) \quad (29)$$

where $*$ *is the convolution operator such that* $a * b = a(1-b) + b(1-a)$.

If we apply Mrs. Gerber's Lemma to our current situation, we have that

$$\begin{aligned}
H(U^k|Y_1^n,\underline{N}) &\ge kH\left(H^{-1}\left(\frac{H(S^k|Y_1^n,\underline{N})}{k}\right) * \kappa\right) \\
&= kH\left(H^{-1}\left(\frac{H(S^k|\underline{N}) - I(S^k;Y_1^n|\underline{N})}{k}\right) * \kappa\right) \\
&\overset{(a)}{=} kH\left(H^{-1}\left(\frac{k - I(S^k;Y_1^n|\underline{N})}{k}\right) * \kappa\right) \\
&\overset{(b)}{\ge} kH\left(H^{-1}\left(\frac{k - I(S^k,X^n;Y_1^n|\underline{N})}{k}\right) * \kappa\right) \\
&\overset{(c)}{=} kH\left(H^{-1}\left(\frac{k - I(X^n;Y_1^n|\underline{N})}{k}\right) * \kappa\right) \quad (30)
\end{aligned}$$

where

(a) follows from the fact that the source is memoryless, equiprobable and independent of the noise processes

(b) follows from the fact that the function $H^{-1}(\cdot)$ is monotonically increasing and the function $H(\cdot)$ is monotonically increasing when $\kappa \in [0, 1/2]$, i.e., $H(\cdot)$ is monotonically increasing when its domain is restricted to be in $[0, 1/2]$

(c) follows from the fact that Y_1^n is a function of X^n and $\underline{N}$

C. An Outer Bound as an Optimization Problem

We now combine the upper and lower bounds of the previous two subsections to give an outer bound in the form of an optimization problem. Combining Equations (14), (22), (28) and (30), we can derive the inequality given in Equation (31).

Define θ such that

$$\theta \triangleq \frac{I(X^n; Y_1^n | \underline{N})}{k}. \qquad (33)$$

We can rearrange right hand inequality in Equation (31) and isolate b to get that $b \geq f_0(\kappa, \theta)$, where $f_0(\kappa, \theta)$ is given by Equation (10) and we express f_0 as a function of only κ and θ, under the assumption that ϵ_1, ϵ_2, D_1 and D_2 are fixed.

We now look to maximize $f_0(\kappa, \theta)$ in order to give the largest lower bound. We note that any code that achieves the distortion pair (D_1, D_2), induces a value of θ, which must satisfy the necessary conditions given in Equation (31). In particular, in seeking the largest lower bound for b, the least favorable θ must satisfy this for any choice of κ, which determines the test channel. Thus, we may formulate the outer bound as the optimization problem given in Equation (9).

References

[1] Y. Li, E. Soljanin, and P. Spasojević, "Three schemes for wireless coded broadcast to heterogeneous users," *Elsevier Physical Communications: Special Issue on Wireless Network Coding*, To appear 2012 `arXiv:submit/0477372`.

[2] L. Tan, A. Khisti, and E. Soljanin, "Quadratic Gaussian source broadcast with individual bandwidth mismatches," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, Jul. 2012, pp. 204–208.

[3] Z. Reznic, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3778–3788, Aug. 2006.

[4] Y. Zhu and D. Guo, "Capacity region of layered erasure one-sided interference channels without CSIT," in *Proc. IEEE Information Theory Workshop (ITW)*, Jan. 2010, pp. 1–5.

[5] L. Tan, "On delay-sensitive lossy source broadcast," MASc. thesis, Dept. Elect. Eng., Univ. Toronto, Toronto, ON, November 2012.

[6] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

$$1 - H(D_1) \leq \frac{I(X^n; Y_1^n|\underline{N})}{k} \leq 1 + \frac{1-\epsilon_1}{1-\epsilon_2}\left(b(1-\epsilon_2) - (1 - H(\kappa * D_2))\right) - H\left(H^{-1}\left(\frac{k - I(X^n; Y_1^n|\underline{N})}{k}\right) * \kappa\right) \quad (31)$$