Slepian-Wolf Coding for Broadcasting with Cooperative Base-Stations

Roy Timo† and Michèle Wigger‡
†Institute for Communications Engineering, Technische Universität München
‡Communications and Electronics Department, Telecom ParisTech‡
roy.timo@tum.de, michele.wigger@telecom-paristech.fr

***Abstract*—We propose a base-station (BS) cooperation model for broadcasting a discrete memoryless source in a cellular or heterogeneous network. The model allows the receivers to use helper BSs to improve network performance, and it permits the receivers to have prior side information about the source. We establish the model's information-theoretic limits in two operational modes: In Mode 1, the helper BSs are given information about the channel codeword transmitted by the main BS, and in Mode 2 they are provided information about the source. Optimal codes for Mode 1 use *hash-and-forward coding* at the helper BSs; while, in Mode 2, optimal codes use source codes from Wyner's *helper side-information problem* at the helper BSs. We prove the optimality of both approaches by way of a new list-decoding generalisation of [8, Thm. 6], and, in doing so, show an operational duality between Modes 1 and 2.**

I. Introduction

The proliferation of wireless communications devices presents significant performance challenges for cellular networks, and it will require more sophisticated heterogeneous networks in the near future [1, 2]. A powerful methodology for improving performance is centered on the idea of base-station (BS) cooperation: Instead of operating independently, future BSs will coordinate encoding and decoding operations using information shared over backbone networks. The tremendous potential of BS cooperation has been widely investigated [3]–[5]; however, despite many advances, there remains significant challenges in understanding and exhausting its benefits. Indeed, the fundamental limits of cooperation are fully understood in very few settings [4].

To help understand the full potential of BS cooperation, we consider a simple, but rather useful, broadcast model. The setup for two receivers is shown in Figure 1. A source $\mathbb{X}$ is to be reliably transmitted over a broadcast channel to many receivers, and the idea is to improve network performance by allowing the receivers to be assisted by *helper* BSs. In a future heterogenous network, for example, the helpers may be pico or femto BSs operating within the main macro cell on orthogonal channels [6]. Alternatively, the helpers may be WiFi hotspots through which traffic is diverted from a heavily loaded cellular network [7]. The purpose of this paper is to characterise the model's information-theoretic limits, and to provide architectural insights for optimal codes.

A more complete version of this work has been submitted for possible publication in the IEEE Transactions on Communications [26].

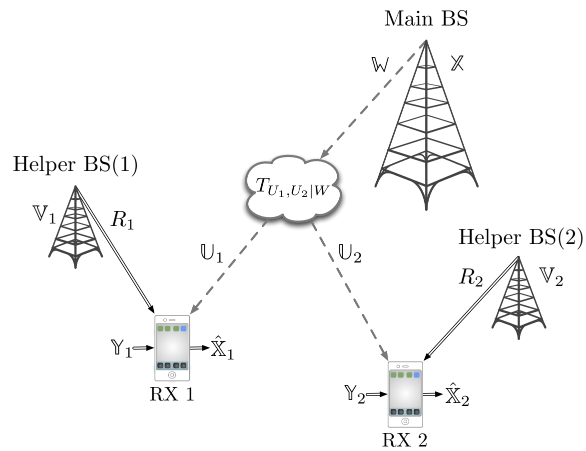

Fig. 1. Broadcasting with helper BSs and receiver side information.

We assume that the broadcast channel from the main BS is discrete and memoryless, and the channels from the helper BSs are noiseless and rate-limited. Although this setup does not capture all modes of cooperation, it nevertheless has enough sophistication to provide insight into some important coding challenges. For example, consider the idea of augmenting traffic flow in a cellular network via a WiFi hotspot: The hotspot's radio-access technology is orthogonal to that of the cellular network, and a cellular network engineer can well approximate the WiFi link by a noiseless rate-limited channel. A natural question is then: What coding techniques at the BSs and WiFi hotspot yield the best overall performance?

We consider two operational modes in the above framework.

- *Mode 1:* The helper BSs are given side information about the channel codeword transmitted by the main BS.
- *Mode 2:* The helper BSs are given correlated side information about the source $\mathbb{X}$.

We will see that optimal codes for Mode 1 combine virtual-binning from *Slepian-Wolf Coding over Broadcast Channels* [8] with hash-and-forward coding for the *primitive relay channel* [9]. Optimal codes for Mode 2, on the other hand, combine virtual binning with source codes from *Wyner's helper side-information problem* [10]. We prove the optimality of both codes by way of a new list-decoding generalisation of [8, Thm. 6], and, in doing so, show an operational duality between Modes 1 and 2.

II. Problem Setup

A. Source and Channel Setup

The main BS is required to communicate a source

$$\mathbb{X} = (X_1, X_2, \ldots, X_{n_\mathrm{s}})$$

over a discrete memoryless broadcast channel to K receivers with side information; the side information at receiver k, for $k \in \{1, 2, \ldots, K\}$, is denoted by

$$\mathbb{Y}_k = (Y_{k,1}, Y_{k,2}, \ldots, Y_{k,n_\mathrm{s}}).$$

For example, $\mathbb{X}$ and $\mathbb{Y}_k$ may be the current and previous states of a mobile application, the global and local contents of a cloud storage drive, or the current and previous frames of a video feed. Alternatively, specific choices of $\mathbb{X}$ and $\mathbb{Y}_k$ lead to the bi-directional broadcast channel and complementary side information model [11]–[14]. For generality, let us only assume that the source and side information are emitted by a discrete memoryless source. That is,

$$(\mathbb{X}, \mathbb{Y}_1, \mathbb{Y}_2, \ldots, \mathbb{Y}_K) := \left\{(X_i, Y_{1,i}, Y_{2,i}, \ldots, Y_{K,i})\right\}_{i=1}^{n_\mathrm{s}}$$

is a sequence of n_s independent and identically distributed (iid) source/side-information tuples $(X, Y_1, Y_2, \ldots, Y_K)$ defined by a fixed, but arbitrary, joint probability mass function (pmf) on the Cartesian product space $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \cdots \times \mathcal{Y}_K$.

Let $\mathcal{W}$ denote the discrete memoryless broadcast channel's input alphabet, $T(\cdot|\cdot)$ its per-symbol transition probabilities, and $\mathcal{U}_k$ its output alphabet at receiver k. The main BS transmits

$$\mathbb{W} := f(\mathbb{X})$$

over the broadcast channel, where $f : \mathcal{X}^{n_\mathrm{s}} \to \mathcal{W}^{n_\mathrm{c}}$ is the BS's encoder and $\mathbb{W} = (W_1, W_2, \ldots, W_{n_\mathrm{c}})$ is a codeword with n_c symbols. Receiver k observes $\mathbb{U}_k = (U_{k,1}, U_{k,2}, \ldots, U_{k,n_\mathrm{c}})$ at the output of the broadcast channel. The ratio of channel symbols to source symbols,

$$\kappa := \frac{n_\mathrm{c}}{n_\mathrm{s}},$$

is called the *bandwidth expansion factor*.

B. No Base-Station Cooperation

Momentarily suppose that there is no BS cooperation, and that the source is to be losslessly reconstructed using only the channel outputs and side information at each receiver. In this setting, reliable communication is possible if (and only if)[1] there exists a pmf P_W on $\mathcal{W}$ such that [8]

$$H(X|Y_k) < \kappa I(W; U_k), \quad \forall\ k, \tag{1}$$

where $(W, U_1, U_2, \ldots, U_K) \sim P_W(\cdot)T(\cdot|\cdot)$. The necessity and sufficiency of (1) for reliable communication is an elegant and powerful result with applications throughout network information theory; for example, consider [11]–[18]. Indeed, a new list-decoding generalisation of (1) will play a central role in this paper.

[1]For the "only if" assertion: Replace the strict inequality $<$ in (1) with an inequality $\leq$ in the same direction.

C. Base-Station Cooperation

Let us now return to the BS cooperation model. The helper BS of receiver k, denoted BS(k), obtains side information

$$\mathbb{V}_k = (V_{k,1}, V_{k,2}, \ldots, V_{k,n_\mathrm{h}})$$

about $\mathbb{X}$ or $\mathbb{W}$ via a backbone network. Here $n_\mathrm{h} = n_\mathrm{s}$ (resp. $n_\mathrm{h} = n_\mathrm{c}$) when BS($k$) has information about $\mathbb{X}$ (resp. $\mathbb{W}$), and a precise definition of $\mathbb{V}_k$ will be given shortly. BS(k) sends

$$M_k := f_k(\mathbb{V}_k)$$

over a noiseless channel to receiver k, where $f_k : \mathcal{V}^{n_\mathrm{h}} \to \{1, 2, \ldots, \lfloor 2^{n_\mathrm{s} R_k} \rfloor\}$ is BS(k)'s encoder and R_k is its rate (in bits per source symbol). Receiver k attempts to recover the source via

$$\hat{\mathbb{X}}_k := g_k(\mathbb{U}_k, \mathbb{Y}_k, M_k),$$

where $g_k : \mathcal{U}_k^{n_\mathrm{c}} \times \mathcal{Y}_k^{n_\mathrm{s}} \times \{1, 2, \ldots, \lfloor 2^{n_\mathrm{s} R_k} \rfloor\} \longrightarrow \mathcal{X}^{n_\mathrm{s}}$ is the receiver's decoder. The collection of all encoders and decoders is called an $(n_\mathrm{s}, n_\mathrm{c}, R_1, R_2, \ldots, R_K)$-*code*.

D. Mode 1 (helper side information about the codeword $\mathbb{W}$)

Suppose that $\mathbb{V}_k$ is the entire codeword $\mathbb{W}$ or a scalar quantised version thereof. Quantisation is appropriate, for example, when the backbone network is rate limited. More formally, let $\phi_k : \mathcal{W} \to \mathcal{V}_k$ be an arbitrary but given deterministic mapping (scalar quantiser) and

$$V_{k,i} := \phi_k(W_i), \quad \forall\ i.$$

The main problem of interest here is to determine when reliable communication is achievable in the following sense.

Definition 1: Fix the bandwidth expansion factor κ, helper BS rates $\boldsymbol{R} := (R_1, \ldots, R_K)$, and scalar quantisers $\boldsymbol{\phi} := (\phi_1, \ldots, \phi_K)$. We say that a source/side information tuple $(X, Y_1, \ldots, Y_K)$ is $(\kappa, \boldsymbol{R}, \boldsymbol{\phi})$-*achievable in Mode 1* if for any $\epsilon > 0$ there exists an $(n_\mathrm{s}, n_\mathrm{c}, R_1, \ldots, R_K)$-code such that

$$n_\mathrm{c}/n_\mathrm{s} = \kappa \quad \text{and} \quad \mathbb{P}[\hat{\mathbb{X}}_k \neq \mathbb{X}] \leq \epsilon, \quad \forall\ k. \tag{2}$$

E. Mode 2 (helper side information about the source $\mathbb{X}$)

Suppose that $\mathbb{V}_k$ is directly correlated with the source and side information. That is, assume $(\mathbb{X}, \mathbb{Y}_1, \ldots, \mathbb{Y}_K, \mathbb{V}_1, \ldots, \mathbb{V}_K)$ is emitted by an arbitrary discrete memoryless source and thus is a sequence of n_s iid tuples $(X, Y_1, \ldots, Y_K, V_1, \ldots, V_K)$. We are interested in the following definition of achievability.

Definition 2: Fix the bandwidth expansion factor κ and helper BS rates $\boldsymbol{R} := (R_1, R_2, \ldots, R_K)$. We say that a source/side information tuple $(X, Y_1, \ldots, Y_K, V_1, \ldots, V_K)$ is $(\kappa, \boldsymbol{R})$-*achievable in Mode 2* if for any $\epsilon > 0$ there exists an $(n_\mathrm{s}, n_\mathrm{c}, R_1, R_2, \ldots, R_K)$-code such that (2) holds.

The problem setups of Modes 1 and 2 are special cases of the multi-relay network in [18, Sec. V]. The achievability result [18, Thm. 3] is based on a decode-and-forward protocol, and it holds for a much broader class of relay problems than Modes 1 and 2. The decode-and-forward approach is quite different to the list decoding approach used later in this paper. The results of the next section solve two previously unknown special cases of the problem in [18, Sec. V].

III. Main Results

A. Mode 1

Theorem 1: Fix the helper BS rates $\boldsymbol{R}$, bandwidth expansion factor κ and quantisers $\boldsymbol{\phi}$. A source/side-information tuple $(X, Y_1, \ldots, Y_K)$ is $(\kappa, \boldsymbol{R}, \boldsymbol{\phi})$-achievable if there exists a pmf P_W on $\mathcal{W}$ such that for all k

$$H(X|Y_k) < \kappa I(W;U_k) + \min\big\{R_k,\ \kappa I(W;V_k|U_k)\big\}, \quad (3)$$

where $(W, U_1, \ldots, U_K) \sim P_W(\cdot)T(\cdot|\cdot)$ and[2] $V_k = \phi_k(W)$. Conversely: If the source/side-information tuple is $(\kappa, \boldsymbol{R}, \boldsymbol{\phi})$-achievable, then there exists a pmf P_W on $\mathcal{W}$ such that (3) holds with inequality $\leq$ for all k.

Proof: See the full paper [26]. ■

We first notice that if the helper rates are all set to zero, then (3) becomes

$$H(X|Y_k) < \kappa I(W;U_k), \quad \forall\ k,$$

and we recover the setup of (1). If for a given pmf P_W and scalar quantisers $\boldsymbol{\phi}$ we have $R_k > \kappa H(V_k|U_k)$ for all k, then (3) evaluates to

$$H(X|Y_k) < \kappa I(W;U_k,V_k), \quad \forall\ k.$$

This latter result can be interpreted as follows. If $(\mathbb{U}_k, \mathbb{V}_k)$ can be considered as the outcome of a discrete memoryless source, then BS(k) can reliably send $\mathbb{V}_k$ to receiver k using a Slepian-Wolf code of rate R_k [20]. The receiver can then consider $(\mathbb{U}_k, \mathbb{V}_k)$ as its effective channel output, and we again return to the setup and result in (1), but where the channel output at receiver k, U_k, needs to be replaced by the pair (U_k, V_k).

For other helper rates, we note the similarity of (3) to Kim's capacity theorem [9, Thm. 1] for the primitive relay channel. Intuitively, the right hand side of (3) is the maximum rate at which information can be sent to receiver k. This intuition, however, should be treated with care because, for example, the classical Shannon approach of *strictly* separating source and channel coding is suboptimal. Nonetheless, it is natural to wonder whether Kim's simple timesharing proof of [9, Thm. 1] can be modified to prove Theorem 1. While we do not take the timesharing approach in this paper, D. Gündüz has noticed that it may indeed be possible to give such a proof of Theorem 1 using the *semiregular encoding* and *backward decoding* techniques developed in [18, App. B] (these techniques, for example, give an alternative proof of the no-cooperation case shown in (1)).

Example 1 (Broadcast capacity with helpers): Consider the bandwidth-matched case $n_{\mathrm{s}} = n_{\mathrm{c}} = n$ and $\kappa = 1$, and fix a positive rate R^*. Suppose that the receivers have no side-information, $Y_k =$ constant, and that the main BS is required to broadcast a rate R^* message M to the receivers, where M is uniformly distributed on $\{1, 2, \ldots, \lfloor 2^{nR^*} \rfloor\}$.

Given helper rates $\boldsymbol{R}$, we can define the *helper capacity* $C(\boldsymbol{R})$ to be the supremum of all achievable message rates R^*; that is, those rates R^* for which there exists a sequence of codes with vanishing probability of decoding error. It can be argued from Theorem 1 that

$$C(\boldsymbol{R}) = \max_{P_W} \min_{k} \big[I(W;U_k) + \min\big\{R_k, I(W;V_k|U_k)\big\}\big], \quad (4)$$

where the maximisation is taken over all pmfs P_W on $\mathcal{W}$.

If the channel outputs are defined over a common alphabet, say $\mathcal{U}_k = \mathcal{U}$ for all k, then (4) is a type of compound channel capacity with relays. Indeed, one recovers the compound channel capacity theorem [19] upon setting $R_k = 0$ in (4).

Example 2 (Bidirectional broadcast channel with helpers): Consider Mode 1 with two receivers for the bandwidth matched case $n_{\mathrm{s}} = n_{\mathrm{c}} = n$, and fix positive rates R_1^* and R_2^*. Recall the bidirectional setup of [11]: The main BS has two independent uniformly distributed messages M_1 and M_2 on $\{1, 2, \ldots, \lfloor 2^{nR_1^*} \rfloor\}$ and $\{1, 2, \ldots, \lfloor 2^{nR_2^*} \rfloor\}$ respectively; receiver 1 has M_1 as side information and requires M_2; and receiver 2 has M_2 as side information and requires M_1.

For fixed helper rates (R_1, R_2), we can define the *helper capacity region* $\mathcal{C}(R_1, R_2)$ to be closure of the set of all (R_1, R_2)-achievable rate pairs (R_1^*, R_2^*). It can be argued from Theorem 1 that $\mathcal{C}(R_1, R_2)$ is equal to the set of all (R_1^*, R_2^*) for which there exists a pmf P_W on $\mathcal{W}$ such that

$$R_1^* \leq I(W;U_2) + \min\big\{R_2, I(W;V_2|U_2)\big\}$$
$$R_2^* \leq I(W;U_1) + \min\big\{R_1, I(W;V_1|U_1)\big\}.$$

B. Mode 2

Theorem 2: Fix the helper BS rates $\boldsymbol{R}$ and bandwidth expansion factor κ. A source/side-information tuple $(X, Y_1, \ldots, Y_K, V_1, \ldots, V_K)$ is $(\kappa, \boldsymbol{R})$-achievable if there exists a pmf P_W on $\mathcal{W}$ and K auxiliary random variables $(A_1, \ldots, A_K)$ such that for all k we have the Markov chain $(X, Y_k) \leftrightarrow V_k \leftrightarrow A_k$,

$$R_k > I(V_k; A_k|Y_k) \quad (5a)$$

and

$$H(X|A_k, Y_k) < \kappa I(W;U_k), \quad (5b)$$

where $(W, U_1, \ldots, U_K) \sim P_W(\cdot)T(\cdot|\cdot)$. Conversely: If the source/side-information tuple is $(\kappa, \boldsymbol{R})$-achievable, then there exists a pmf P_W on $\mathcal{W}$ and auxiliary random variables $(A_1, \ldots, A_K)$ such that (5) holds with inequalities and $(X, Y_k) \leftrightarrow V_k \leftrightarrow A_k$ for all k.

Proof: See the full paper [26]. ■

When computing Theorem 2, we can assume that the alphabet of A_k has a cardinality of at most $|\mathcal{V}_k|$. The following example is inspired by Wyner's *binary helper source coding problem* [10].

Example 3: Let $\kappa = 1$, and fix $(\rho_1, \rho_2, \ldots, \rho_K) \in [0, 1/2]^K$. Suppose that the source is uniform and binary, $X \sim$ Bern(1/2);

[2] Since V_k is a function of W, we also have $I(W;V_k|U_k) = H(V_k|U_k)$.

there is no receiver side information, Y_k = constant; and helper BS(k)'s side information is

$$V_k := X + Z_k, \qquad \text{(modulo 2)},$$

where $Z_k := \text{Bern}(\rho_k)$ is independent additive binary noise.

The source/side-information tuple $(X, V_1, V_2, \ldots, V_K)$ is $(1, \boldsymbol{R})$-achievable if there exists a pmf P_W on $\mathcal{W}$ and constants $(\alpha_1, \alpha_2, \ldots, \alpha_K) \in [0, 1/2]^K$ such that

$$R_k > 1 - h(\alpha_k) \tag{6a}$$

and

$$h(\alpha_k \star \rho_k) < \kappa I(W; U_k) \tag{6b}$$

holds for all k. Here $h(\cdot)$ denotes the binary entropy function and

$$a \star b := a(1-b) + (1-a)b, \quad 0 \leq a, b \leq 1.$$

Conversely: If the source/side-information tuple is $(1, \boldsymbol{R})$-achievable, then there exists a pmf P_W on $\mathcal{W}$ and $(\alpha_1, \alpha_2, \ldots, \alpha_K) \in [0, 1/2]^K$ such that (6) holds with inequalities.

C. Source-channel separation

The single-letter expressions in Theorems 1 and 2 depend only on the marginal source and channel distributions, instead of the complete joint source-channel distribution[3] — the latter being more typical in the joint source-channel coding literature, e.g., see [25]. Nevertheless, strict separation of source and channel coding is not optimal in either mode, as can be seen for the special case of $R_1 = \ldots = R_K = 0$ in [8]. The separation of source and channel variables in Theorems 1 and 2 is best understood in the sense of *operational separation* described in [8]. Roughly, we separate the source, channel and helper codebooks as well as the encoders, but joint decoding across all three codebooks is required. More detailed discussions on the various types of source-channel separation can be found in [8, 15, 18].

D. Mixing Modes 1 and 2

Suppose that some helper BSs have information about the codeword $\mathbb{W}$, while others have information about the source $\mathbb{X}$ — a mix of Modes 1 and 2. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ denote the index sets of Mode 1 and 2 helper BSs respectively. From Theorems 1 and 2:

Example 4: A source/side-information tuple is achievable if there exists a pmf P_W on $\mathcal{W}$ and $|\mathcal{K}_2|$ auxiliary random variables $\{A_k;\ k \in \mathcal{K}_2\}$ such that

$$H(X|Y_k) < \kappa I(W; U_k) + \min\{R_k,\ \kappa I(W; V_k|U_k)\} \tag{7}$$

for all $k \in \mathcal{K}_1$, and

$$R_k > I(V_k; A_k|Y_k) \tag{8a}$$

$$H(X|A_k, Y_k) < \kappa I(W; U_k) \tag{8b}$$

[3]All of the entropy and mutual information functions in Theorems 1 and 2 depend on either the source variables or the channel variables, but not both.

and $A_k \leftrightarrow V_k \leftrightarrow (X, Y_k)$ for all $k \in \mathcal{K}_2$. Conversely: If the source/side-information tuple is achievable, then there exists a pmf P_W on $\mathcal{W}$ and $|\mathcal{K}_2|$ auxiliary random variables $\{A_k;\ k \in \mathcal{K}_2\}$ such that (7) and (8) hold with inequalities and $A_k \leftrightarrow V_k \leftrightarrow (X, Y_k)$.

IV. Slepian-Wolf Coding over Broadcast Channels with List decoding

To prove Theorems 1 and 2, it is useful to consider a list-decoding extension to (1). In this section, suppose that there is *no* BS cooperation and the receivers employ list decoding.

A. Setup and result

Let

$$\Omega(L) := \{\mathcal{L} \subseteq \mathcal{X}^{n_s} : |\mathcal{L}| = L\}$$

denote the collection of all subsets of $\mathcal{X}^{n_s}$ with cardinality L. An $(n_s, n_c, L_1, L_2, \ldots, L_K)$ *list code* is a collection of $(K+1)$ mappings $(f, g_1, g_2, \ldots, g_K)$, where

$$f : \mathcal{X}^{n_s} \longrightarrow \mathcal{W}^{n_c}$$

is the encoder at the transmitter and

$$g_k : \mathcal{U}_k^{n_c} \times \mathcal{Y}_k^{n_s} \longrightarrow \Omega(L_k)$$

is the list decoder at receiver k. Upon observing the channel output $\mathbb{U}_k$ and side information $\mathbb{Y}_k$, receiver k computes

$$\mathcal{L}_k := g_k(\mathbb{U}_k, \mathbb{Y}_k).$$

An error is declared at receiver k if $\mathbb{X} \notin \mathcal{L}_k$.

If (1) holds, then [8, Thm. 6] guarantees the existence of a sequence of list codes with $|\mathcal{L}_k| = 1$ and $\mathbb{P}[\mathbb{X} \notin \mathcal{L}_k] \to 0$ for all k. On the other hand: If (1) does not hold, then $|\mathcal{L}_k|$ must grow exponentially in n_s to ensure $\mathbb{P}[\mathbb{X} \notin \mathcal{L}_k] \to 0$. Here we are concerned with the smallest such exponent.

Definition 3: Fix the bandwidth expansion factor κ and list exponents $\boldsymbol{D} = (D_1, D_2, \ldots, D_K)$, with $D_k \geq 0$, for all k. We say that $(\kappa, \boldsymbol{D})$ is *achievable* if for any $\epsilon > 0$ there exists a $(n_s, n_c, L_1, L_2, \ldots, L_K)$ list code such that

$$\frac{n_c}{n_s} = \kappa, \tag{9a}$$

$$L_k \leq 2^{n_s D_k} \quad \text{and} \quad \mathbb{P}[\mathbb{X} \notin \mathcal{L}_k] \leq \epsilon,\ \forall\ k. \tag{9b}$$

The next lemma shows that the best exponent of receiver k's list size can be larger, but not smaller, than the equivocation (or, uncertainty) in X given Y_k minus the information conveyed over the channel.

Lemma 3: $(\kappa, \boldsymbol{D})$ is achievable if there exists a pmf P_W on $\mathcal{W}$ such that

$$D_k > \max\{H(X|Y_k) - \kappa I(W; U_k),\ 0\}, \quad \forall\ k, \tag{10}$$

where $(W, U_1, \ldots, U_K) \sim P_W(\cdot)T(\cdot|\cdot)$. Conversely: If $(\kappa, \boldsymbol{D})$ is achievable, then there exists a pmf P_W on $\mathcal{W}$ such that (10) holds with an inequality.

Proof: See the full paper [26]. ■

B. Related problems and results

Definition 3 is a lossy generalisation of the setup for (1). The standard (per-letter / average distortion) generalisation of (1) is called *"Wyner-Ziv Coding over broadcast channels"* [16], and it is a formidable open problem that includes Heegard and Berger's rate-distortion function [14, 22, 23] as well as the broadcast channel capacity region.

Definition 3 and Lemma 3 are related to Chia's recent list-decoding result [24, Prop. 1] for Heegard and Berger's rate-distortion problem [22]. For example, suppose that $\kappa = 1$ and we replace the memoryless BC $T(\cdot|\cdot)$ in our model with a noiseless source-coding 'index' channel defined on the alphabet $\{1, 2, \ldots, \lfloor 2^{n_s R_s} \rfloor\}$. In this case, the mutual information $I(W; U_k)$ transforms to the source-coding rate R_s and Lemma 3 reduces to [24, Prop. 1]

$$R_s > \max_k \left\{ H(X|Y_k) - D_k \right\}.$$

Lemma 3 is also consistent with Tuncel's result for unique decoding (1): Suppose that we are interested in unique decoding and hence the all-zero list exponent vector $\boldsymbol{D} = (0, 0, \ldots, 0)$. The reverse (converse) assertion of Lemma 3 shows that $(\kappa, \boldsymbol{D})$ is achievable *only if*

$$H(X|Y_k) \leq \kappa I(W; U_k), \quad \forall k. \tag{11}$$

The forward (achievability) assertion of Lemma 3, unfortunately, does not include the all-zero list exponent. It does, however, say the following: Any arbitrarily small positive list exponent $\boldsymbol{D}$ is achievable if (11) holds.

V. Sketch of Coding Schemes Achieving Theorems 1 and 2

In both modes, it turns out that the following approach to BS cooperation is optimal (see [26] for details): Use a good list code on the broadcast channel, and task BS(k) with helping receiver k determine which element of its decoded list $\mathcal{L}_k$ is equal to the source $\mathbb{X}$. This list will, with high probability, include $\mathbb{X}$ and have $|\mathcal{L}_k| \approx 2^{n_s D_k}$ elements. To resolve receiver k's uncertainty, BS(k) needs to encode its side information $\mathbb{V}_k$ at a rate R_k that is proportional to the list exponent D_k. In both modes, the smallest achievable rate R_k is fundamentally determined by Lemma 3.

Theorems 1 and 2 are duals in the operational sense that changing from Mode 1 to Mode 2 (or, vice versa) does not change the underlying coding problem at the main BS and the receivers — it only changes BS(k)'s approach to the problem. The side information $\mathbb{V}_k$ in Mode 2 is directly correlated with the source $\mathbb{X}$, and, in this setting, it is optimal for BS(k) to use a good source code from Wyner's 'helper' source coding problem [10]. In Mode 1, on the other hand, the side information $\mathbb{V}_k$ is a scalar quantised version of the channel codeword, and it is optimal for BS(k) to use a version of Kim's 'random-hashing' for the relay channel [9].

References

[1] A. Damnjanovic, J. Montojo, W. Yongbin, J. Tingfang, L. Tao, M. Vajapeyam, Y. Taesang, S. Osok, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun*, vol. 18, no. 3, pp. 10 – 21, 2011.

[2] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: from theory to practice," *IEEE Commun. Magazine*, vol. 50, no. 6, pp. 54 – 64, 2012.

[3] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56–61, 2006.

[4] D. Gesbert, S. Hanly, H. Huang, S. Shamai (Shitz), O. Simeone, and Y. Wei, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, 2010.

[5] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Commun.*, vol. 12, no. 2, 2005.

[6] V. Chandrasekhar and J. G. Andrews, "Spectrum allocation in tiered cellular networks," *IEEE Trans. Commun.*, vol. 57, no. 10, 2009.

[7] S. Singh, H. S. Dhillon, and J. A. Andrews, "Offloading in heterogeneous networks: modeling, analysis, and design insights," *IEEE Trans. Wireless Commun.*, vol. (to appear), 2014.

[8] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1469 –1482, 2006.

[9] Y.-H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 3, pp. 1328 –1329, 2008.

[10] A. Wyner, "On source coding with side-information at the decoder," *IEEE Trans. Inform. Theory*, vol. 21, no. 3, pp. 294–300, 1975.

[11] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 454–458, 2008.

[12] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *proc. IEEE Inform. Theory Workshop*, Lake Tahoe, California, 2007.

[13] A. Wyner, J. Wolf, and F. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1243–1249, 2002.

[14] R. Timo, A. Grant, and G. Kramer, "Lossy broadcasting with complementary side information," *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 104 – 131, 2013.

[15] D. Gunduz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3927–3944, 2009.

[16] J. Nayak, E. Tuncel, and D. Gunduz, "Wyner-Ziv coding over broadcast channels: digital schemes," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1782–1799, 2010.

[17] R. Timo, G. Lecher, L. Ong, and S. Johnson, "Multi-way relay networks: orthogonal uplink, source-channel separation and code design," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 753 – 768, 2013.

[18] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Reliable joint source-channel cooperative transmission over relay networks," *IEEE Trans. Inform. Theory*, vol. 59, no. 4, 2013.

[19] A. Lapidoth and E. Telatar, "The compound channel capacity of a class of finite-state channels," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 973–983, 1998.

[20] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[21] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[22] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, 1985.

[23] R. Timo, T. Chan and A. Grant, "Rate distortion with side-information at many decoders," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5240–5257, 2011.

[24] Y.-K. Chia, "On multiterminal source coding with list decoding constraints," in *Intl. Symp. Inform. Theory*, HI, USA, 2014.

[25] T. Cover, A. E. Gamal and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. 26, no. 6, pp. 648 – 657, 1980.

[26] R. Timo and M. Wigger, "Slepian-Wolf coding for broadcasting with cooperative base-stations," to appear in *IEEE Trans. Commun.*, 2015.