

A Unified Approach to Hybrid Coding

Paolo Minero, *Member, IEEE*, Sung Hoon Lim, *Member, IEEE*, and Young-Han Kim, *Fellow, IEEE*

Abstract—Hybrid analog–digital coding has been used for several communication scenarios, such as joint source–channel coding of Gaussian sources over Gaussian channels and relay communication over Gaussian networks. In this paper, a generalized hybrid coding technique is proposed for communication over discrete memoryless and Gaussian systems, and its utility is demonstrated via three examples—lossy joint source–channel coding over multiple access channels, channel coding over two-way relay channels, and channel coding over diamond networks. The corresponding coding schemes recover and extend several existing results in the literature.

Index Terms—Analog–digital coding, hybrid coding, joint source–channel coding, network information theory, relay networks.

I. INTRODUCTION

THE fundamental architecture of most of today’s communication systems is inspired by Shannon’s source–channel separation theorem [1]–[4]. This fundamental theorem states that a source can be optimally communicated over a point-to-point channel by concatenating an *optimal* source coder that compresses the source into “bits” at the rate of its entropy (or rate–distortion function) with an *optimal* channel coder that communicates those “bits” reliably over the channel at the rate of its capacity. The appeal of Shannon’s separation theorem is twofold. First, it suggests a simple system architecture in which source coding and channel coding are separated by a universal digital interface. Second, it guarantees that this separation architecture does not incur any asymptotic performance loss.

The optimality of the source–channel separation architecture, however, does not extend to communication systems with multiple users. Except for a few special network models in which sources and channels are suitably “matched” [5]–[10], the problem of lossy communication over a general multiuser

network requires joint optimization of the source coding and channel coding operations. Consequently, there is a vast body of literature on joint source–channel coding schemes for multiple access channels [11]–[15], broadcast channels [16]–[26], interference channels [27], [28], and other multiuser channels [29]–[31].

Hybrid analog–digital coding has been recently shown to be an efficient joint source–channel coding technique for communicating analog Gaussian sources over memoryless Gaussian channels under a mean-square-error fidelity criterion. In the point-to-point setting, Mittal and Phamdo [32] showed that a graceful degradation of performance can be achieved at low signal-to-noise ratios by transmitting a linear combination of the source sequence and the result of its quantization using a high-dimensional Gaussian vector quantizer. Similar schemes have been proposed [33]–[36] for point-to-point channels with state and source side information at the receiver. In the multi-user setting, Lapidath and Tinguely [37], [38] showed that hybrid analog–digital coding is asymptotically optimal in the limiting regime of high signal-to-noise ratios for the problem of sending a bivariate Gaussian source over a Gaussian multiple access channel, where each transmitter has access to only one source component, while Tian, Diggavi, and Shamai [24] showed that hybrid analog–digital coding is optimal for the dual problem of sending a bivariate Gaussian source over a Gaussian broadcast channel, where each receiver is interested in only one component of the source.

Hybrid analog–digital coding has also been applied to channel coding over Gaussian relay networks. Yao and Skoglund [39] considered a Gaussian relay channel with flat fading and no channel state information at the transmitter and used hybrid analog–digital coding as a technique to increase the robustness against the unknown fading in the transmission from the relay to the destination. Khormuji and Skoglund [40] considered the Gaussian two-way relay channel and proposed three hybrid analog–digital coding schemes that outperform several existing relaying schemes. Kochman, Khina, Erez, and Zamir [41] considered a Gaussian parallel relay network with colored noise wherein the source-to-relay and relay-to-destination links have mismatched bandwidths, and used hybrid analog–digital coding techniques to compress the received signal at the relay and forward it to the destination.

In this paper, we propose a generalized hybrid coding technique as a basic building block in joint source–channel coding and channel coding for discrete memoryless systems. When adapted to the corresponding Gaussian settings, it includes several existing results as special cases and sometimes leads to strictly better performances via more flexibility in generating the digital part and in combining the analog and

Manuscript received June 3, 2013; revised July 15, 2014; accepted December 22, 2014. Date of publication February 6, 2015; date of current version March 13, 2015. This work was supported in part by the National Science Foundation under Grant CCF-1117728 and in part by the European ERC Starting under Grant 259530-ComCom. This paper was presented at the 2010 Allerton Conference on Communication, Control, and Computing, at the 2011 IEEE International Symposium of Information Theory, and at the 2013 IEEE Global Conference on Signal and Information Processing.

P. Minero is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, 46556 USA, and also with Qualcomm Inc., San Diego, CA, 92121 USA (e-mail: pminero@nd.edu).

S. H. Lim is with the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland (e-mail: sunglim@epfl.ch).

Y.-H. Kim is with the Department of Electrical and Computer Engineering, University of California at San Diego, San Diego, CA 92093 USA (e-mail: yhk@ucsd.edu).

Communicated by A. Wagner, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2015.2401000

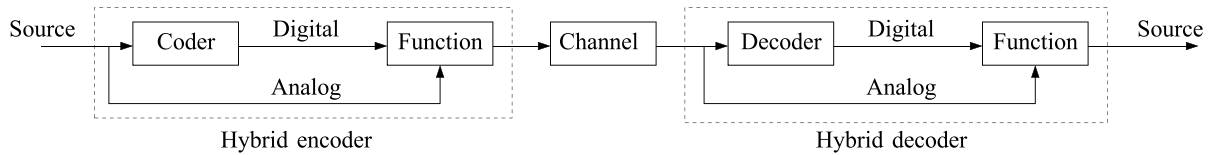


Fig. 1. A joint source-channel coding system architecture based on hybrid coding.

digital parts. The proposed hybrid coding technique employs the architecture depicted in Fig. 1 and is characterized by the following features:

- 1) An encoder generates a (digital) codeword from the (analog) source and selects the channel input as a symbol-by-symbol function of the codeword and the source.
- 2) A decoder recovers the (digital) codeword from the (analog) channel output and selects the source estimate as a symbol-by-symbol function of the codeword and the channel output.

The basic components in this architecture are not new. The idea of using a symbol-by-symbol function traces back to the Shannon strategy [42] for channels with states. The encoder structure is similar to the coding scheme by Gel'fand and Pinsker [43] for channels with state. The decoder structure is again similar to the Wyner-Ziv coding scheme [44] for lossy source coding with side information.

The main contribution of this paper lies in combining all these known techniques into a unifying framework that can be used to construct efficient coding schemes for various network communication scenarios, the performances of which can be characterized in computable (i.e., single-letter) expressions. One of the most appealing features of the resulting hybrid coding schemes is that the first-order performance analysis can be done by separately studying the conditions for source coding and for channel coding, exactly as in the source-channel separation theorem.

We illustrate the advantages of our hybrid coding framework by focusing on two specific problems:

- 1) *Joint Source-Channel Coding Over Multiple Access Channels:* We construct a joint source-channel coding scheme for lossy communications over discrete memoryless multiple access channels whereby each encoder/decoder in the network operates according to the hybrid coding architecture in Fig. 1. In Section II, we establish a sufficient condition for lossy communication that recovers and generalizes several existing results on joint source-channel coding over this channel model. In particular, when specialized to the Gaussian communication problem studied in [37], our result recovers the hybrid analog-digital scheme derived by Lapidath and Tinguely.
- 2) *Relay Networks:* Following [39], we propose a channel coding scheme for noisy relay networks based on hybrid coding in Section III. This coding scheme operates in a similar manner to the noisy network coding scheme proposed in [45], [46], except that each relay node uses the hybrid coding interface to transmit a symbol-by-symbol function of the received sequence and its compressed version. This coding scheme unifies

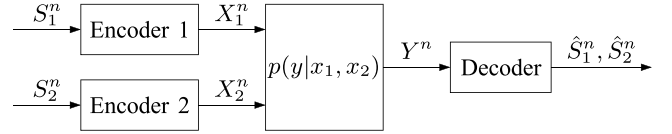


Fig. 2. Communication of a 2-DMS over a DM-MAC.

both amplify-forward [47] and compress-forward [48], and can strictly outperform both. The potential of the hybrid coding interface for relaying is demonstrated through two specific examples—communication over two-way relay channels [49] and over diamond relay networks [47].

Throughout we closely follow the notation in [50]. In particular, for a discrete random variable $X \sim p(x)$ on an alphabet \mathcal{X} and $\epsilon \in (0, 1)$, we define the set of ϵ -typical n -sequences x^n (or the typical set in short) [51] as $\mathcal{T}_\epsilon^{(n)}(X) = \{x^n : |\#\{i : x_i = x\}/n - p(x)| \leq \epsilon p(x) \text{ for all } x \in \mathcal{X}\}$. We use $\delta(\epsilon) > 0$ to denote a generic function of $\epsilon > 0$ that tends to zero as $\epsilon \rightarrow 0$. Similarly, we use $\epsilon_n \geq 0$ to denote a generic sequence in n that tends to zero as $n \rightarrow \infty$.

II. JOINT SOURCE-CHANNEL CODING OVER MULTIPLE ACCESS CHANNELS

Consider the problem of communicating a pair of correlated discrete memoryless sources (2-DMS) $(S_1, S_2) \sim p(s_1, s_2)$ over a discrete memoryless multiple access channel (DM-MAC) $p(y|x_1, x_2)$, as depicted in Fig. 2. Here each sender $j = 1, 2$ wishes to communicate in n transmissions its source S_j to a common receiver so the sources can be reconstructed within desired distortions.

- An $(|\mathcal{S}_1|^n, |\mathcal{S}_2|^n, n)$ joint source-channel code consists of
- two encoders, where encoder $j = 1, 2$ assigns a sequence $x_j^n(s_j^n) \in \mathcal{X}_j^n$ to each sequence $s_j^n \in \mathcal{S}_j^n$, and
 - a decoder that assigns an estimate $(\hat{s}_1^n, \hat{s}_2^n) \in \hat{\mathcal{S}}_1^n \times \hat{\mathcal{S}}_2^n$ to each sequence $y^n \in \mathcal{Y}^n$.

Let $d_1(s_1, \hat{s}_1)$ and $d_2(s_2, \hat{s}_2)$ be two distortion measures. A distortion pair (D_1, D_2) is said to be achievable for communication of the 2-DMS (S_1, S_2) over the DM-MAC $p(y|x_1, x_2)$ if there exists a sequence of $(|\mathcal{S}_1|^n, |\mathcal{S}_2|^n, n)$ joint source-channel codes such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{E}(d_j(S_{ji}, \hat{S}_{ji}) \leq D_j, \quad j = 1, 2.$$

The optimal distortion region is the closure of the set of all achievable distortion pairs (D_1, D_2) . A computable characterization of the optimal distortion region is not known in general.

We prove the following inner bound on the optimal distortion region.

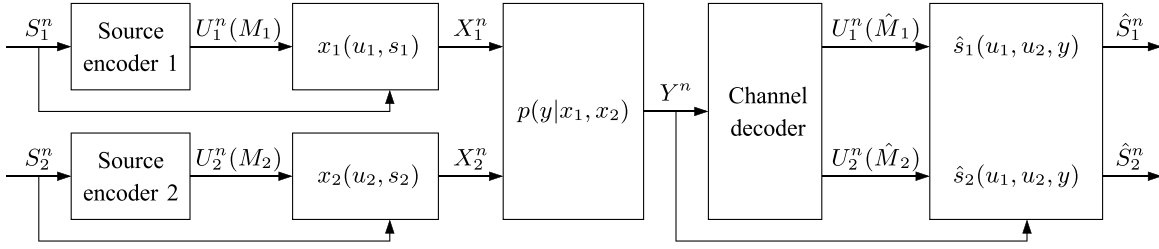


Fig. 3. The joint source–channel coding system architecture for communicating a 2-DMS over a DM-MAC.

Theorem 1: A distortion pair (D_1, D_2) is achievable for communication of the 2-DMS (S_1, S_2) over the DM-MAC $p(y|x_1, x_2)$ if

$$\begin{aligned} I(U_1; S_1|U_2, Q) &< I(U_1; Y|U_2, Q), \\ I(U_2; S_2|U_1, Q) &< I(U_2; Y|U_1, Q), \\ I(U_1, U_2; S_1, S_2|Q) &< I(U_1, U_2; Y|Q) \end{aligned}$$

for some conditional pmf $p(q)p(u_1|s_1, q)p(u_2|s_2, q)$ and functions $x_1(q, u_1, s_1)$, $x_2(q, u_2, s_2)$, $\hat{s}_1(q, u_1, u_2, y)$, and $\hat{s}_2(q, u_1, u_2, y)$ such that $\mathbf{E}(d_j(S_j, \hat{S}_j)) \leq D_j$, $j = 1, 2$.

The complete proof of the theorem is given in Appendix A, but in the next subsection we describe the gist of the proof. Here we digress a little and apply Theorem 1 to obtain the following results as special cases:

- a) *Lossless Communication:* When specialized to the case in which d_1 and d_2 are Hamming distortion measures and $D_1 = D_2 = 0$, Theorem 1 recovers the following sufficient condition for lossless communication of a 2-DMS over a DM-MAC.

Corollary 1 (Cover, El Gamal, and Salehi [11]): A 2-DMS (S_1, S_2) can be communicated losslessly over the DM-MAC $p(y|x_1, x_2)$ if

$$\begin{aligned} H(S_1|S_2) &< I(X_1; Y|X_2, S_2, Q), \\ H(S_2|S_1) &< I(X_2; Y|X_1, S_1, Q), \\ H(S_1, S_2) &< I(X_1, X_2; Y|Q) \end{aligned}$$

for some conditional pmf $p(q)p(x_1|s_1, q)p(x_2|s_2, q)$.

Proof: In Theorem 1, set $U_j = (X_j, S_j)$, $x_j(q, u_j, s_j) = x_j$, and $\hat{s}_j(q, u_1, u_2, y) = s_j$, $j = 1, 2$, under a conditional pmf of the form $p(q)p(x_1|s_1, q)p(x_2|s_2, q)$. ■

- b) *Distributed Lossy Source Coding:* When specialized to the case of a noiseless DM-MAC $Y = (X_1, X_2)$ with $H(X_1) = R_1$, $H(X_2) = R_2$, and (X_1, X_2) independent of the sources, Theorem 1 recovers the Berger–Tung inner bound on the rate–distortion region for distributed lossy source coding.

Corollary 2 (Berger [52] and Tung [53]): A distortion pair (D_1, D_2) with rate pair (R_1, R_2) is achievable for distributed lossy source coding of a 2-DMS (S_1, S_2) if

$$\begin{aligned} R_1 &> I(S_1; U_1|U_2, Q), \\ R_2 &> I(S_2; U_2|U_1, Q), \\ R_1 + R_2 &> I(S_1, S_2; U_1, U_2|Q) \end{aligned}$$

for some conditional pmf $p(q)p(u_1|s_1, q)p(u_2|s_2, q)$ and functions $\hat{s}_1(q, u_1, u_2)$ and $\hat{s}_2(q, u_1, u_2)$ such that $\mathbf{E}(d_j(S_j, \hat{S}_j)) \leq D_j$, $j = 1, 2$.

Proof: In Theorem 1, set $U_j = (X_j, \tilde{U}_j)$, $x_j(q, u_j, s_j) = x_j$, and $\hat{s}_j(q, u_1, u_2, y) = \hat{s}_j(q, \tilde{u}_1, \tilde{u}_2)$, $j = 1, 2$, under a conditional pmf of the form $p(q)p(\tilde{u}_1|s_1, q)p(\tilde{u}_2|s_2, q)$, and relabel the random variables with tilde. ■

- c) *Bivariate Gaussian Source Over a Gaussian MAC:* Suppose that the source is a bivariate Gaussian pair with equal variance σ^2 and that each source component has to be reconstructed by the decoder under quadratic (i.e., mean-square-error) distortion measures $d_j(s_j, \hat{s}_j) = (s_j - \hat{s}_j)^2$, $j = 1, 2$. In addition, assume that the channel is the Gaussian MAC $Y = X_1 + X_2 + Z$, where Z is Gaussian and the channel inputs X_1 and X_2 are subject to average power constraints. Theorem 1 can be adapted to this case via the standard discretization method [50, Secs. 3.4 and 3.8]. Suppose that in Theorem 1 we choose (U_1, U_2) as jointly Gaussian random variables conditionally independent given (S_1, S_2) , the encoding function $x_j(u_j, s_j)$, $j = 1, 2$, as a linear function of u_j and s_j , and the decoding functions $\hat{s}_j(u_1, u_2, y)$ as the minimum mean-square error (MMSE) estimate of S_j given U_1, U_2 , and Y . Then, Theorem 1 recovers the sufficient condition for lossy communication established by Lapidath and Tinguely [37, Th. IV.6] via a hybrid analog–digital scheme that combines uncoded transmission and vector quantization. We remark that the encoders used in the proof of Theorem 1 are the same as the ones described in [37], but the exact operation of the decoder is somewhat different. Lapidath and Tinguely developed a Gaussian-specific minimum-distance decoding technique to go around the issue of dependency between message and codebook. As shown in Appendix A, we develop a joint typicality decoding technique that resolves the dependency issue in general.

A. Hybrid Coding Architecture

The joint source–channel coding scheme used in the proof of Theorem 1 is based on the architecture depicted in Fig. 3.

Here the source sequence S_j^n is mapped by source encoder $j = 1, 2$ into a sequence $U_j^n(\hat{M}_j)$ from a randomly generated codebook $\mathcal{C}_j = \{U_j^n(m_j) : m_j \in [1 : 2^{nR_j}]\}$ of independently distributed codewords. The selected sequence and the source S_j^n are then mapped symbol-by-symbol through an encoding function $x_j(s_j, u_j)$ to a sequence X_j^n , which is transmitted over the MAC. Upon receiving Y^n , the decoder finds the estimates $U_1^n(\hat{M}_1)$ and $U_2^n(\hat{M}_2)$ of $U_1^n(M_1)$ and $U_2^n(M_2)$,

respectively, and reconstructs \hat{S}_1^n and \hat{S}_2^n from $U_1^n(\hat{M}_1)$, $U_2^n(\hat{M}_2)$, and Y^n by symbol-by-symbol mappings $\hat{s}_j(u_1, u_2, y)$, $j = 1, 2$.

The conditions under which a distortion pair (D_1, D_2) is achievable can be obtained by studying the conditions for source coding and channel coding separately. By the covering lemma [50, Sec. 3.7], the source encoding operation is successful if

$$\begin{aligned} R_1 &> I(U_1; S_1), \\ R_2 &> I(U_2; S_2), \end{aligned}$$

while by the packing lemma [50, Sec. 3.2], suitably modified to account for the dependence between the indices and the codebook, the channel decoding operation is successful if

$$\begin{aligned} R_1 &< I(U_1; Y, U_2), \\ R_2 &< I(U_2; Y, U_1), \\ R_1 + R_2 &< I(U_1, U_2; Y) + I(U_1; U_2). \end{aligned}$$

Then, the sufficient condition in Theorem 1 (with $Q = \emptyset$) is established by combining the above inequalities and eliminating the intermediate rate pair (R_1, R_2) . The sufficient condition with a general Q can be proved by introducing a time sharing random variable Q and using coded time sharing [50, Sec. 4.5.3].

B. Remarks

The proposed joint source–channel coding scheme is conceptually similar to separate source and channel coding and, loosely speaking, is obtained by concatenating the source coding scheme in [51] and [52] for distributed lossy source coding with a channel code for multiple access communication, except that the same codeword is used by both the source encoder and the channel encoder. Similar to the coding scheme by Cover, El Gamal, and Salehi in [11] for lossless communication over a DM-MAC, the coding scheme in Theorem 1 enables coherent communication over the MAC by preserving the correlation between the sources at the channel inputs chosen by the two senders.

The achievable distortion region in Theorem 1 can be increased when the 2-DMS (S_1, S_2) has a nontrivial common part in the sense of Gács–Körner [54] and Witsenhausen [55]. In this case, the encoders can jointly compress the common part and use it to establish coherent communication over the MAC. This extension is considered elsewhere [56], where a hybrid coding scheme is proposed by combining the distributed lossy source coding scheme in [57] for sources with a nonempty common part and the channel coding scheme in [58] for multiple access communication with a common message shared by the two encoders. The result in Theorem 1 can also be generalized to the setting in which the source consists of a random triple (S, S_1, S_2) , the distortion measures are $d_j : S \times S_1 \times S_2 \rightarrow \hat{S}_j$, $j = 1, 2$, but encoder j can only observe the source component S_j , $j = 1, 2$. This setting includes as special cases the CEO problem [59]–[62] and the Gaussian sensor network [63].

The modular approach presented here for lossy communication over multiple access channels can be adapted

to construct joint source–channel coding schemes for other channel models. In [56], extensions to several canonical channel models studied in the literature are presented—the broadcast channel and the interference channel as well as channels with state or noiseless output feedback. In all these examples, we can systematically establish sufficient conditions for lossy communication based on hybrid coding. The basic design principle lies with combining a source coding scheme with a suitably “matched” channel coding scheme.

Finally, it should be remarked that the proposed architecture can also be extended to the case of source–channel bandwidth mismatch, whereby k samples of a source are transmitted per n channel uses. This can be accomplished by replacing the source and channel symbols in Fig. 3 by supersymbols of lengths k and n (or their co-prime factors), respectively. It should be noted, however, that the computation complexity of characterizing the performances grows accordingly in co-prime k and n .

III. RELAY NETWORKS

In this section we explore applications of hybrid coding in the context of relay networks, wherein a source node wishes to send a message to a destination node with the help of intermediate relay nodes. Over the past decades, three dominant paradigms have been proposed for relay communication: decode–forward, compress–forward, and amplify–forward.

- In decode–forward, each relay recovers the transmitted message by the source either fully or partially and forwards it to the receiver (digital-to-digital interface) while coherently cooperating with the source node. Decode–forward was originally proposed in [48] for the relay channel and has been generalized to multiple relay networks, for example, in [63] and [64] and further improved by combining it with structured coding [66], [67].
- In amplify–forward, each relay sends a scaled version of its received sequence and forwards it to the receiver (analog-to-analog interface). Amplify–forward was proposed in [47] for the Gaussian two-relay diamond network and subsequently studied for the Gaussian relay channel in [68]. Generalizations of amplify–forward to general nonlinear analog mappings for relay communication have been proposed in [69].
- In compress–forward, each relay vector-quantizes its received sequence and forwards the compression index to the receiver (analog-to-digital interface). Compress–forward was proposed in [48] for the relay channel and has been generalized to arbitrary noisy networks in [45], [46] as noisy network coding.

In this section, we consider hybrid analog–digital coding at the relay nodes. This idea is originally due to Yao and Skoglund [39] in the context of Gaussian fading relay channels. Here we fully develop this idea for general discrete memoryless relay networks. The proposed scheme naturally extends both amplify–forward and compress–forward since each relay node uses the hybrid coding architecture introduced in Section II to transmit a symbol-by-symbol function of the received sequence and its quantized version

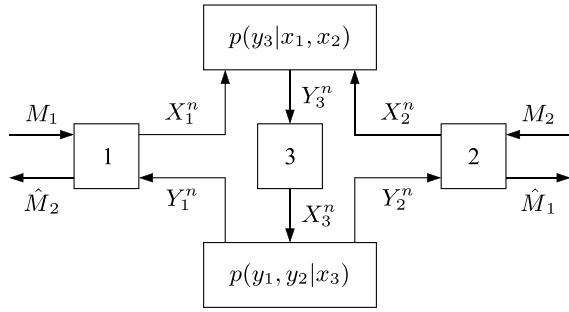


Fig. 4. The two-way relay channel.

(analog-to-analog/digital interface). More important than this conceptual unification is the performance improvement of hybrid coding. We demonstrate through two specific examples, the two-way relay channel (Section III-A) and the two-relay diamond network (Section III-B), that hybrid coding can strictly outperform the existing coding schemes.

A. Two-Way Relay Channel

Consider the relay network depicted in Fig. 4, where two source/destination nodes communication with the help of one relay. Node 1 wishes to send the message $M_1 \in [1 : 2^{nR_1}]$ to node 2 and node 2 wishes to send the message $M_2 \in [1 : 2^{nR_2}]$ to node 1 with the help of the relay node 3. Nodes 1 and 2 are connected to the relay through the MAC $p(y_3|x_1, x_2)$, while the relay is connected to nodes 1 and 2 via the broadcast channel $p(y_1, y_2|x_3)$.

This network is modeled by a 3-node discrete memoryless two-way relay channel (DM-TWRC) $p(y_1, y_2|x_3)p(y_3|x_1, x_2)$. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the DM-TWRC consists of

- two message sets $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$,
- two encoders, where at time $i \in [1 : n]$ encoder $j = 1, 2$ assigns a symbol $x_{ji}(m_j, y_j^{i-1}) \in \mathcal{X}_j$ to each message $m_j \in [1 : 2^{nR_j}]$ and past received output sequence $y_j^{i-1} \in \mathcal{Y}_j^{i-1}$,
- a relay encoder that assigns a symbol $x_{3i}(y_3^{i-1})$ to each past received output sequence $y_3^{i-1} \in \mathcal{Y}_3^{i-1}$, and
- two decoders, where decoder 1 assigns an estimate \hat{m}_2 or an error message to each m_1 and each received sequence $y_1^n \in \mathcal{Y}_1^n$ and decoder 2 assigns an estimate \hat{m}_1 or an error message to each m_2 and each received sequence $y_2^n \in \mathcal{Y}_2^n$.

We assume that the message pair (M_1, M_2) is uniformly distributed over $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. The average probability of error is defined as $P_e^{(n)} = \mathbb{P}\{(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\}$. A rate pair (R_1, R_2) is said to be achievable for the DM-TWRC if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $\lim_{n \rightarrow \infty} P_e^{(n)} = 0$. The capacity region of the DM-TWRC is the closure of the set of achievable rate pairs (R_1, R_2) and the sum-capacity is the supremum of the achievable sum rates $R_1 + R_2$.

The capacity region of the DM-TWRC is not known in general. Rankov and Wittneben [49] characterized inner bounds on the capacity region based on decode-forward, compress-forward, and amplify-forward. Another inner bound based on noisy network coding is given in [45]. In the

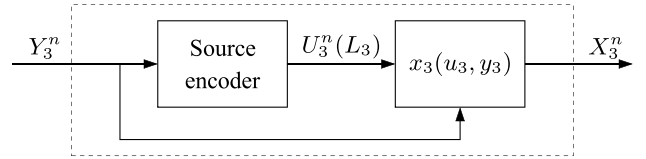


Fig. 5. The hybrid coding system architecture for the two-way relay channel.

special case of a Gaussian TWRC, Nam, Chung, and Lee [67] proposed a coding scheme based on nested lattice codes and structured binning that achieves within 1/2 bit per dimension from the capacity region for all underlying channel parameters.

Hybrid coding yields the following inner bound on the capacity region, the proof of which is given in Appendix B.

Theorem 2: A rate pair (R_1, R_2) is achievable for the DM-TWRC $p(y_1, y_2|x_3)p(y_3|x_1, x_2)$ if

$$\begin{aligned} R_1 &< \min(I(X_1; Y_2, U_3|X_2), \\ &\quad I(X_1, U_3; X_2, Y_2) - I(Y_3; U_3|X_1)), \\ R_2 &< \min(I(X_2; Y_1, U_3|X_1), \\ &\quad I(X_2, U_3; X_1, Y_1) - I(Y_3; U_3|X_2)) \end{aligned} \quad (1)$$

for some pmf $p(x_1)p(x_2)p(u_3|y_3)$ and function $x_3(u_3, y_3)$.

Remark 1: Theorem 2 includes both the noisy network coding inner bound, which is recovered by letting $U_3 = (\hat{Y}_3, X_3)$ under a pmf $p(\hat{y}_3, x_3|y_3) = p(\hat{y}_3|y_3)p(x_3)$, and the amplify-forward inner bound, which is obtained by setting $U_3 = \emptyset$, and the inclusion can be strict in general.

1) *Gaussian Two-Way Relay Channel:* As an application of Theorem 2, consider the special case of the Gaussian TWRC, where the channel outputs corresponding to the inputs X_1, X_2 , and X_3 are

$$\begin{aligned} Y_1 &= g_{13}X_3 + Z_1, \\ Y_2 &= g_{23}X_3 + Z_2, \\ Y_3 &= g_{31}X_1 + g_{32}X_2 + Z_3, \end{aligned}$$

and the noise components Z_k , $k = 1, 2, 3$, are i.i.d. $\mathcal{N}(0, 1)$. The channel gains g_{kj} from node j to node k are assumed to be real, constant over time, and known throughout the network. Each sender is subject to expected power constraint P . We denote the received SNR $S_{jk} = g_{jk}^2 P$.

Theorem 2 yields the following inner bound on the capacity region.

Corollary 3: A rate pair (R_1, R_2) is achievable for the Gaussian TWRC if

$$\begin{aligned} R_1 &< \frac{1}{2} \log \frac{\left(\frac{\alpha S_{23}(S_{31}+1)}{S_{31}+S_{32}+1} + \beta S_{23} + 1\right) (S_{31} + 1 + \sigma^2) - S_{23}a_1^2}{\left(\frac{\alpha S_{23}}{S_{31}+S_{32}+1} + \beta S_{23} + 1\right) (1 + \sigma^2) - S_{23}b^2}, \\ R_1 &< \frac{1}{2} \log \frac{\left(\frac{\alpha S_{23}(S_{31}+1)}{S_{31}+S_{32}+1} + (1-\alpha)S_{23} + 1\right) \sigma^2}{\left(\frac{\alpha S_{23}}{S_{31}+S_{32}+1} + \beta S_{23} + 1\right) (1 + \sigma^2) - S_{23}b^2}, \\ R_2 &< \frac{1}{2} \log \frac{\left(\frac{\alpha S_{13}(S_{32}+1)}{S_{31}+S_{32}+1} + \beta S_{13} + 1\right) (S_{32} + 1 + \sigma^2) - S_{13}a_2^2}{\left(\frac{\alpha S_{13}}{S_{31}+S_{32}+1} + \beta S_{13} + 1\right) (1 + \sigma^2) - S_{13}b^2}, \\ R_2 &< \frac{1}{2} \log \frac{\left(\frac{\alpha S_{13}(S_{32}+1)}{S_{31}+S_{32}+1} + (1-\alpha)S_{13} + 1\right) \sigma^2}{\left(\frac{\alpha S_{13}}{S_{31}+S_{32}+1} + \beta S_{13} + 1\right) (1 + \sigma^2) - S_{13}b^2} \end{aligned}$$

for some $\alpha, \beta \in [0, 1]$ such that $\alpha + \beta \leq 1$ and $\sigma^2 > 0$, where $a_k = \sqrt{\frac{\alpha(S_{3k+1})}{S_{31}+S_{32+1}}} + \sqrt{\beta\sigma^2}$ and $b = \sqrt{\frac{\alpha}{S_{31}+S_{32+1}}} + \sqrt{\beta\sigma^2}$.

Proof: In Theorem 2, set X_1 and X_2 as i.i.d. $\sim N(0, P)$, $U_3 = (V_3, \hat{Y}_3)$, where $\hat{Y}_3 = Y_3 + \hat{Z}_3$, \hat{Z}_3 and V_3 are i.i.d. zero-mean Gaussian independent of (X_1, X_2, Y_3) with variance σ^2 and 1, respectively, and

$$x_3(u_3, y_3) = \sqrt{\frac{\alpha P}{S_{31} + S_{32} + 1}} y_3 + \sqrt{\frac{\beta P}{\sigma^2}} (y_3 - \hat{y}_3) + \sqrt{(1 - \alpha - \beta)P} v_3, \quad (2)$$

for some $\alpha, \beta \in [0, 1]$ such that $\alpha + \beta \leq 1$. ■

Note from (2) that the channel input sequence produced by the relay node is a linear combination of the (analog) sequence Y_3 , the (digital) quantized sequence $\hat{Y}_3 = Y_3 + \hat{Z}_3$, whose resolution is determined by σ_j^2 , and the (digital) sequence V_3 . Hence by varying α and β , we can vary the amount of power allocated to the digital and analog parts in order to optimize the achievable rate region. In particular, by letting $\alpha = \beta = 0$, the hybrid coding inner bound in Corollary 3 reduces to the noisy network coding inner bound [45] that consists of all rate pairs (R_1, R_2) such that

$$R_1 < \min \left(C \left(\frac{S_{31}}{1 + \sigma^2} \right), C(S_{23}) - C(1/\sigma^2) \right), \\ R_2 < \min \left(C \left(\frac{S_{32}}{1 + \sigma^2} \right), C(S_{13}) - C(1/\sigma^2) \right) \quad (3)$$

for some $\sigma^2 > 0$. If instead we let $\alpha = 1, \beta = 0$, and $\sigma^2 \rightarrow \infty$, then the hybrid coding inner bound reduces to the amplify-forward inner bound [49] that consists of all rate pairs (R_1, R_2) such that

$$R_1 < C \left(\frac{S_{23}S_{31}}{1 + S_{23} + S_{31} + S_{32}} \right), \\ R_2 < C \left(\frac{S_{13}S_{32}}{1 + S_{13} + S_{31} + S_{32}} \right). \quad (4)$$

Similarly, by letting $\alpha = 0$ and $\beta = 1$, then the hybrid coding inner bound in Corollary 3 reduces to the set of rate pairs (R_1, R_2) such that

$$R_1 < \min \left(C \left(\frac{S_{31}(1 + S_{23})}{1 + \sigma^2 + S_{23}} \right), C \left(\frac{S_{23}\sigma^2}{1 + \sigma^2 + S_{23}} \right) - C(1/\sigma^2) \right), \\ R_2 < \min \left(C \left(\frac{S_{32}(1 + S_{13})}{1 + \sigma^2 + S_{13}} \right), C \left(\frac{S_{13}\sigma^2}{1 + \sigma^2 + S_{13}} \right) - C(1/\sigma^2) \right) \quad (5)$$

for some $\sigma^2 > 0$. Finally, by setting $\beta = 0$, Corollary 3 includes as a special case a hybrid coding scheme recently proposed in [40].

Fig. 6 compares the cutset bound [50] on the sum-capacity with the inner bound achieved by decode-forward [49], noisy network coding (3), amplify-forward (4), hybrid coding (5), the compute-forward-based scheme in [67], and the hybrid coding scheme in [40]. The plots in the figure assume that

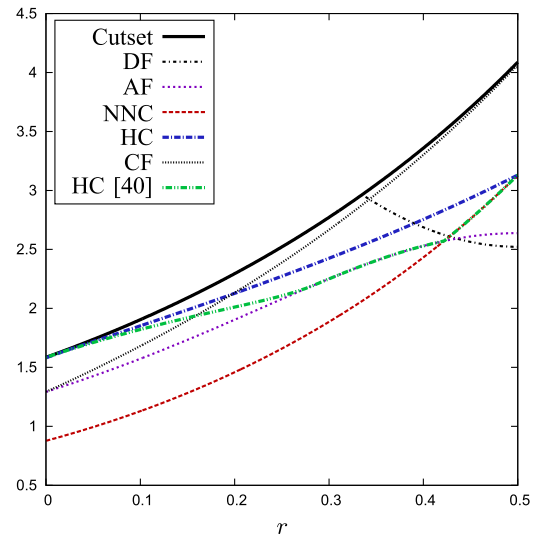


Fig. 6. Comparison of the cutset bound, the decode-forward lower bound (DF), the amplify-forward lower bound (AF), the noisy network coding lower bound (NNC), the hybrid coding lower bound (HC), the compute-forward lower bound (CF), and the hybrid coding lower bound in [40] on the sum-capacity for the Gaussian TWRC as a function of the distance r between nodes 1 and 3.

nodes 1 and 2 are unit distance apart and node 3 is at distance $r \in [0, 1]$ from node 1 along the line between nodes 1 and 2; the channel gains are of the form $g_{jk} = r_{jk}^{-3/2}$, where r_{jk} is the distance between nodes j and k , hence $g_{13} = g_{31} = r^{-3/2}$, $g_{23} = g_{32} = (1 - r)^{-3/2}$, and the power $P = 2$. Note that the hybrid coding bound in (5) strictly outperforms amplify-forward, noisy network coding, and the hybrid scheme in [40] for every $r \in (0, 1/2)$.

2) *Hybrid Coding Architecture:* The proposed relay coding scheme can be described as follows. A channel encoder at source node $j = 1, 2$ maps the message M_j into one of 2^{nR_j} sequences $X_j^n(M_j)$ generated i.i.d. according to $\prod_{i=1}^n p_{X_j}(x_{ji})$. The relay node uses the hybrid coding architecture depicted in Fig. 5. Specifically, at the relay node, the “source” sequence Y_3^n is mapped to one of 2^{nR_3} independently generated sequences $U_3^n(L_3)$ via joint typicality encoding and then the pair $(Y_3^n, U_3^n(L_3))$ is mapped to X_3^n via the symbol-by-symbol map $x_3(u_3, y_3)$. Decoding at node 1 is performed by searching for the unique message $\hat{M}_2 \in [1 : 2^{nR_2}]$ such that the tuple $(X_1^n(M_1), U_3^n(L_3), X_2^n(\hat{M}_2), Y_4^n)$ is jointly typical for some $L_3 \in [1 : 2^{nR_3}]$. In other words, node 1 nonuniquely decodes the sequence $U_3^n(L_3)$ selected by the relay node.

The conditions under which a rate pair (R_1, R_2) is achievable can be obtained by studying the conditions for channel decoding at the destinations and for hybrid encoding at the relay separately. By the covering lemma, the encoding operation at the relay node is successful if

$$R_3 > I(Y_3; U_3).$$

On the other hand, by the packing lemma, suitably modified to account for the dependence between the index and the codebook at the relay node, the channel decoding operation at

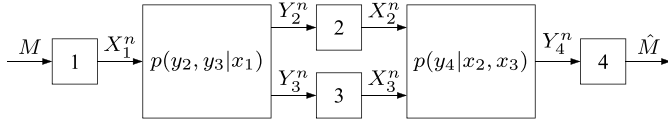


Fig. 7. The two-relay diamond network.

node 1 is successful if

$$\begin{aligned} R_2 &< I(X_2; Y_1, U_3|X_1), \\ R_2 + R_3 &< I(X_2, U_3; X_1, Y_1) + I(X_2; U_3). \end{aligned}$$

Similar conditions hold for the case of decoder 2. The lower bound (1) is then established by combining the above inequalities and eliminating the intermediate rate R_3 .

B. Diamond Relay Network

A canonical channel model used to feature the benefits of node cooperation in relay networks is the diamond network introduced in [47]; see Fig. 7. This two-hop network consists of a source node (node 1) that wishes to send a message $M \in [1 : 2^{nR}]$ to a destination (node 4) with the help of two relay nodes (nodes 2 and 3). The source node is connected through the broadcast channel $p(y_2, y_3|x_1)$ to the two relay nodes that are in turn connected to the destination node through the multiple access channel $p(y_4|x_2, x_3)$. A diamond network $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3, p(y_2, y_3|x_1)p(y_4|x_2, x_3), \mathcal{Y}_2 \times \mathcal{Y}_3 \times \mathcal{Y}_4)$ consists of six alphabet sets and a collection of conditional pmfs on $\mathcal{Y}_2 \times \mathcal{Y}_3 \times \mathcal{Y}_4$. A $(2^{nR}, n)$ code for the diamond network consists of

- a message set $[1 : 2^{nR}]$,
- an encoder that assigns a codeword $x_1^n(m)$ to each message $m \in [1 : 2^{nR}]$,
- two relay encoders, where relay encoder $j = 2, 3$ assigns a symbol $x_{ji}(y_j^{i-1})$ to each past received output sequence $y_j^{i-1} \in \mathcal{Y}_j^{i-1}$, and
- a decoder that assigns an estimate \hat{m} or an error message to each received sequence $y_4^n \in \mathcal{Y}_4^n$.

We assume that the message M is uniformly distributed over $[1 : 2^{nR}]$. The average probability of error is defined as $P_e^{(n)} = \mathbb{P}\{\hat{M} \neq M\}$. A rate R is said to be achievable for the diamond network if there exists a sequence of $(2^{nR}, n)$ codes such that $\lim_{n \rightarrow \infty} P_e^{(n)} = 0$. The capacity C of the diamond network is the supremum of the achievable rates R .

The capacity of the diamond network is not known in general. Schein and Gallager [47] characterized lower bounds on the capacity based on decode–forward and amplify–forward.

Hybrid coding yields the following lower bound on the capacity, the proof of which is given in Appendix C.

Theorem 3: The capacity of the diamond network $p(y_2, y_3|x_1)p(y_4|x_2, x_3)$ is lower bounded as

$$\begin{aligned} C \geq \max \min \{ & I(X_1; U_2, U_3, Y_4), \\ & I(X_1, U_2; U_3, Y_4) - I(U_2; Y_2|X_1), \\ & I(X_1, U_3; U_2, Y_4) - I(U_3; Y_3|X_1), \\ & I(X_1, U_2, U_3; Y_4) - I(U_2, U_3; Y_2, Y_3|X_1) \}, \end{aligned} \quad (6)$$

where the maximum is over all conditional pmfs $p(x_1)p(u_2|y_2)p(u_3|y_3)$ and functions $x_2(u_2, y_2)$, $x_3(u_3, y_3)$.

Remark 2: Theorem 3 includes both the noisy network coding lower bound, which is recovered by setting $U_j = (X_j, \hat{Y}_j)$ with $p(x_j)p(\hat{y}_j|y_j)$, $j = 2, 3$, and the amplify–forward lower bound, which is obtained by setting $U_j = \emptyset$ for $j = 2, 3$, and it can be shown that the inclusion is strict in general.

1) Deterministic Diamond Network: Consider the special case where the multiple access channel $p(y_2, y_3|x_1)$ and the broadcast channel $p(y_4|x_2, x_3)$ are deterministic, i.e., the channel outputs are functions of the corresponding inputs. In this case, Theorem 3 simplifies to the following.

Corollary 4: The capacity of the deterministic diamond network is lower bounded as

$$C \geq \max_{p(x_1)p(x_2|y_2)p(x_3|y_3)} R(Y_2, Y_3, Y_4|X_2, X_3), \quad (7)$$

where

$$\begin{aligned} R(Y_2, Y_3, Y_4|X_2, X_3) \\ = \min \{ & H(Y_2, Y_3), H(Y_2) + H(Y_4|X_2, Y_2), \\ & H(Y_3) + H(Y_4|X_3, Y_3), H(Y_4) \}. \end{aligned}$$

Proof: In Theorem 3, set $U_2 = (Y_2, X_2)$, $U_3 = (Y_3, X_3)$, $x_2(u_2, y_2) = x_2$, and $x_3(u_3, y_3) = x_3$ under a conditional pmf $p(x_2|y_2)p(x_3|y_3)$. ■

We can compare the result in Corollary 4 with the existing upper and lower bounds for this channel model. An upper bound on the capacity is given by the cutset bound [70], which in this case simplifies to

$$C \leq \max_{p(x_1)p(x_2, x_3)} R(Y_2, Y_3, Y_4|X_2, X_3) \quad (8)$$

On the other hand, specializing the scheme in [71] for deterministic relay networks, we obtain the lower bound

$$C \geq \max_{p(x_1)p(x_2)p(x_3)} R(Y_2, Y_3, Y_4|X_2, X_3). \quad (9)$$

Note that (7), (8), and (9) differ only in the set of allowed maximizing input pmfs. In particular, (7) improves upon the lower bound (9) by allowing X_2 and X_3 to depend on Y_2 and Y_3 and thereby enlarging the set of distributions on them.

2) Hybrid Coding Architecture: The proof of Theorem 3 is based on a hybrid coding architecture similar to the one used in the proof of Theorem 1 and can be described as follows. At the source node, the message M is mapped to one of 2^{nR_1} sequences $X_1^n(M)$ i.i.d. according to $\prod_{i=1}^n p_{X_1}(x_{1i})$ as in point-to-point communication. At the relay nodes, the “source” sequence Y_j^n , $j = 2, 3$, is separately mapped into one of 2^{nR_j} independently generated sequences $U_j^n(M_j)$. Then, the pair $(Y_j^n, U_j^n(M_j))$ is mapped by node j to X_j^n via a symbol-by-symbol map. By the covering lemma, the source encoding operation at the relays is successful if

$$\begin{aligned} R_2 &> I(U_2; Y_2), \\ R_3 &> I(U_3; Y_3). \end{aligned}$$

At the destination node, decoding is performed by joint typicality and indirect decoding of the sequences (U_2^n, U_3^n) ,

that is, by searching for the unique message $\hat{M} \in [1 : 2^{nR}]$ such that the tuple $(X_1^n(\hat{M}), U_2^n(\hat{M}_2), U_3^n(\hat{M}_3), Y_4^n)$ is jointly typical for some $M_2 \in [1 : 2^{nR_2}]$ and $M_3 \in [1 : 2^{nR_3}]$. By the packing lemma, combined with the technique introduced in Section II, the channel decoding operation at the destination node is successful if

$$\begin{aligned} R &< I(X_1; U_2, U_3, Y_4), \\ R + R_2 &< I(X_1, U_2; U_3, Y_4) + I(X_1; U_2), \\ R + R_3 &< I(X_1, U_3; U_2, Y_4) + I(X_1; U_3), \\ R + R_2 + R_3 &< I(X_1, U_2, U_3; Y_4) + I(X_1; U_2, U_3) \\ &\quad + I(U_2; U_3). \end{aligned}$$

Hence, the lower bound (6) is obtained by combining the conditions for source coding at the relay nodes with those for channel decoding at the destination and by eliminating the auxiliary rates R_2 and R_3 from the resulting system of inequalities.

IV. CONCLUDING REMARKS

In this paper we first studied the problem of lossy communications over multiple access channels, for which we presented a joint source–channel coding scheme based on hybrid coding that unifies and generalizes several existing results in the literature. The proposed scheme is conceptually similar to separate source and channel coding and is obtained by concatenating the source coding scheme in [51] and [52] for distributed lossy source coding with the channel coding scheme for multiple access communication, except that the same codeword is used for source coding as well as for channel coding. The same design principle can be readily adapted to other joint source–channel coding problems for which separate source coding and channel coding have matching index structures [56].

Next, we explored applications of hybrid coding in the context of relay networks. We introduced a general coding technique for DM relay networks based on hybrid coding, whereby each relay uses the hybrid coding interface to transmit a symbol-by-symbol function of the received sequence and its quantized version (analog-to-analog/digital interface). We demonstrated via two specific examples, the two-relay diamond network and the two-way relay channel, that the proposed hybrid coding can strictly outperform both amplify–forward (analog-to-analog interface) and compress–forward/noisy network coding (analog-to-digital interfaces). For simplicity, we assumed that the relay nodes do not attempt to decode the message transmitted by the source, but the presented results can be further improved by combining hybrid coding with other coding techniques such as decode–forward and structured coding [66]. In this case, hybrid coding provides a general analog/digital-to-analog/digital interface for relay communication. While we focused on two specific examples, similar ideas can be applied to general layered network model, provided that the proposed hybrid coding scheme is repeated at each layer in the network [56], [72]. In principle, hybrid coding can also be applied to the full-duplex relay channel and other nonlayered relay networks. However, in this case hybrid coding (or even amplify–forward) would not yield inner

bounds on the capacity region in a single-letter form, due to the dependency between the channel input at each relay node and the previously received analog channel outputs.

APPENDIX A PROOF OF THEOREM 1

For simplicity, we consider the case $Q = \emptyset$. Achievability for an arbitrary Q can be proved using coded time sharing [50, Sec. 4.5.3].

Codebook Generation: Let $\epsilon > \epsilon' > 0$. Fix a conditional pmf $p(u_1|s_1)p(u_2|s_2)$, channel encoding functions $x_1(u_1, s_1)$ and $x_2(u_2, s_2)$, and source decoding functions $\hat{s}_1(u_1, u_2, y)$ and $\hat{s}_2(u_1, u_2, y)$ such that $\mathbf{E}(d_j(S_j, \hat{S}_j)) \leq D_j/(1 + \epsilon)$, $j = 1, 2$. For each $j = 1, 2$, randomly and independently generate 2^{nR_j} sequences $u_j^n(m_j)$, $m_j \in [1 : 2^{nR_j}]$, each according to $\prod_{i=1}^n p_{U_j}(u_{ji})$. The codebook $\mathcal{C} = \{(u_1^n(m_1), u_2^n(m_2)) : m_1 \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]\}$ is revealed to both the encoders and the decoder.

Encoding: Upon observing a sequence s_j^n , encoder $j = 1, 2$ finds an index $m_j \in [1 : 2^{nR_j}]$ such that $(s_j^n, u_j^n(m_j)) \in \mathcal{T}_{\epsilon'}^{(n)}$. If there is more than one such index, it chooses one of them at random. If there is no such index, it chooses an index at random from $[1 : 2^{nR_j}]$. Encoder j then transmits $x_{ji} = x_j(u_{ji}(m_j), s_{ji})$ for $i \in [1 : n]$.

Decoding: Upon receiving y^n , the decoder finds the unique index pair (\hat{m}_1, \hat{m}_2) such that $(u_1^n(\hat{m}_1), u_2^n(\hat{m}_2), y^n) \in \mathcal{T}_{\epsilon}^{(n)}$ and sets the estimates as $\hat{s}_{ji} = \hat{s}_j(u_{1i}(m_1), u_{2i}(m_2), y_i)$, $i \in [1 : n]$, for $j = 1, 2$.

Analysis of the Expected Distortion: We bound the distortion averaged over (S_1^n, S_2^n) , the random choice of the codebook \mathcal{C} , and the random index assignments at the encoders. Let M_1 and M_2 be the random indices chosen at encoder 1 and at encoder 2, respectively. Define the “error” event

$$\mathcal{E} = \{(S_1^n, S_2^n, U_1^n(\hat{M}_1), U_2^n(\hat{M}_2), Y^n) \notin \mathcal{T}_{\epsilon}^{(n)}\}$$

such that the desired distortion pair is achieved if $\mathbf{P}(\mathcal{E})$ tends to zero as $n \rightarrow \infty$. Partition \mathcal{E} into the events

$$\mathcal{E}_j = \{(S_j^n, U_j^n(m_j)) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } m_j\}, \quad j = 1, 2,$$

$$\mathcal{E}_3 = \{(S_1^n, S_2^n, U_1^n(M_1), U_2^n(M_2), Y^n) \notin \mathcal{T}_{\epsilon}^{(n)}\},$$

$$\mathcal{E}_4 = \{(U_1^n(m_1), U_2^n(m_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)} \\ \text{for some } m_1 \neq M_1, m_2 \neq M_2\},$$

$$\mathcal{E}_5 = \{(U_1^n(m_1), U_2^n(M_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)} \text{ for some } m_1 \neq M_1\},$$

$$\mathcal{E}_6 = \{(U_1^n(M_1), U_2^n(m_2), Y^n) \in \mathcal{T}_{\epsilon}^{(n)} \text{ for some } m_2 \neq M_2\}.$$

Then by the union of events bound,

$$\begin{aligned} \mathbf{P}(\mathcal{E}) &\leq \mathbf{P}(\mathcal{E}_1) + \mathbf{P}(\mathcal{E}_2) + \mathbf{P}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c) \\ &\quad + \mathbf{P}(\mathcal{E}_4) + \mathbf{P}(\mathcal{E}_5) + \mathbf{P}(\mathcal{E}_6). \end{aligned}$$

By the covering lemma, $\mathbf{P}(\mathcal{E}_1)$ and $\mathbf{P}(\mathcal{E}_2)$ tend to zero as $n \rightarrow \infty$, if

$$R_1 > I(U_1; S_1) + \delta(\epsilon'), \quad (10)$$

$$R_2 > I(U_2; S_2) + \delta(\epsilon'). \quad (11)$$

By the Markov lemma [50, Sec. 12.1.1], the third term tends to zero as $n \rightarrow \infty$. By the symmetry of the random codebook

generation and encoding, we analyze the remaining probability terms conditioned on the event $\mathcal{M} = \{M_1 = 1, M_2 = 1\}$. First, by the union of events bound,

$$\begin{aligned} \mathbb{P}(\mathcal{E}_4|\mathcal{M}) &\leq \sum_{m_1=2}^{2^{nR_1}} \sum_{m_2=2}^{2^{nR_2}} \sum_{(u_1^n, u_2^n, y^n) \in \mathcal{T}_\epsilon^{(n)}} p(u_1^n) \\ &\quad \mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(m_2) = u_2^n, Y^n = y^n | \mathcal{M}\}. \end{aligned} \quad (12)$$

Let $\tilde{U}^n = (U_1^n(1), U_2^n(1), S_1^n, S_2^n)$ and $\tilde{u}^n = (\tilde{u}_1^n, \tilde{u}_2^n, s_1^n, s_2^n)$ in short. Then, by the law of total probability, for $m_1 \neq 1$ and $m_2 \neq 1$,

$$\begin{aligned} &\mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(m_2) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &= \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(m_2) = u_2^n, \\ &\quad Y^n = y^n, \tilde{U}^n = \tilde{u}^n | \mathcal{M}\} \\ &\stackrel{(a)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n | \mathcal{M}, \tilde{U}^n = \tilde{u}^n\} \\ &\quad \times \mathbb{P}\{U_2^n(m_2) = u_2^n | \mathcal{M}, U_1^n(m_1) = u_1^n, \tilde{U}^n = \tilde{u}^n\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, Y^n = y^n\} \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \\ &\stackrel{(b)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n | M_1 = 1, U_1^n(1) = \tilde{u}_1^n, S_1^n = s_1^n\} \\ &\quad \times \mathbb{P}\{U_2^n(m_2) = u_2^n | M_2 = 1, U_2^n(1) = \tilde{u}_2^n, S_2^n = s_2^n\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, Y^n = y^n\} \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \\ &\stackrel{(c)}{\leq} (1 + \epsilon) \sum_{\tilde{u}^n} \left(\prod_{i=1}^n p_{U_1}(u_{1i}) p_{U_2}(u_{2i}) \right) \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, Y^n = y^n\} \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \\ &= (1 + \epsilon) \left(\prod_{i=1}^n p_{U_1}(u_{1i}) p_{U_2}(u_{2i}) \right) \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \end{aligned} \quad (13)$$

for n sufficiently large. Here, (a) follows since given \mathcal{M} ,

$$\begin{aligned} &(U_1^n(m_1), U_2^n(m_2)) \\ &\rightarrow (S_1^n, S_2^n, U_1^n(M_1), U_2^n(M_2)) \rightarrow Y^n \end{aligned} \quad (14)$$

form a Markov chain for all $m_1 \neq 1$ and $m_2 \neq 1$, while (b) follows from the independence of the sequences and the encoding procedure. For step (c), we apply Lemma 1 at the end of this section twice (first with $U \leftarrow U_1, S \leftarrow S_1, M \leftarrow M_1$ and second with $U \leftarrow U_2, S \leftarrow S_2, M \leftarrow M_2$). Combining (12) and (13), it follows that for n sufficiently large

$$\begin{aligned} \mathbb{P}(\mathcal{E}_4|\mathcal{M}) &\leq (1 + \epsilon) \sum_{m_1=2}^{2^{nR_1}} \sum_{m_2=2}^{2^{nR_2}} \sum_{(u_1^n, u_2^n, y^n) \in \mathcal{T}_\epsilon^{(n)}} p(u_1^n) \\ &\quad \times p(u_2^n) \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \\ &\leq (1 + \epsilon) 2^{n(R_1+R_2)} \sum_{y^n \in \mathcal{T}_\epsilon^{(n)}} \mathbb{P}\{Y^n = y^n | \mathcal{M}\} \\ &\quad \times 2^{-n(I(U_1, U_2; Y) + I(U_1; U_2) - \delta(\epsilon))} \\ &\leq (1 + \epsilon) 2^{n(R_1+R_2 - I(U_1, U_2; Y) - I(U_1; U_2) + \delta(\epsilon))}. \end{aligned}$$

Hence, $\mathbb{P}(\mathcal{E}_4)$ tends to zero as $n \rightarrow \infty$ if

$$R_1 + R_2 < I(U_1, U_2; Y) + I(U_1; U_2) - \delta(\epsilon). \quad (15)$$

Following similar steps, $\mathbb{P}(\mathcal{E}_5)$ is upper bounded by

$$\begin{aligned} &\mathbb{P}\{(U_1^n(m_1), U_2^n(1), Y^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_1 \neq 1 | \mathcal{M}\} \\ &\leq \sum_{m_1=2}^{2^{nR_1}} \mathbb{P}\{U_1^n(m_1), U_2^n(1), Y^n \in \mathcal{T}_\epsilon^{(n)} | \mathcal{M}\} \\ &= \sum_{m_1=2}^{2^{nR_1}} \sum_{(u_1^n, u_2^n, y^n) \in \mathcal{T}_\epsilon^{(n)}} \mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(1) = u_2^n, \\ &\quad Y^n = y^n | \mathcal{M}\} \\ &\stackrel{(a)}{\leq} (1 + \epsilon) \sum_{m_1=2}^{2^{nR_1}} \sum_{(u_1^n, u_2^n, y^n) \in \mathcal{T}_\epsilon^{(n)}} p(u_1^n) \\ &\quad \times \mathbb{P}\{Y^n = y^n, U_2^n(1) = u_2^n | \mathcal{M}\} \\ &\leq (1 + \epsilon) 2^{nR_1} \sum_{(u_2^n, y^n) \in \mathcal{T}_\epsilon^{(n)}} \mathbb{P}\{U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &\quad \times 2^{-n(I(U_1; Y, U_2) - \delta(\epsilon))} \\ &\leq (1 + \epsilon) 2^{n(R_1 - I(U_1; Y, U_2) + \delta(\epsilon))} \end{aligned}$$

for n sufficiently large, which implies that $\mathbb{P}(\mathcal{E}_5)$ tends to zero as $n \rightarrow \infty$ if

$$R_1 < I(U_1; Y, U_2) - \delta(\epsilon). \quad (16)$$

In the above chain of inequalities step (a) is justified as follows. Let $\tilde{U}^n = (U_1^n(1), S_1^n, S_2^n)$ and $\tilde{u}^n = (\tilde{u}_1^n, s_1^n, s_2^n)$ in short. Then, by the law of total probability, for $m_1 \neq 1$,

$$\begin{aligned} &\mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &= \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n, U_2^n(1) = u_2^n, Y^n = y^n, \\ &\quad \tilde{U}^n = \tilde{u}^n | \mathcal{M}\} \\ &\stackrel{(b)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n | \mathcal{M}, U_2^n(1) = u_2^n, \tilde{U}^n = \tilde{u}^n\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, U_2^n(1) = u_2^n, Y^n = y^n\} \\ &\quad \times \mathbb{P}\{U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &\stackrel{(c)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_1^n(m_1) = u_1^n | M_1 = 1, U_1^n(1) = \tilde{u}_1^n, S_1^n = s_1^n\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, U_2^n(1) = u_2^n, Y^n = y^n\} \\ &\quad \times \mathbb{P}\{U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &\stackrel{(d)}{\leq} (1 + \epsilon) \sum_{\tilde{u}^n} \left(\prod_{i=1}^n p_{U_1}(u_{1i}) \right) \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, U_2^n(1) = u_2^n, Y^n = y^n\} \\ &\quad \times \mathbb{P}\{U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \\ &= (1 + \epsilon) \left(\prod_{i=1}^n p_{U_1}(u_{1i}) \right) \mathbb{P}\{U_2^n(1) = u_2^n, Y^n = y^n | \mathcal{M}\} \end{aligned}$$

for n sufficiently large. Here, (b) follows from the Markov chain (14) for all $m_1 \neq 1$, (c) follows from the independence of the sequences and the encoding procedure, while (d) follows by Lemma 1 at the end of this section. Finally, $\mathbb{P}(\mathcal{E}_6)$ can be bounded in a similar manner, provided that the subscripts 1 and 2 are interchanged in the upper bound for $\mathbb{P}(\mathcal{E}_5)$. It follows that $\mathbb{P}(\mathcal{E}_6)$ tends to zero as $n \rightarrow \infty$ if

$$R_2 < I(U_2; Y, U_1) - \delta(\epsilon). \quad (17)$$

Therefore, if (10), (11), (15), (16), and (17) hold, the probability of “error” tends to zero as $n \rightarrow \infty$ and the average distortions over the random codebooks is bounded as desired. Thus, there exists at least one sequence of codes achieving the desired distortions. By letting $\epsilon \rightarrow 0$ and eliminating the intermediate rate pair (R_1, R_2) , the sufficient condition in Theorem 1 (with $Q = \emptyset$) for lossy communication over a DM-MAC via hybrid coding is established.

Lemma 1: Let $(U, S) \sim p(u, s)$ and $\epsilon > \epsilon' > 0$. Let $S^n \sim \prod_{i=1}^n p_S(s_i)$ and $U^n(m)$, $m \in [1 : 2^{nR}]$, be independently generated sequences, each drawn according to $\prod_{i=1}^n p_U(u_i)$, independent of S^n . Let $\mathcal{I} = \{m \in [1 : 2^{nR}] : (U^n(m), S^n) \in \mathcal{T}_{\epsilon'}^{(n)}(U, S)\}$ be a set of random indices and let $M \sim \text{Unif}(\mathcal{I})$, if $|\mathcal{I}| > 0$, and $M \sim \text{Unif}([1 : 2^{nR}])$, otherwise. Then, for every (u^n, \tilde{u}^n, s^n) ,

$$\begin{aligned} \mathbb{P}\{U^n(2) = u^n \mid U^n(1) = \tilde{u}^n, S^n = s^n, M = 1\} \\ \leq (1 + \epsilon) \cdot \prod_{i=1}^n p_U(u_i) \end{aligned}$$

for n sufficiently large.

Proof: Given \tilde{u}^n and s^n , let $\mathcal{A} = \{U^n(1) = \tilde{u}^n, S^n = s^n\}$ in short. Let $\mathcal{C}' = \{U^n(m) : m \in [3 : 2^{nR}]\}$. Then, by the law of total probability and the Bayes rule, for every u^n ,

$$\begin{aligned} \mathbb{P}\{U^n(2) = u^n \mid M = 1, \mathcal{A}\} \\ &= \sum_{\mathcal{C}'} \mathbb{P}\{\mathcal{C}' = c', U^n(2) = u^n \mid M = 1, \mathcal{A}\} \\ &= \sum_{\mathcal{C}'} \mathbb{P}\{\mathcal{C}' = c' \mid M = 1, \mathcal{A}\} \mathbb{P}\{U^n(2) = u^n \mid \mathcal{A}, \mathcal{C}' = c'\} \\ &\quad \times \frac{\mathbb{P}\{M = 1 \mid \mathcal{A}, U^n(2) = u^n, \mathcal{C}' = c'\}}{\mathbb{P}\{M = 1 \mid \mathcal{A}, \mathcal{C}' = c'\}} \\ &= \sum_{\mathcal{C}'} \mathbb{P}\{\mathcal{C}' = c' \mid M = 1, \mathcal{A}\} \left(\prod_{i=1}^n p_U(u_i) \right) \\ &\quad \times \frac{\mathbb{P}\{M = 1 \mid \mathcal{A}, U^n(2) = u^n, \mathcal{C}' = c'\}}{\mathbb{P}\{M = 1 \mid \mathcal{A}, \mathcal{C}' = c'\}}. \end{aligned} \quad (18)$$

For each $(\tilde{u}^n, s^n, u^n, c')$ such that $\mathbb{P}\{\mathcal{C}' = c' \mid M = 1, \mathcal{A}\} > 0$, let

$$\begin{aligned} n(\tilde{u}^n, s^n, u^n, c') &= n(s^n, c') \\ &= |\{u^n \in c' : (u^n, s^n) \in \mathcal{T}_{\epsilon'}^{(n)}\}| \end{aligned}$$

denote the number of unique sequences in c' that are jointly typical with s^n and let

$$i(\tilde{u}^n, s^n, u^n, c') = i(\tilde{u}^n, s^n, c') = \begin{cases} 1, & (\tilde{u}^n, s^n) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ and } n(s^n, c') = 0, \\ 0, & \text{otherwise,} \end{cases}$$

be the indicator function for the case that neither \tilde{u}^n nor any codeword in c' is jointly typical with s^n . Then, by the way the random index M is generated, it can be easily verified that

$$\begin{aligned} \mathbb{P}\{M = 1 \mid \mathcal{A}, U^n(2) = u^n, \mathcal{C}' = c'\} \\ \leq \frac{1}{2^{nR}} i(\tilde{u}^n, s^n, c') + \frac{1}{n(s^n, c') + 1} (1 - i(\tilde{u}^n, s^n, c')). \end{aligned}$$

Similarly, since $U^n(2) \sim \prod_{i=1}^n p_U(u_i)$, independent of S^n and $U^n(m)$, $m \neq 2$,

$$\begin{aligned} \mathbb{P}\{M = 1 \mid \mathcal{A}, \mathcal{C}' = c'\} \\ &\geq \mathbb{P}\{M = 1 \mid \mathcal{A}, \mathcal{C}' = c', 2 \notin \mathcal{I}\} \cdot \mathbb{P}\{2 \notin \mathcal{I} \mid \mathcal{A}, \mathcal{C}' = c'\} \\ &\geq \mathbb{P}\{M = 1 \mid \mathcal{A}, \mathcal{C}' = c', 2 \notin \mathcal{I}\} \left(1 - 2^{-n(I(U;S) - \delta(\epsilon'))}\right) \\ &= \left(\frac{1}{2^{nR}} i(\tilde{u}^n, s^n, c') + \frac{1}{n(s^n, c') + 1} (1 - i(\tilde{u}^n, s^n, c')) \right) \\ &\quad \times (1 - 2^{-n(I(U;S) - \delta(\epsilon'))}). \end{aligned}$$

It follows that

$$\frac{\mathbb{P}\{M = 1 \mid U^n(2) = u^n, E, \mathcal{C}' = c'\}}{\mathbb{P}\{M = 1 \mid E, \mathcal{C}' = c'\}} \leq \frac{1}{1 - 2^{-n(I(U;S) - \delta(\epsilon'))}} \leq 1 + \epsilon \quad (19)$$

for n sufficiently large. By combining (18) and (19), the claim follows. \blacksquare

APPENDIX B PROOF OF THEOREM 2

We use b transmission blocks, each consisting of n transmissions, as in the proof of the multihop lower bound for the relay channel [50, Sec. 16.4.1]. A sequence of $(b-1)$ message pairs $(M_{1j}, M_{2j}) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$, $j \in [1 : b-1]$, each selected independently and uniformly over $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ is sent over b blocks. Note that the average rate pair over the b blocks is $((b-1)/b)(R_1, R_2)$, which can be made arbitrarily close to (R_1, R_2) by letting $b \rightarrow \infty$.

Codebook Generation: Let $\epsilon > \epsilon' > 0$. Fix a conditional pmf $p(x_1)p(x_2)p(u_3|y_3)$ and an encoding function $x_3(u_3, y_3)$. We randomly and independently generate a codebook for each block. For $j \in [1 : b]$, randomly and independently generate 2^{nR_3} sequences $u_3^n(l_{3j})$, $l_{3j} \in [1 : 2^{nR_3}]$, each according to $\prod_{i=1}^n p_{U_3}(u_{3i})$. For each $k = 1, 2$, randomly and independently generate 2^{nR_k} sequences $x_k^n(m_{kj})$, $m_{kj} \in [1 : 2^{nR_k}]$, each according to $\prod_{i=1}^n p_{X_k}(x_{ki})$. This defines the codebook

$$\begin{aligned} \mathcal{C}_j &= \{(x_1^n(m_{1j}), x_2^n(m_{2j}), u_3^n(l_{3j})) : \\ &\quad m_1 \in [1 : 2^{nR_1}], m_2 \in [1 : 2^{nR_2}], l_3 \in [1 : 2^{nR_3}]\} \end{aligned}$$

for $j \in [1 : b]$.

Encoding: Let $m_{kj} \in [1 : 2^{nR_k}]$ be the message to be sent in block $j \in [1 : b-1]$ by node $k = 1, 2$. Then, node k transmits $x_k^n(m_{kj})$ from codebook \mathcal{C}_j .

Relay Encoding: Upon receiving $y_3^n(j)$ in block $j \in [1 : b-1]$, relay node 3 finds an index l_{3j} such that $(u_3^n(l_{3j}), y_3^n(j)) \in \mathcal{T}_{\epsilon'}^{(n)}$. If there is more than one index, it chooses one of them at random. If there is no such index, it chooses an index at random from $[1 : 2^{nR_3}]$. In block $j+1$, relay 3 transmits $x_{3i} = x_3(u_{3i}(l_{3j}), y_{3i}(j))$ for $i \in [1 : n]$.

Decoding: Upon receiving $y_1^n(j)$, $j \in [2 : b]$, decoder 1 finds the unique message $\hat{m}_{2,j-1}$ such that

$$(x_1^n(m_{1,j-1}), u_3^n(l_{3,j-1}), x_2^n(\hat{m}_{2,j-1}), y_1^n(j)) \in \mathcal{T}_{\epsilon'}^{(n)}$$

for some $l_{3,j-1} \in [1 : 2^{nR_3}]$. Decoding at node 2 is performed in a similar manner.

Analysis of the Probability of Error: We analyze the probability of decoding error at node 1 in block $j = 2, \dots, b$,

averaged over the random codebooks and index assignment in the encoding procedure at the relay. Let $L_{3,j-1}$ be the random index chosen in block $j-1$ at relay 3. Decoder 1 makes an error only if one or more of the following events occur:

$$\begin{aligned}\mathcal{E}_1 &= \{(Y_3^n(j-1), U_3^n(l)) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } l\}, \\ \mathcal{E}_2 &= \{(X_1^n(M_{1,j-1}), U_3^n(L_{3,j-1}), \\ &\quad X_2^n(M_{2,j-1}), Y_1^n(j)) \notin \mathcal{T}_{\epsilon}^{(n)}\}, \\ \mathcal{E}_3 &= \{(X_1^n(M_{1,j-1}), U_3^n(L_{3,j-1}), X_2^n(m), Y_1^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } m \neq M_{2,j-1}\}, \\ \mathcal{E}_4 &= \{(X_1^n(M_{1,j-1}), U_3^n(l), X_2^n(m), Y_1^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } m \neq M_{2,j-1}, l \neq L_{3,j-1}\}.\end{aligned}$$

Then, by the union of events bound the probability of decoding error is upper bounded as

$$\begin{aligned}\mathbf{P}(\hat{M}_{2,j-1} \neq M_{2,j-1}) &\leq \mathbf{P}(\mathcal{E}_1) + \mathbf{P}(\mathcal{E}_2 \cap \mathcal{E}_1^c) \\ &\quad + \mathbf{P}(\mathcal{E}_3) + \mathbf{P}(\mathcal{E}_4).\end{aligned}$$

By the covering lemma, $\mathbf{P}(\mathcal{E}_1)$ tends to zero as $n \rightarrow \infty$, if

$$R_3 > I(U_3; Y_3) + \delta(\epsilon').$$

By the Markov lemma, $\mathbf{P}(\mathcal{E}_2 \cap \mathcal{E}_1^c)$ tends to zero as $n \rightarrow \infty$. By the symmetry of the random codebook generation and the random index assignment at the relays, it suffices to consider the conditional probabilities of the remaining error events conditioned on the event that

$$\mathcal{M} = \{M_{1,j-1} = 1, M_{2,j-1} = 1, L_{3,j-1} = 1\}. \quad (20)$$

Then, by the packing lemma, $\mathbf{P}(\mathcal{E}_3)$ tends to zero as $n \rightarrow \infty$ if

$$R_2 < I(X_2; Y_1, U_3|X_1) - \delta(\epsilon).$$

Next, for n sufficiently large, $\mathbf{P}(\mathcal{E}_4)$ is upper bounded by

$$\begin{aligned}&\mathbf{P}\{(X_1^n(1), U_3^n(l), X_2^n(m), Y_1^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } l \neq 1, m \neq 1 | \mathcal{M}\} \\ &\leq \sum_{m=2}^{2^{nR_2}} \sum_{l=2}^{2^{nR_3}} \sum_{(x_1^n, u_3^n, x_2^n, y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}} \\ &\quad \mathbf{P}\{X_1^n(1) = x_1^n, U_3^n(l) = u_3^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &\stackrel{(a)}{\leq} (1 + \epsilon) \sum_{m=2}^{2^{nR_2}} \sum_{l=2}^{2^{nR_3}} \sum_{(x_1^n, u_3^n, x_2^n, y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}} p(u_3^n) \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &= (1 + \epsilon) \sum_{m=2}^{2^{nR_2}} \sum_{l=2}^{2^{nR_3}} \sum_{(x_1^n, u_3^n, x_2^n, y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}} p(u_3^n) \cdot p(x_2^n) \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &\leq (1 + \epsilon) 2^{n(R_2+R_3)} \\ &\quad \cdot \sum_{(x_1^n, y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}} \mathbf{P}\{X_1^n(1) = x_1^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &\quad \times 2^{-n(I(X_2, U_3; Y_1, X_1) + I(X_2; U_3) - \delta(\epsilon))} \\ &\leq (1 + \epsilon) 2^{n(R_2+R_3 - I(X_2, U_3; Y_1, X_1) - I(X_2; U_3) + \delta(\epsilon))}.\end{aligned}$$

Here, step (a) is justified as follows. Let $\tilde{U}^n = (U_3^n(1), Y_3^n(j-1))$ and $\tilde{u}^n = (\tilde{u}_3^n, \tilde{y}_3^n)$ in short. Then, by the law of total probability, for $l \neq 1$ and $m \neq 1$,

$$\begin{aligned}&\mathbf{P}\{X_1^n(1) = x_1^n, U_3^n(l) = u_3^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &= \sum_{\tilde{u}^n} \mathbf{P}\{X_1^n(1) = x_1^n, U_3^n(l) = u_3^n, X_2^n(m) = x_2^n, \\ &\quad Y_1^n(j) = y_1^n, \tilde{U}^n = \tilde{u}^n | \mathcal{M}\} \\ &\stackrel{(b)}{=} \sum_{\tilde{u}^n} \mathbf{P}\{U_3^n(l) = u_3^n | \mathcal{M}, X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, \\ &\quad \tilde{U}^n = \tilde{u}^n\} \\ &\quad \times \mathbf{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n\} \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &\stackrel{(c)}{=} \sum_{\tilde{u}^n} \mathbf{P}\{U_3^n(l) = u_3^n | L_{3,j-1} = 1, U_3^n(1) = \tilde{u}_3^n, \\ &\quad Y_3^n(j-1) = \tilde{y}_3^n\} \\ &\quad \times \mathbf{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n\} \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &\stackrel{(d)}{\leq} (1 + \epsilon) \sum_{\tilde{u}^n} \left(\prod_{i=1}^n p_{U_3}(u_{3i}) \right) \\ &\quad \times \mathbf{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n\} \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\} \\ &= (1 + \epsilon) \left(\prod_{i=1}^n p_{U_1}(u_{1i}) \right) \\ &\quad \times \mathbf{P}\{X_1^n(1) = x_1^n, X_2^n(m) = x_2^n, Y_1^n(j) = y_1^n | \mathcal{M}\}\end{aligned}$$

for n sufficiently large, where (b) follows since given \mathcal{M} , $U_3^n(l) \rightarrow (X_1^n(1), X_2^n(m), U_3^n(L_{3,j-1}), Y_3^n(j-1)) \rightarrow Y_1^n(j)$ form a Markov chain for all $l \neq 1$ and $m \neq 1$, step (c) follows from the independence of the sequences and the encoding procedure, while (d) follows by Lemma 1. It follows that $\mathbf{P}(\mathcal{E}_4)$ tends to zero as $n \rightarrow \infty$ if

$$R_2 + R_3 < I(X_2, U_3; X_1, Y_1) + I(X_2; U_3) - \delta(\epsilon).$$

By similar steps, the decoding error probability at node 2 tends to zero as $n \rightarrow \infty$ if $R_1 < I(X_1, U_3; Y_2|X_2) - \delta(\epsilon)$ and

$$R_1 + R_3 < I(X_1, U_3; X_2, Y_2) + I(X_1; U_3) - \delta(\epsilon).$$

Finally, by eliminating R_3 from the above inequalities, the probability of error tends to zero as $n \rightarrow \infty$ if the conditions in Theorem 2 are satisfied.

APPENDIX C PROOF OF THEOREM 3

As in the previous section, we use block Markov coding with b transmission blocks and communicate a sequence of $(b-1)$ messages $M_j \in [1 : 2^{nR}]$, $j \in [1 : b-1]$.

Codebook Generation: Let $\epsilon > \epsilon' > 0$. Fix a conditional pmf $p(x_1)p(u_2|y_2)p(u_3|y_3)$ and two encoding functions $x_2(u_2, y_2)$ and $x_3(u_3, y_3)$. We randomly and independently generate a codebook for each block. For $j \in [1 : b]$, randomly and independently generate 2^{nR} sequences $x_1^n(m_j)$, $M_j \in [1 : 2^{nR}]$, each according to $\prod_{i=1}^n p_{X_1}(x_{1i})$. For each $k = 2, 3$, randomly and independently generate 2^{nR_k} sequences $u_k^n(l_{kj})$, $l_{kj} \in [1 : 2^{nR_k}]$, each according to $\prod_{i=1}^n p_{U_k}(u_{ki})$.

Encoding: Let $m_j \in [1 : 2^{nR}]$ be the message to be sent in block $j \in [1 : b-1]$. Then, the source node transmits $x_1^n(m_j)$.

Relay Encoding: Upon receiving $y_k^n(j)$ in block $j \in [1 : b-1]$, relay node $k = 2, 3$ finds an index l_{kj} such that $(u_k^n(l_{kj}), y_k^n(j)) \in \mathcal{T}_{\epsilon'}^{(n)}$. If there is more than one index, it chooses one of them at random. If there is no such index, it chooses an index at random from $[1 : 2^{nR_k}]$. In block $j+1$, relay k transmits $x_{ki} = x_k(u_{ki}(l_{kj}), y_{ki}(j))$ for $i \in [1 : n]$.

Decoding: Upon receiving $y_4^n(j)$, $j \in [2 : b]$, the decoder finds the unique message \hat{m}_{j-1} such that

$$(x_1^n(\hat{m}_{j-1}), u_2^n(l_{2,j-1}), u_3^n(l_{3,j-1}), y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)},$$

for some $l_{2,j-1} \in [1 : 2^{nR_2}]$ and $l_{3,j-1} \in [1 : 2^{nR_3}]$.

Analysis of the Probability of Error: We analyze the probability of decoding error for the message M_{j-1} in block $j = 2, \dots, b$, averaged over the random codebooks and index assignments. Let $L_{2,j-1}$ and $L_{3,j-1}$ be the random indices chosen in block $j-1$ at relay nodes 2 and 3, respectively. The decoder makes an error only if one or more of the following events occur:

$$\begin{aligned} \mathcal{E}_1 &= \{(Y_2^n(j-1), U_2^n(l_2)) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } l_2\}, \\ \mathcal{E}_2 &= \{(Y_3^n(j-1), U_3^n(l_3)) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } l_3\}, \\ \mathcal{E}_3 &= \{(X_1^n(M_{j-1}), U_2^n(L_{2,j-1}), U_3^n(L_{3,j-1}), Y_4^n(j)) \notin \mathcal{T}_{\epsilon}^{(n)}\}, \\ \mathcal{E}_4 &= \{(X_1^n(m), U_2^n(L_{2,j-1}), U_3^n(L_{3,j-1}), Y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } m \neq M_{j-1}\}, \\ \mathcal{E}_5 &= \{(X_1^n(m), U_2^n(l_2), U_3^n(L_{3,j-1}), Y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } l_2 \neq L_{2,j-1}, m \neq M_{j-1}\}, \\ \mathcal{E}_6 &= \{(X_1^n(m), U_2^n(L_{2,j-1}), U_3^n(l_3), Y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } l_3 \neq L_{3,j-1}, m \neq M_{j-1}\}, \\ \mathcal{E}_7 &= \{(X_1^n(m), U_2^n(l_2), U_3^n(l_3), Y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } l_2 \neq L_{2,j-1}, l_3 \neq L_{3,j-1}, m \neq M_{j-1}\}. \end{aligned}$$

Then by the union of events bound, the probability of decoding error is upper bounded as

$$\begin{aligned} \mathbb{P}(\hat{M}_{j-1} \neq M_{j-1}) &\leq \mathbb{P}(\mathcal{E}_1) + \mathbb{P}(\mathcal{E}_2) + \mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c) \\ &\quad + \mathbb{P}(\mathcal{E}_4) + \mathbb{P}(\mathcal{E}_5) + \mathbb{P}(\mathcal{E}_6) + \mathbb{P}(\mathcal{E}_7). \end{aligned}$$

By the covering lemma, $\mathbb{P}(\mathcal{E}_1)$ and $\mathbb{P}(\mathcal{E}_2)$ tend to zero as $n \rightarrow \infty$, if

$$\begin{aligned} R_2 &> I(U_2; Y_2) + \delta(\epsilon'), \\ R_3 &> I(U_3; Y_3) + \delta(\epsilon'), \end{aligned}$$

respectively. By the Markov lemma, $\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c)$ tends to zero as $n \rightarrow \infty$. By the symmetry of the random codebook generation and the random index assignment at the relays, it suffices to consider the conditional probabilities of the remaining error events conditioned on the event that

$$\mathcal{M} = \{M_{j-1} = 1, L_{2,j-1} = 1, L_{3,j-1} = 1\}. \quad (21)$$

Then, by the packing lemma, $\mathbb{P}(\mathcal{E}_4)$ tends to zero as $n \rightarrow \infty$ if

$$R < I(X_1; U_2, U_3, Y_4) - \delta(\epsilon).$$

Next we bound $\mathbb{P}(\mathcal{E}_5)$. Let $\tilde{U}^n = (U_2^n(1), Y_2^n(j-1), Y_3^n(j-1))$ and $\tilde{u}^n = (\tilde{u}_2^n, \tilde{y}_2^n, \tilde{y}_3^n)$ in short. Then, by the law of total probability, for $l_2 \neq 1, l_3 \neq 1$, and $m \neq 1$,

$$\begin{aligned} &\mathbb{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= \sum_{\tilde{u}^n} \mathbb{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(1) = u_3^n, \\ &\quad Y_4^n(j) = y_4^n, \tilde{U}^n = \tilde{u}^n | \mathcal{M}\} \\ &\stackrel{(a)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_2^n(l_2) = u_2^n | \mathcal{M}, X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, \\ &\quad \tilde{U}^n = \tilde{u}^n\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n\} \\ &\quad \times \mathbb{P}\{X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\stackrel{(b)}{=} \sum_{\tilde{u}^n} \mathbb{P}\{U_2^n(l_2) = u_2^n | L_{2,j-1} = 1, U_2^n(1), Y_2^n(j-1)\} \\ &\quad \times \mathbb{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n\} \\ &\quad \times \mathbb{P}\{X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\stackrel{(c)}{\leq} (1 + \epsilon) \left(\prod_{i=1}^n p_{U_2}(u_{2i}) \right) \\ &\quad \times \mathbb{P}\{X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \end{aligned}$$

for n sufficiently large. Here, (a) follows since given \mathcal{M} ,

$$\begin{aligned} U_2^n(l_2) &\rightarrow (U_2^n(L_{2,j-1}), U_3^n(L_{3,j-1}), Y_2^n(j-1), Y_3^n(j-1)) \\ &\rightarrow Y_4^n(j) \end{aligned}$$

form a Markov chain for all $l_2 \neq 1$, (b) follows from the independence of the sequences and the encoding procedure, and (c) follows by Lemma 1. Hence, for n sufficiently large,

$$\begin{aligned} \mathbb{P}(\mathcal{E}_5) &= \mathbb{P}\{(X_1^n(m), U_2^n(l_2), U_3^n(1), Y_4^n(j)) \in \mathcal{T}_{\epsilon}^{(n)} \\ &\quad \text{for some } l_2 \neq 1, m \neq 1 | \mathcal{M}\} \\ &\leq \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_{\epsilon}^{(n)}} \mathbb{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(1) = u_3^n, \\ &\quad Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\leq (1 + \epsilon) \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_{\epsilon}^{(n)}} p(u_2^n) \\ &\quad \times \mathbb{P}\{X_1^n(m) = x_1^n, U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= (1 + \epsilon) \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_{\epsilon}^{(n)}} p(u_2^n) \\ &\quad \times p(x_1^n) \mathbb{P}\{U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= (1 + \epsilon) 2^{n(R+R_2)} \\ &\quad \times \sum_{(u_2^n, y_4^n) \in \mathcal{T}_{\epsilon}^{(n)}} \mathbb{P}\{U_3^n(1) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\quad \times 2^{-n(I(X_1, U_2; U_3, Y_4) + I(X_1; U_2) - \delta(\epsilon))}, \end{aligned}$$

which implies that $\mathbb{P}(\mathcal{E}_5)$ tends to zero as $n \rightarrow \infty$ if

$$R + R_2 < I(X_1, U_2; U_3, Y_4) + I(X_1; U_2) - \delta(\epsilon).$$

We can bound $\mathbf{P}(\mathcal{E}_6)$ in a similar manner, provided that the subscripts 2 and 3 are interchanged in the upper bound for $\mathbf{P}(\mathcal{E}_5)$. Thus, $\mathbf{P}(\mathcal{E}_6)$ tends to zero as $n \rightarrow \infty$ if

$$R + R_3 < I(X_1, U_3; U_2, Y_4) + I(X_1; U_3) - \delta(\epsilon).$$

Next we bound $\mathbf{P}(\mathcal{E}_7)$. Let $\tilde{U}^n = (U_2^n(1), U_3^n(1), Y_2^n(j-1), Y_3^n(j-1))$ and $\tilde{u}^n = (\tilde{u}_2^n, \tilde{u}_3^n, \tilde{y}_2^n, \tilde{y}_3^n)$ in short. Then, by the law of total probability, for $l_2 \neq 1, l_3 \neq 1$, and $m \neq 1$,

$$\begin{aligned} & \mathbf{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(l_3) = u_3^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= \sum_{\tilde{u}^n} \mathbf{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(l_3) = u_3^n, \\ & \quad Y_4^n(j) = y_4^n, \tilde{U}^n = \tilde{u}^n | \mathcal{M}\} \\ &\stackrel{(a)}{=} \sum_{\tilde{u}^n} \mathbf{P}\{U_2^n(l_2) = u_2^n, U_3^n(l_3) = u_3^n | \mathcal{M}, X_1^n(m) = x_1^n, \\ & \quad \tilde{U}^n = \tilde{u}^n\} \\ & \quad \times \mathbf{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n\} \\ & \quad \times \mathbf{P}\{X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\stackrel{(b)}{=} \sum_{\tilde{u}^n} \mathbf{P}\{U_2^n(l_2) = u_2^n | L_{2,j-1} = 1, U_2^n(1), Y_2^n(j-1)\} \\ & \quad \times \mathbf{P}\{U_3^n(l_3) = u_3^n | L_{3,j-1} = 1, U_3^n(1), Y_3^n(j-1)\} \\ & \quad \times \mathbf{P}\{\tilde{U}^n = \tilde{u}^n | \mathcal{M}, X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n\} \\ & \quad \times \mathbf{P}\{X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\stackrel{(c)}{\leq} (1 + \epsilon) \left(\prod_{i=1}^n p_{U_2}(u_{2i}) p_{U_3}(u_{3i}) \right) \\ & \quad \times \mathbf{P}\{X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \end{aligned}$$

for n sufficiently large. Here, step (a) follows since given \mathcal{M} , $(U_2^n(l_2), U_2^n(l_3)) \rightarrow (U_2^n(L_{2,j-1}), U_3^n(L_{3,j-1}), Y_2^n(j-1), Y_3^n(j-1)) \rightarrow Y_4^n(j)$ form a Markov chain for all $l_2 \neq 1$ and $l_3 \neq 1$, (b) follows from the independence of the sequences and the encoding procedure, and (c) follows by applying Lemma 1 twice. Thus, for n sufficiently large,

$$\begin{aligned} \mathbf{P}(\mathcal{E}_7) &= \mathbf{P}\{(X_1^n(m), U_2^n(l_2), U_3^n(l_3), Y_4^n(j)) \in \mathcal{T}_\epsilon^{(n)} \\ & \quad \text{for some } l_2 \neq 1, l_3 \neq 1, m \neq 1 | \mathcal{M}\} \\ &\leq \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{l_3=2}^{2^{nR_3}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_\epsilon^{(n)}} \\ & \quad \mathbf{P}\{X_1^n(m) = x_1^n, U_2^n(l_2) = u_2^n, U_3^n(l_3) = u_3^n, \\ & \quad Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &\leq (1 + \epsilon) \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{l_3=2}^{2^{nR_3}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_\epsilon^{(n)}} p(u_2^n) \\ & \quad \times p(u_3^n) \cdot \mathbf{P}\{X_1^n(m) = x_1^n, Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= (1 + \epsilon) \sum_{m=2}^{2^{nR}} \sum_{l_2=2}^{2^{nR_2}} \sum_{l_3=2}^{2^{nR_3}} \sum_{(x_1^n, u_2^n, u_3^n, y_4^n) \in \mathcal{T}_\epsilon^{(n)}} p(u_2^n) \\ & \quad \times p(u_3^n) p(x_1^n) \cdot \mathbf{P}\{Y_4^n(j) = y_4^n | \mathcal{M}\} \\ &= (1 + \epsilon) 2^{n(R+R_2+R_3)} \sum_{y_4^n \in \mathcal{T}_\epsilon^{(n)}} \mathbf{P}\{Y_4^n(j) = y_4^n | \mathcal{M}\} \\ & \quad \times 2^{-n(I(X_1, U_2, U_3; Y_4) + I(X_1; U_2, U_3) + I(U_2; U_3) - \delta(\epsilon))}, \end{aligned}$$

which implies that $\mathbf{P}(\mathcal{E}_7)$ tends to zero as $n \rightarrow \infty$ if

$$R + R_2 + R_3 < I(X_1, U_2, U_3; Y_4) + I(X_1; U_2, U_3) + I(U_2; U_3) - \delta(\epsilon).$$

Finally, by eliminating R_2 and R_3 , the probability of error tends to zero as $n \rightarrow \infty$ if the conditions in Theorem 3 are satisfied.

ACKNOWLEDGMENTS

The authors would like to thank the Associate Editor and the anonymous reviewers for their insightful comments that helped improve the presentation of the paper.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, nos. 3–4, pp. 379–423, 623–656, 1948.
- [2] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, vol. 7, pt. 4, pp. 142–163.
- [3] R. M. Fano, *Transmission of Information*. Cambridge, MA, USA: MIT Press, 1961.
- [4] J. L. Massey, "Joint source and channel coding," in *Communication Systems and Random Process Theory* (NATO Advanced Studies Institutes Series), J. K. Skwirzynski, Ed. Alphen aan den Rijn, The Netherlands: Noordhoff, 1978, pp. 279–293.
- [5] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.
- [6] L. Song, R. W. Yeung, and N. Cai, "A separation theorem for single-source network coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1861–1871, May 2006.
- [7] A. Ramamoorthy, K. Jain, P. A. Chou, and M. Effros, "Separating distributed source coding from network coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2785–2795, Jun. 2006.
- [8] M. Agarwal and S. K. Mitter. (2010). "Communication to within a fidelity criterion over unknown networks by reduction to reliable communication problems over unknown networks." [Online]. Available: <http://arxiv.org/abs/1002.1300>
- [9] S. Jalali and M. Effros, "On the separation of lossy source-network coding and channel coding in wireline networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 500–504.
- [10] C. Tian, J. Chen, S. N. Diggavi, and S. Shamai (Shitz), "Optimality and approximate optimality of source-channel separation in networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 904–918, Feb. 2014.
- [11] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 648–657, Nov. 1980.
- [12] K. de Bruyn, V. V. Prelov, and E. van der Meulen, "Correlated sources over an asymmetric multiple access channel with one distortion criterion (corresp.)," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 716–718, Sep. 1987.
- [13] R. Rajesh, V. K. Varshneya, and V. Sharma, "Distributed joint source channel coding on a multiple access channel with side information," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 2707–2711.
- [14] S. H. Lim, P. Minero, and Y.-H. Kim, "Lossy communication of correlated sources over multiple access channels," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep./Oct. 2010, pp. 851–858.
- [15] A. Jain, D. Gündüz, S. R. Kulkarni, H. V. Poor, and S. Verdú, "Energy-distortion tradeoffs in Gaussian joint source-channel coding problems," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3153–3168, May 2012.
- [16] T. S. Han and M. M. H. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 641–650, Sep. 1987.
- [17] 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.
- [18] E. Tuncel, "Slepian–Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

- [19] K. Wei and G. Kramer, "Broadcast channel with degraded source random variables and receiver side information," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1711–1715.
- [20] G. Kramer and C. Nair, "Comments on 'Broadcast channels with arbitrarily correlated sources,'" in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun./Jul. 2009, pp. 2777–2779.
- [21] P. Minero and Y.-H. Kim, "Correlated sources over broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun./Jul. 2009, pp. 2780–2784.
- [22] G. Kramer, Y. Liang, and S. Shamai (Shitz), "Outer bounds on the admissible source region for broadcast channels with dependent sources," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2009, pp. 169–172.
- [23] R. Soundararajan and S. Vishwanath, "Hybrid coding for Gaussian broadcast channels with Gaussian sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun./Jul. 2009, pp. 2790–2794.
- [24] C. Tian, S. Diggavi, and S. Shamai (Shitz), "The achievable distortion region of sending a bivariate Gaussian source on the Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6419–6427, Oct. 2011.
- [25] J. Nayak, E. Tuncel, and D. Gündüz, "Wyner–Ziv coding over broadcast channels: Digital schemes," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1782–1799, Apr. 2010.
- [26] Y. Gao and E. Tuncel, "Wyner–Ziv coding over broadcast channels: Hybrid digital/analog schemes," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5660–5672, Sep. 2011.
- [27] W. Liu and B. Chen, "Communicating correlated sources over interference channels: The lossy case," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 345–349.
- [28] W. Liu and B. Chen, "Interference channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 8027–8037, Dec. 2011.
- [29] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3927–3944, Sep. 2009.
- [30] T. P. Coleman, E. Martinian, and E. Ordentlich, "Joint source–channel coding for transmitting correlated sources over broadcast networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3864–3868, Aug. 2009.
- [31] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Reliable joint source–channel cooperative transmission over relay networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2442–2458, Apr. 2013.
- [32] U. Mittal and N. Phamdo, "Hybrid digital–analog (HDA) joint source–channel codes for broadcasting and robust communications," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.
- [33] S. Shamai (Shitz), S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.
- [34] Y. Kochman and R. Zamir, "Joint Wyner–Ziv/dirty-paper coding by modulo–lattice modulation," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4878–4889, Nov. 2009.
- [35] M. P. Wilson, K. Narayanan, and G. Caire, "Joint source channel coding with side information using hybrid digital analog codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4922–4940, Oct. 2010.
- [36] Y. Gao and E. Tuncel, "New hybrid digital/analog schemes for transmission of a Gaussian source over a Gaussian channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6014–6019, Dec. 2010.
- [37] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian MAC," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2714–2752, Jun. 2010.
- [38] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian source over a Gaussian MAC with feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1852–1864, Apr. 2010.
- [39] S. Yao and M. Skoglund, "Hybrid digital–analog relaying for cooperative transmission over slow fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 944–951, Mar. 2009.
- [40] M. N. Khormuji and M. Skoglund, "Hybrid digital–analog noisy network coding," in *Proc. Int. Symp. Netw. Coding (NetCod)*, Jul. 2011, pp. 1–5.
- [41] Y. Kochman, A. Khina, U. Erez, and R. Zamir, "Rematch-and-forward: Joint source–channel coding for parallel relaying with spectral mismatch," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 605–622, Jan. 2014.
- [42] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, no. 4, pp. 289–293, 1958.
- [43] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [44] R. M. Gray and A. D. Wyner, "Source coding for a simple network," *Bell Syst. Tech. J.*, vol. 53, no. 9, pp. 1681–1721, Nov. 1974.
- [45] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.
- [46] M. H. Yassaee and M. R. Aref, "Slepian–Wolf coding over cooperative relay networks," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 3462–3482, Jun. 2011, doi: 10.1109/TIT.2011.2143990.
- [47] B. Schein and R. G. Gallager, "The Gaussian parallel relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 22.
- [48] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [49] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1668–1672.
- [50] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [51] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.
- [52] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. New York, NY, USA: Springer-Verlag, 1978, pp. 171–231.
- [53] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School Elect. Eng., Cornell Univ., Ithaca, NY, USA, 1978.
- [54] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems Control Inf. Theory*, vol. 2, no. 2, pp. 149–162, 1973.
- [55] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, no. 1, pp. 100–113, 1975.
- [56] P. Minero, S. H. Lim, and Y.-H. Kim, "Hybrid coding: A universal architecture for joint source–channel coding and network communication," in preparation.
- [57] A. B. Wagner, B. G. Kelly, and Y. Altuğ, "Distributed rate–distortion with common components," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4035–4057, Aug. 2011.
- [58] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, Sep. 1973.
- [59] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
- [60] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep. 1997.
- [61] V. Prabhakaran, D. N. C. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. Int. Symp. Inf. Theory*, Chicago, IL, USA, Jun./Jul. 2004, p. 117.
- [62] Y. Oohama, "Rate–distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.
- [63] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian 'sensor' network," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5247–5251, Nov. 2008.
- [64] M. R. Aref, "Information flow in relay networks," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, Oct. 1980.
- [65] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
- [66] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
- [67] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.
- [68] J. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [69] M. N. Khormuji and M. Skoglund, "On instantaneous relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3378–3394, Jul. 2010.
- [70] A. El Gamal, "On information flow in relay networks," in *Proc. IEEE Nat. Telecom Conf.*, vol. 2, Nov. 1981, pp. D4.1.1–D4.1.4.
- [71] S. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [72] P. Minero, S. H. Lim, and Y.-H. Kim, "Hybrid coding: A new paradigm for relay communication," in *Proc. 1st IEEE Global Conf. Signal Inf. Process.*, Dec. 2013, pp. 919–922.

Paolo Minero (S'05–M'11) received the Laurea degree (with highest honors) in electrical engineering from the Politecnico di Torino, Torino, Italy, in 2003, the M.S. degree in electrical engineering from the University of California at Berkeley in 2006, and the Ph.D. degree in electrical engineering from the University of California at San Diego in 2010. He is currently an Adjunct Assistant Professor in the Department of Electrical Engineering at the University of Notre Dame, where he was an Assistant Professor from 2011 to 2014. He is also currently a Staff Engineer in the Modem System Engineering group at Qualcomm Inc. His research interests are in communication systems theory and include information theory, wireless communication, and control over networks. Dr. Minero received the U.S. Vodafone Fellowship in 2004 and 2005, and the Shannon Memorial Fellowship in 2008.

Sung Hoon Lim (S'08–M'12) received the B.S. degree with honors in electrical and computer engineering from Korea University, Korea, in 2005 and the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 2007 and 2011, respectively. From March 2012 to May 2014, he was with Samsung Electronics. He is currently a postdoctoral associate in the School of Computer and Communication Sciences at Ecole polytechnique federale de Lausanne (EPFL), Lausanne, Switzerland. His research interests are in information theory, communication systems, and data compression.

Young-Han Kim (S'99–M'06–SM'12–F'15) received the B.S. degree with honors in electrical engineering from Seoul National University, Korea, in 1996 and the M.S. degrees in electrical engineering and statistics, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2001, 2006, and 2006, respectively. In July 2006, he joined the University of California, San Diego, where he is currently an Associate Professor of Electrical and Computer Engineering. His research interests are in statistical signal processing and information theory, with applications in communication, control, computation, networking, data compression, and learning.

Dr. Kim is a recipient of the 2008 NSF Faculty Early Career Development (CAREER) Award the 2009 US-Israel Binational Science Foundation Bergmann Memorial Award, and the 2012 IEEE Information Theory Paper Award. He served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and a Distinguished Lecturer for the IEEE Information Theory Society.