# Is Non-Unique Decoding Necessary?

Shirin Saeedi Bidokhti
EPFL, Switzerland
shirin.saeedi@epfl.ch

Vinod M. Prabhakaran
TIFR, India
vinodmp@tifr.res.in

Suhas. N. Diggavi
U.C. Los Angeles, USA
suhas@ee.ucla.edu

***Abstract*—In mutiterminal communication systems, signals carrying messages meant for different destinations are often observed together at any given destination receiver. Han and Kobayashi (1981) proposed a receiving strategy which performs a joint unique decoding of messages of interest along with a subset of messages which are not of interest. It is now well-known that this provides an achievable region which is, in general, larger than if the receiver treats all messages not of interest as noise. Nair and El Gamal (2009) and Chong, Motani, Garg, and El Gamal (2008) independently proposed a generalization called indirect or non-unique decoding where the receiver uses the codebook structure of the messages to only uniquely decode its messages of interest. Indirect (non-unique) decoding has since been used in various scenarios. The main result in this paper is to provide an interpretation and a systematic proof technique for why indirect decoding, in all known cases where it has been employed, can be replaced by a particularly designed joint unique decoding strategy, without any penalty from a rate region viewpoint[1].**

## I. Introduction

Coding schemes for multiterminal systems with many information sources and many destinations try to exploit the broadcast and interference nature of the communication media. A consequence of this is that in many schemes the signals received at a destination carry information, not only about messages that are expected to be decoded at the destination (*messages of interest*), but also about messages that are not of interest to that destination.

Standard methods in (random) code design (at the encoder) are rate splitting, superposition coding and Marton's coding [1], [2]. On the other hand, standard decoding techniques are successive decoding and joint unique decoding schemes [1], [3]. In [3], Han and Kobayashi proposed a receiving strategy which performs a joint unique decoding of messages of interest along with a subset of messages which are not of interest. We refer to a decoder with such a decoding strategy, as a joint unique decoder. It is now well-known that employing such a joint unique decoder in the code design provides an achievable region which is, in general, larger than if the receiver decodes the messages of interest while treating all messages not of interest as noise. Recently, Nair and El Gamal [4] and Chong, Motani, Garg, and El Gamal [5] independently proposed a generalization called indirect or non-unique decoding where the decoder looks for the unique messages of interest while using the codebook structure of all the messages (including the ones not of interest). Such a decoder does not uniquely decode messages not of interest, though it might narrow it to a smaller list. We refer to such a decoder, as an indirect decoder. Coding schemes which employ indirect decoders have since played a role in achievability schemes in different multi-terminal problems such as [6], [7], [8], [9], [10]. It is of interest, therefore, to see if they can achieve higher reliable transmission rates compared to codes that employ joint unique decoders. In this paper, we develop our intuition and ideas within the framework of [4]. While much of the discussion in this paper is confined to this framework, the technique applies more generally to problems studied in [8], [9], [10], as we show in [11].

In [4], the idea of indirect decoding is studied in the context of broadcast channels with degraded message sets. Nair and El Gamal consider a 3-receiver general broadcast channel where a source communicates a common message $M_0$ to three receivers $Y_1$, $Y_2$, and $Y_3$ and a private message $M_1$ only to one of the receivers, $Y_1$ (Fig 1). They characterize an inner-bound

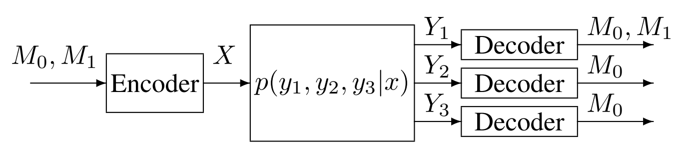


Fig. 1. The 3-receiver broadcast channel with two degraded message sets.

to the capacity region of this problem using indirect decoding and show its tightness for some special cases. It turns out that the same inner-bound of [4] can be achieved using a joint unique decoding strategy at all receivers. The equivalence of the rate region achievable by indirect decoding and that of joint unique decoding was observed in [4], but it was arrived at by comparing single letter expressions for the two rate regions[2]. This led the authors to express the hope that in general such an equivalence may not exist.

In this paper we will provide an interpretation together with a proof technique which, we believe, systematically, shows an equivalence between the rate region achievable through indirect decoders and joint unique decoders in several

[1]This work was supported in part by ERC Starting Grant Project NOWIRE ERC-2009-StG-240317. Vinod M. Prabhakaran was partially supported by a Ramanujan Fellowship from the Department of Science and Technology, Government of India. The work of Suhas Diggavi was partially supported by NSF-CPS award 1136174.

[2]A similar equivalence was also noticed in [5], again by comparing single-letter expressions. Similarly, for wireless network information flow of [6], the fact that non-unique decoding at the receiver can be replaced by joint unique decoding is implied by the combinatorial algorithms in [12], [13] for linear deterministic networks and in [14] for memoryless networks.

examples. Our technique is based on designing a special auxiliary joint unique decoder which replaces the indirect decoder and sheds some light on why this equivalence holds. This line of argument is applicable to all known instances where non-unique (indirect) decoding has been employed in the literature [11]. However, we would like to note that analysis using non-unique decoding can often give a more compact representation of the rate-region – a fact observed in [4], [5] – which still makes it a valuable tool for analysis.

## II. Why Joint Unique Decoding Suffices in the Inner-Bounds of Nair and El Gamal in [4]

We start this section by briefly reviewing the work of [4] where inner and outer bounds are derived for the capacity region of a 3-receiver broadcast channel with degraded message sets. In particular, we consider the case where a source communicates a common message (of rate $R_0$) to all receivers, and a private message (of rate $R_1$) only to one of the receivers. A coding scheme is a sequence of $((2^{nR_0}, 2^{nR_1}), n)$ codes consisting of an encoder and a decoder and is said to achieve a rate-tuple $(R_0, R_1)$ if the probability of error at the decoders tend to zero as $n \to \infty$.

*Joint unique decoder vs. Indirect decoder:* We consider typical set decoding. A decoder at a certain destination may, in general, *examine* a subset of messages which includes, but is not necessarily limited to, the messages of interest to that destination. By the term examine, we mean that the decoder will try to make use of the structure (of the codebook) associated with the messages it examines. We say a coding scheme employs a *joint unique decoder* if every decoder tries to uniquely decode all the messages they consider (and declare an error if there is ambiguity in any of the messages, irrespective of whether such messages are of interest to the destination or not). In contrast, we say that a coding scheme employs an *indirect decoder* if the decoder tries to decode uniquely only the messages of interest to the destination and tolerates ambiguity in messages which are not of interest. Within this framework, Proposition 5 of [4] establishes an achievable rate region of this problem through a coding scheme that employs an indirect decoder.

It turns out that employing a joint unique decoder, one can still achieve the same inner-bound of [4]. In this section, we develop a new proof technique to show this equivalence systematically. The same technique allows us to show the equivalence in all the examples considered in [11].

### A. Indirect decoding in the achievable scheme of Nair and El Gamal

The main problem studied in [4] is that of sending two messages over a 3-user discrete memoryless broadcast channel $p(y_1, y_2, y_3|x)$. The source intends to communicate messages $M_0$ and $M_1$ to receiver 1 and message $M_0$ to receivers 2 and 3. Rates of messages $M_0$ and $M_1$ are denoted by $R_0$ and $R_1$, respectively. In [4] an inner-bound to the capacity region is proved using a standard encoding scheme based on superposition coding and Marton's coding, and indirect (or non-unique) decoding. We briefly review this scheme.

*1) Random codebook generation and encoding:* To design the codebook, split the private message $M_1$ into four independent parts, $M_{10}$, $M_{11}$, $M_{12}$, and $M_{13}$ of non-negative rates $S_0, S_1, S_2, S_3$, respectively. Let $R_1 = S_0 + S_1 + S_2 + S_3$, $T_2 \geq S_2$ and $T_3 \geq S_3$. Fix a joint probability distribution $p(u, v_2, v_3, x)$.

Randomly and independently generate $2^{n(R_0+S_0)}$ sequences $U^n(m_0, s_0)$, $m_0 \in [1 : 2^{nR_0}]$ and $s_0 \in [1 : 2^{nS_0}]$, each distributed uniformly over the set of typical sequences $U^n$. For each sequence $U^n(m_0, s_0)$, generate randomly and conditionally independently $(i)$ $2^{nT_2}$ sequences $V_2^n(m_0, s_0, t_2)$, $t_2 \in [1 : 2^{nT_2}]$, each distributed uniformly over the set of conditionally typical sequences $V_2^n$, and $(ii)$ $2^{nT_3}$ sequences $V_3^n(m_0, s_0, t_3)$, $t_3 \in [1 : 2^{nT_3}]$, each distributed uniformly over the set of conditionally typical sequences $V_3^n$. Randomly partition sequences $V_2^n(m_0, s_0, t_2)$ into $2^{nS_2}$ bins $\mathcal{B}_2(m_0, s_0, s_2)$ and sequences $V_3^n(m_0, s_0, t_3)$ into $2^{nS_3}$ bins $\mathcal{B}_3(m_0, s_0, s_3)$. In each product bin $\mathcal{B}_2(m_0, s_0, s_2) \times \mathcal{B}_3(m_0, s_0, s_3)$, choose one (random) jointly typical sequence pair $(V_2^n(m_0, s_0, t_2), V_3^n(m_0, s_0, t_3))$. If there is no such pair, declare an error whenever the message $(m_0, s_0, s_2, s_3)$ is to be transmitted. Finally for each chosen jointly typical pair $(V_2^n(m_0, s_0, t_2), V_3^n(m_0, s_0, t_3))$ in each product bin $(s_2, s_3)$, randomly and conditionally independently generate $2^{nS_1}$ sequences $X^n(m_0, s_0, s_2, s_3, s_1)$, $s_1 \in [1 : 2^{nS_1}]$, each distributed uniformly over the set of conditionally typical $X^n$ sequences. To send the message pair $(m_0, m_1)$, where $m_1$ is expressed as $(s_0, s_1, s_2, s_3)$, the encoder sends the codeword $X^n(m_0, s_0, s_2, s_3, s_1)$.

*2) Indirect decoding:* Receiver $Y_1$ jointly uniquely decodes all messages $M_0$, $M_{10}$, $M_{11}$, $M_{12}$, and $M_{13}$. Receivers $Y_2$ and $Y_3$, however, decode $M_0$ indirectly. More precisely,

- Receiver $Y_1$ declares that the message tuple $(m_0, s_0, s_2, s_3, s_1)$ was sent if it is the unique quintuple such that the received signal $Y_1^n$ is jointly typical with $(U^n(m_0, s_0), V_2^n(m_0, s_0, t_2), V_3^n(m_0, s_0, t_3), X^n(m_0, s_0, s_2, s_3, s_1))$, where index $s_2$ is the product bin number of $V_2^n(m_0, s_0, t_2)$ and index $s_3$ is the product bin number of $V_3^n(m_0, s_0, t_3)$.
- Receiver $Y_2$ declares that the message pair $(M_0, M_{10}) = (m_0, s_0)$ was sent if it finds a unique pair of indices $(m_0, s_0)$ for which the received signal $Y_2^n$ is jointly typical with $(U^n(m_0, s_0), V_2^n(m_0, s_0, t_2))$ for some index $t_2 \in [1 : 2^{nT_2}]$.
- Receiver $Y_3$ is similar to receiver $Y_2$ with $V_3$ and $t_3$, respectively, instead of $V_2$ and $t_2$.

The above encoding/decoding scheme achieves rate pairs $(R_0, R_1)$ for which inequalities (1) to (10) below are satisfied for a joint distribution $p(u, v_2, v_3, x)$. The reader is referred to [4] for the analysis of the error probabilities.

$$R_1 = S_0 + S_1 + S_2 + S_3 \quad (1)$$

$$S_0, S_1, S_2, S_3 \geq 0,\ T_2 \geq S_2,\ T_3 \geq S_3 \quad (2)$$

$$T_2 + T_3 \geq S_2 + S_3 + I(V_2; V_3|U) \quad (3)$$

$$S_1 \leq I(X; Y_1|U, V_2, V_3) \quad (4)$$

$$S_1 + S_2 \le I(X; Y_1|U, V_3) \quad (5)$$
$$S_1 + S_3 \le I(X; Y_1|U, V_2) \quad (6)$$
$$S_1 + S_2 + S_3 \le I(X; Y_1|U) \quad (7)$$
$$R_0 + S_0 + S_1 + S_2 + S_3 \le I(X; Y_1) \quad (8)$$
$$R_0 + S_0 + T_2 \le I(U, V_2; Y_2) \quad (9)$$
$$R_0 + S_0 + T_3 \le I(U, V_3; Y_3). \quad (10)$$

### *B. Joint unique decoding suffices in the achievable scheme of Nair and El Gamal*

Fix the codebook generation and encoding scheme to be that of Section II-A. We will demonstrate how a joint unique decoding scheme suffices by following these steps:

(1) We first analyze the indirect decoder to characterize regimes where it uniquely decodes all the messages it considers and regimes where it decodes some of the messages non-uniquely.

(2) For each of the regimes, we deduce that the indirect decoder may be replaced by a joint unique decoder.

For the rest of this section, we only consider decoding schemes at receiver $Y_2$. Similar arguments are valid for receiver $Y_3$ due to the symmetry of the problem. We refer to inequality (9), which is shown in [4] to ensure reliability of the indirect decoder at receiver $Y_2$, as the indirect decoding constraint (9).

Let the rate pair $(R_0, R_1)$ be such that the indirect decoder of receiver $Y_2$ decodes message $M_0$ with high probability; i.e., the indirect decoding constraint (9) is satisfied. Consider the following two regimes:

(a) $R_0 + S_0 < I(U; Y_2)$.

(b) $R_0 + S_0 > I(U; Y_2)$,

In regime (a), it is clear from the defining condition that a joint unique decoder which decodes $(M_0, M_{10}) = (m_0, s_0)$ by finding the unique sequence $U^n(m_0, s_0)$ such that $(U^n(m_0, s_0), Y_2^n)$ is jointly typical will succeed with high probability. This is the joint unique decoder we may use in place of the indirect decoder for this regime. Notice that in this regime, while the indirect decoder obtains $(m_0, s_0)$ uniquely with high probability, it may not necessarily succeed in uniquely decoding $t_2$. Indeed, in this regime insisting on joint unique decoding of $U^n(m_0, s_0)$, $V_2^n(m_0, s_0, t_2)$ could, in some cases, result in a strictly smaller achievable region.

Regime (b) is the more interesting regime. Here it is clear that simply decoding for $(M_0, M_{10})$ and treating all other messages as noise will not work. Indirect decoding must indeed be taking advantage of the codeword $V_2^n$ as well. The indirect decoder looks for a unique pair of messages $(m_0, s_0)$ such that there exists some $t_2$ for which $(U^n(m_0, S_0), V^n(m_0, s_0, t_2), Y_2^n)$ is jointly typical. One may, in general, expect that there could be several choices of $t_2$ even in this regime. An important observation is that, in this regime, there is (with high probability) only one choice for $t_2$. In other words, *in this regime, receiver 2 decodes $t_2$ uniquely along with $m_0$ and $s_0$*. To see this, notice that using inequality (9), we have

$$T_2 \le I(V_2; Y_2|U). \quad (11)$$

Inequalities (9) and (11) together guarantee that a joint unique decoder can decode messages $M_0, M_{10}$, and $M_{12}$ with high probability; In other words, in regime (b) the indirect decoder ends up with a unique decoding of the satellite codeword $V_2^n(m_0, s_0, t_2)$ with high probability. i.e., we may replace the indirect decoder with a joint unique decoder for messages $M_0$, $M_{10}$, $M_{12}$. To summarize loosely, whenever the indirect decoder is forced to derive information from the codeword $V_2^n$ (i.e., when treating $V_2^n$ as noise will not result in correct decoding), the indirect decoder will recover this codeword also uniquely. We make this loose intuition more concrete in Section II-C.

The same argument goes through for receiver $Y_3$ and this shows that insisting on jointly uniquely decoding at all receivers is not restrictive in this problem. Thus, we arrive at the following:

*Theorem 1:* For every rate pair $(R_0, R_1)$ satisfying the inner-bound of (1)-(10), there exists a coding scheme employing a joint unique decoder which achieves the same rate pair.

The idea behind the proof of Theorem 1 was simple and general. Consider an indirect decoder which is decoding some messages of interest. The message of interest in our problem is $M_0$. Along with this message of interest, the decoder might also decode certain other messages, $M_{10}$ and $M_{12}$ for example. The two main steps of the proof is then as follows.

(1) Analyze the indirect decoder to determine what messages it decodes uniquely. Depending on the regime of operation, the indirect decoder ends up uniquely decoding a subset of its intended messages, and non-uniquely the rest of its intended messages. For example in regime (a) above, the indirect decoder uniquely decodes only $M_0$ and $M_{10}$ and it might not be able to settle on $M_{12}$. While in regime (b), the indirect decoder ends up decoding all of its three messages $M_0$, $M_{10}$, and $M_{12}$ uniquely.

(2) In each regime of operation characterized in step (1), use a joint unique decoder to only decode the messages that the indirect decoder uniquely decodes. In the above proof, this would be a joint unique decoder that decodes $M_0$ and $M_{10}$ in regime (a) and a joint unique decoder that decodes messages $M_0$, $M_{10}$, and $M_{12}$ in regime (b). Verify that the resulting joint unique decoder does support the corresponding part of the rate region achieved by the indirect decoding scheme.

Though the idea is generalizable, analyzing the indirect decoder in step (1) is a tedious task. Even for this very specific problem, it may not be entirely clear how the condition dividing cases (a) and (b) can be derived. Next, we try to resolve this using an approach which generalizes more easily.

### *C. An alternative proof to Theorem 1: an auxiliary decoder*

We take an alternative approach in this section to prove Theorem 1. The proof technique we present here has the same spirit as the proof in Section II-B, but the task of determining which subset of messages should be decoded in what regimes will be implicit rather than explicit as before. To this end, we introduce an auxiliary decoder which serves as a tool to help us develop the proof ideas. We do not propose this more complicated auxiliary decoder as a new decoding technique, but only as a proof technique to show sufficiency of joint unique decoding in the problem of [4]. We analyze the auxiliary decoder at receiver $Y_2$ and show that under the random coding experiment, it decodes correctly with high probability if the indirect decoding constraint (9) holds. From this auxiliary decoder and its performance, we will then be able to conclude that there exists a joint unique decoding scheme that succeeds with high probability.

We now define the auxiliary decoder. The auxiliary decoder at receiver $Y_2$ is a more involved decoder which has access to two component (joint unique) decoders:

- Decoder 1 is a joint unique decoder which decodes messages $M_0$ and $M_{10}$. It finds $M_0$, and $M_{10}$ by looking for the unique sequence $U^n(m_0, s_0)$ for which the pair $(U^n(m_0, s_0), Y_2^n)$ is jointly typical, and declares an error if there exists no such unique sequence.
- Decoder 2 is a joint unique decoder which decodes messages $M_0$, $M_{10}$ and $M_{12}$. It finds $M_0$, $M_{10}$, and $M_{12}$ by looking for the unique sequences $U^n(m_0, s_0)$ and $V_2^n(m_0, s_0, t_2)$ such that the triple $(U^n(m_0, s_0), V_2^n(m_0, s_0, t_2), Y_2^n)$ is jointly typical, and declares an error when such sequences do not exist.

The auxiliary decoder declares an error if either (a) both component decoders declare errors, or (b) if both of them decode and their decoded $(M_0, M_{10})$ messages do not match. In all other cases it declares the $(M_0, M_{10})$ output of a component decoder which did not declare an error as the decoded message.

We analyze the error probability under the random coding experiment of such an auxiliary decoder at receiver $Y_2$ and prove that for any $\epsilon > 0$, there is a large enough $n$ such that

$$\begin{aligned}&\Pr(\text{error at the auxiliary decoder})\\ &\leq \quad \epsilon + 2^{1+n(R_0+S_0+T_2-I(U,V_2;Y_2)+6\epsilon)}. \end{aligned} \tag{12}$$

Inequality (12) shows that for large enough $n$ and under the indirect decoding constraint (9), the auxiliary decoder has an arbitrary small probability of failure.

We start by stating the following lemma and the reader is referred to [11] for the proof.

*Lemma 1:* Fix the probability distribution $p_{U,V,Y}(u, v, y)$ and the typical set $\mathcal{T}_\epsilon^n(U, V, Y)$ corresponding to it. Consider a quadruple of sequences $(U^n, \tilde{U}^n, \hat{V}^n, Y^n)$, such that

- $\tilde{U}^n$ is independent of $(U^n, \hat{V}^n, Y^n)$ and has the distribution $\prod_i p_U(\tilde{u}_i)$,
- $U^n$ has the distribution $\prod_i p_U(u_i)$,
- $Y^n$ and $\hat{V}^n$ are independent conditioned on $U^n$,
- $(U^n, Y^n)$ has the joint distribution $\prod_i p_{U,Y}(u_i, y_i)$,
- $(U^n, \hat{V}^n)$ has the joint distribution $\prod_i p_{U,V}(u_i, \hat{v}_i)$.

Then, probability $\Pr((\tilde{U}^n, Y^n) \in \mathcal{T}_\epsilon^n(U, Y),\ (U^n, \hat{V}^n, Y^n) \in \mathcal{T}_\epsilon^n(U, V, Y))$ is upper-bounded by $2^{-n(I(U,V;Y)-6\epsilon)}$.

Assume now that $(m_0, s_0, s_1, s_2, s_3) = (1, 1, 1, 1, 1)$ is sent and indices $t_1$ and $t_2$ in the encoding procedure are $(t_2, t_3) = (1, 1)$. We analyze in the rest of this section the probability that receiver $Y_2$ declares $M_0 \neq 1$. By the symmetry of the random code construction, the conditional probability of error does not depend on which tuple of indices is sent. Thus, the conditional probability of error is the same as the unconditional probability of error and there is no loss of generality in our assumption.

Conditioned on $(m_0, s_0, s_1, s_2, s_3, t_2, t_3) = (1, 1, 1, 1, 1, 1, 1)$, receiver $Y_2$ makes an error in decoding $M_0$ only if at least one of the following events occur:

$\mathcal{E}_1$: *The channel is atypical:* the triple $(U^n(1, 1), V_2^n(1, 1, 1), Y_2^n)$ is not jointly typical.

$\mathcal{E}_2$: *The first or the second decoder (uniquely) decodes, but incorrectly:* there is a unique pair $(\tilde{m}_0, \tilde{s}_0) \neq (1, 1)$ such that the triple $(U^n(\tilde{m}_0, \tilde{s}_0), Y_2^n)$ is jointly typical, or there is a unique triple $(\hat{m}_0, \hat{s}_0, \hat{t}_2) \neq (1, 1, 1)$ such that $(U^n(\hat{m}_0, \hat{s}_0), V_2^n(\hat{m}_0, \hat{s}_0, \hat{t}_2), Y_2^n)$ is jointly typical.

$\mathcal{E}_3$: *Both decoders fail to decode uniquely and declare errors:* there are at least two distinct pairs $(\tilde{m}_0, \tilde{s}_0)$ and $(\check{m}_0, \check{s}_0)$ such that both pairs $(U^n(\tilde{m}_0, \tilde{s}_0), Y_2^n)$ and $(U^n(\check{m}_0, \check{s}_0), Y_2^n)$ are jointly typical; and similarly there are at least two distinct triples $(\hat{m}_0, \hat{s}_0, \hat{t}_2)$ and $(\check{m}_0, \check{s}_0, \check{t}_2)$ such that both triples $(U^n(\hat{m}_0, \hat{s}_0), V_2^n(\hat{m}_0, \hat{s}_0, \hat{t}_2),\ Y_2^n)$ and $(U^n(\check{m}_0, \check{s}_0), V_2^n(\check{m}_0, \check{s}_0, \check{t}_2),\ Y_2^n)$ are jointly typical.

Therefore, the probability that receiver $Y_2$ makes an error is upper-bounded in terms of the above events:

$$\begin{aligned}&\Pr(\text{error at the auxiliary decoder})\\ &\leq \quad \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2|\overline{\mathcal{E}_1}) + \Pr(\mathcal{E}_3)\\ &\leq \quad \epsilon + 0 + \Pr(\mathcal{E}_3). \end{aligned} \tag{13}$$

where (13) follows because $\Pr(\mathcal{E}_1) = \Pr((U^n(1, 1), V_2^n(1, 1, 1), Y_2^n) \notin \mathcal{T}_\epsilon^n) \leq \epsilon$ (ensured by the encoding and the Asymptotic Equipartition Property), and $\Pr(\mathcal{E}_2|\overline{\mathcal{E}_1}) = 0$. To upper-bound $\Pr(\mathcal{E}_3)$, we write

$$\begin{aligned}\Pr(\mathcal{E}_3) &\overset{(a)}{\leq} \Pr\begin{pmatrix}(U^n(\tilde{m}_0, \tilde{s}_0), Y_2^n) \in A_\epsilon^n\\ \text{for some } (\tilde{m}_0, \tilde{s}_0) \neq (1, 1)\\ (U^n(\hat{m}_0, \hat{s}_0), V_2^n(\hat{m}_0, \hat{s}_0, \hat{t}_2), Y_2^n) \in A_\epsilon^n\\ \text{for some } (\hat{m}_0, \hat{s}_0, \hat{t}_2) \neq (1, 1, 1)\end{pmatrix}\\ &\leq \Pr\begin{pmatrix}(U^n(\tilde{m}_0, \tilde{s}_0), Y_2^n) \in A_\epsilon^n\\ \text{for some } (\tilde{m}_0, \tilde{s}_0) \neq (1, 1)\\ (U^n(\hat{m}_0, \hat{s}_0), V_2^n(\hat{m}_0, \hat{s}_0, \hat{t}_2), Y_2^n) \in A_\epsilon^n\\ \text{for some } (\hat{m}_0, \hat{s}_0) \neq (1, 1) \text{ and } \ \hat{t}_2\end{pmatrix} \qquad (14)\\ &+ \Pr\begin{pmatrix}(U^n(\tilde{m}_0, \tilde{s}_0), Y_2^n) \in A_\epsilon^n\\ \text{for some } (\tilde{m}_0, \tilde{s}_0) \neq (1, 1), \text{ and all the}\\ (U^n(\hat{m}_0, \hat{s}_0), V_2^n(\hat{m}_0, \hat{s}_0, \hat{t}_2), Y_2^n) \in A_\epsilon^n \text{ are s.t.}\\ (\hat{m}_0, \hat{s}_0) = (1, 1) \text{ with at least one s.t. } \hat{t}_2 \neq 1\end{pmatrix}\end{aligned}$$

In the above chain of inequalities, $(a)$ holds because event $\mathcal{E}_3$ is a subset of the event in the right hand side.

It is worthwhile to interpret inequality (14). The error event of interest, roughly speaking, is partitioned into the following two events:

(1) The auxiliary decoder makes an error and the indirect decoder of Section II-A also makes an error.

(2) The auxiliary decoder makes an error but the indirect decoder of Section II-A decodes correctly. We will show that the probability of this event is small. Note that under this error event, (a) component decoder 1 fails (i.e., it is not possible to decode $(M_0, M_{10})$ by treating $V_2^n$ as noise), but still (b) indirect decoder succeeds (i.e., the indirect decoder must be deriving useful information by considering $V_2^n$). By showing that this error event has a small probability, we in effect show that whenever (a) and (b) hold, it is possible to jointly uniquely decode the $V_2^n$ codeword as well. This makes the rough intuition from Section II-B more concrete.

To bound the error probability we bound the two terms of inequality (14) separately. First term of (14) is bounded by the probability of the indirect decoder making an error:

$$\begin{aligned}&\Pr\begin{pmatrix} (U^n(\hat{m}_0,\hat{s}_0), V_2^n(\hat{m}_0,\hat{s}_0,\hat{t}_2), Y_2^n) \in A_\epsilon^n \\ \text{for some } (\hat{m}_0,\hat{s}_0) \neq (1,1) \text{ and } \hat{t}_2 \end{pmatrix}\\ \leq\ & 2^{nT_2} 2^{n(R_0+S_0)} 2^{-n(I(U,V_2;Y_2)-3\epsilon)}. \end{aligned} \quad (15)$$

The second term of (14) is upper-bounded as follows.

$$\begin{aligned}&\Pr\begin{pmatrix} (U^n(\tilde{m}_0,\tilde{s}_0), Y_2^n) \in A_\epsilon^n \\ \text{for some } (\tilde{m}_0,\tilde{s}_0) \neq (1,1), \text{ and all the} \\ (U^n(\hat{m}_0,\hat{s}_0), V_2^n(\hat{m}_0,\hat{s}_0,\hat{t}_2), Y_2^n) \in A_\epsilon^n \text{ are s.t.} \\ (\hat{m}_0,\hat{s}_0) = (1,1) \text{ with at least one s.t. } \hat{t}_2 \neq 1 \end{pmatrix}\\
\leq & \sum_{\hat{t}_2\neq 1,\ (\tilde{m}_0,\tilde{s}_0)\neq(1,1)} \Pr\begin{pmatrix}(U^n(\tilde{m}_0,\tilde{s}_0), Y_2^n) \in A_\epsilon^n \text{ and} \\ (U^n(1,1), V_2^n(1,1,\hat{t}_2), Y_2^n) \in A_\epsilon^n\end{pmatrix}\\
\leq\ & 2^{n(R_0+S_0+T_2)} \Pr\begin{pmatrix}(U^n(\tilde{m}_0,\tilde{s}_0), Y_2^n) \in A_\epsilon^n \text{ and} \\ (U^n(1,1), V_2^n(1,1,\hat{t}_2), Y_2^n) \in A_\epsilon^n\end{pmatrix} \quad (16)\\
\overset{(a)}{\leq}\ & 2^{n(R_0+S_0+T_2)} 2^{-n(I(U,V_2;Y_2)-6\epsilon)}, \quad (17)\end{aligned}$$

where we have $(\tilde{m}_0, \tilde{s}_0) \neq 1$ and $\hat{t}_2 \neq 1$ in the event in inequality (16) and $(a)$ follows from Lemma 1.

We conclude the error probability analysis by putting together inequalities (13), (14), (15), and (17) to obtain that the error probability at the auxiliary decoder is bounded as in inequality (12). So for large enough $n$, the auxiliary decoder succeeds with high probability if the indirect decoding constraint (9) is satisfied; i.e., when the indirect decoder succeeds with high probability.

One can now argue that if the auxiliary decoder succeeds with high probability for an operating point, then there also exists a joint unique decoding scheme that succeeds with high probability. The idea is that for all operating points (except in a subset of the rate region of measure zero), each of the two component joint unique decoders 1 and 2 have either a high or a low probability of success. So, if the operating point is such that the auxiliary decoder decodes correctly with high probability, then at least one of the component decoders should also decode correctly with high probability, giving us the joint unique decoding scheme we were looking for. This is summarized in Lemma 2, and the reader is referred to [11] for the proof.

*Lemma 2:* Given any operating point (except in a subset of the rate region of measure zero), if the auxiliary decoder succeeds with high probability under the random coding experiment, then there exists a joint unique decoding scheme that also succeeds with high probability.

A similar argument goes through for receiver $Y_3$. The random coding argument for the joint unique decoding scheme can now be completed as usual.

### D. Discussion

*Remark 1:* In Sections II-B and II-C, we did not consider cases where $R_0 + S_0 = I(U; Y_2)$ or $R_0 + S_0 = I(U; Y_3)$ (i.e., a subset of measure zero). This is enough since we may get arbitrarily close to such points.

*Remark 2:* In Sections II-B and II-C, we fixed the encoding scheme to be that of [4]. The message splitting and the structure of the codebook is therefore a priori assumed to be that of [4], even when $R_0 + S_0 < I(U; Y_2)$ and message $M_{12}$ is not jointly decoded at $Y_2$. However, in such cases this extra message structure is not required and one can consider message $M_{12}$ as a part of message $M_{11}$.

We refer the readers to [11] for details on how we may adapt this technique to several instances [8], [9], [10] of non-unique decoding in the literature. We believe this technique may be applicable more generally to show the equivalence of rate regions achievable using random coding employing an indirect decoder and a joint unique decoder.

## References


[1] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[2] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 306–311, May 1979.

[3] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, January 1981.

[4] C. Nair and A. El Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4479–4493, October 2009.

[5] H. F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the Han-Kobayashi region for the interference channel," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 3188–3195, July 2008.

[6] S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: a deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[7] S. Lim, Y.-H. Kim, A. El-Gamal, and S-Y.Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, p. 31323152, May 2011.

[8] Y. K. Chia and A. El Gamal, "3-receiver broadcast channels with common and confidential messages," June 2011. [Online]. Available: http://arxiv.org/abs/0910.1407

[9] B. Bandemer and A. El Gamal, "Interference decoding for deterministic channels," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2966–2975, May 2011.

[10] C. Nair, A. El Gamal, and Y. K. Chia, "An achievability scheme for the compound channel with state noncausally available at the encoder," April 2010. [Online]. Available: http://arxiv.org/abs/1004.3427

[11] S. Saeedi, V. M. Prabhakaran, and S. N. Diggavi, "Is non-unique decoding necessary?" February 2012. [Online]. Available: http://infoscience.epfl.ch/record/174716

[12] J. Ebrahimi and C. Fragouli, "Combinatorial algorithms for wireless information flow," *ACM Transactions on Algorithms*, 2012, see also Proc. SODA, Jan 2009 and http://arxiv.org/abs/0909.4808, 2009.

[13] M. Goemans, S. Iwata, and R. Zenklusen, "An algorithmic framework for wireless information flow," in *Allerton*, September 2009.

[14] G. Kramer and J. Hou, "On message lengths for noisy network coding," in *ITW*, October 2011, pp. 430–431.