Robust Communication via Decentralized Processing With Unreliable Backhaul Links

Osvaldo Simeone, *Member, IEEE*, Oren Somekh, *Member, IEEE*, Elza Erkip, *Fellow, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—A source communicates with a remote destination via a number of distributed relays. Communication from source to relays takes place over a (discrete or Gaussian) broadcast channel, while the relays are connected to the receiver via orthogonal finite-capacity links. Unbeknownst to the source and relays, link failures may occur between any subset of relays and the destination in a nonergodic fashion. Upper and lower bounds are derived on average achievable rates with respect to the prior distribution of the link failures, assuming the relays to be oblivious to the source codebook. The lower bounds are obtained by proposing strategies that combine the broadcast coding approach, previously investigated for quasi-static fading channels, and different robust distributed compression techniques. Numerical results show that lower and upper bounds are quite close over most operating regimes, and provide insight into optimal transmission design choices for the scenario at hand. Extension to the case of nonoblivious relays is also discussed.**

***Index Terms*—Broadcast coding, distributed source coding, erasure channel, relay channel, robust channel coding.**

I. Introduction

WIRELESS or wired link failures are of a nonergodic nature whenever the delay tolerated by the application at hand is smaller than, or of the same order of magnitude of, the link outage duration. Such failures are often unpredictable to the transmitter, typically because of the absence of sufficiently fast feedback signaling, to simplify transmitter design, or simply because packet losses may be caused by remote events such as network congestion, see, e.g., [1]. In these situations, conventional channel coding is not effective in coping with link failures.

In the context of wireless channels, where outage is caused by poor fading conditions, a standard approach considers fixed-rate transmission, for a given signal-to-noise-ratio (SNR), and evaluates the best possible tradeoff between rate and reliability (outage) [2]. However, in a number of important applications, one may accept *variable-rate* data delivery, as in the case of video broadcasting: the receiver will simply experience variable reception quality according to the current channel state, and benefit from potentially good fading conditions [3]–[7]. As first proposed in [6], such variable-rate delivery can be achieved, without channel state information, by layering a number of transmission streams via superposition coding. This strategy is referred to as the *broadcast (BC) coding approach*. Layering can then be optimized in terms of *average achievable rate* with respect to a given prior distribution over the fading gains [6].

The issue of nonergodic link failures has also been widely studied in the context of wired networks, especially in recent years in the field of network coding, see, e.g., [9]. When the wired network is used for conveying information regarding correlated sources, the problem of transmission in the presence of (unpredictable) link failures is one of *robust* source coding, which has been studied in [10] and [11] for the case of a single encoder and in [12] and [13] for multiple distributed encoders. These works show that it is generally advantageous to make provision for the entire range of possible link conditions in order to fully exploit the available tradeoffs in the reconstruction quality at the receiver. This conclusion and approach are apparently synergic with the BC coding strategy of [6]. Such a synergy, also exploited in [3]–[7], and references therein in a different context, discussed later, motivates this paper.

We consider a scenario in which a single source communicates with a remote destination via a number of relays, also referred to as agents in related literature. Communication between source and agents is over a broadcast channel, either discrete or Gaussian, while the agents are connected to the destination via orthogonal limited-capacity channels. The scenario can be seen as a special case of a multirelay channel, without a direct link between source and destination, and with no multiaccess interference at the destination. In this sense, it is related to the "diamond network" of [14], to "Aref networks" [15], where the broadcast channel is deterministic, and to primitive relay channels [16], where one relay is available and there is a broadcast channel from source to relay and destination. In [17] the multirelay network described above was studied under the assumption that the relays are *oblivious* to the codebook used by the source; that is, processing at the relays cannot depend on the specific codebook selected by the source, as in, e.g., compress-and-forward

Manuscript received December 21, 2008; revised October 25, 2010; accepted January 12, 2011. Date of current version June 22, 2011. This work was supported in part by the U.S. National Science Foundation under Grants CCF-0914899, CNS-09-05398, CNS-06-26611, CCF-0635177, CNS-0905446, and IIP-1032035, in part by a Marie Curie Outgoing International Fellowship, in part by the NEWCOM++ network of excellence, and in part by the Israel Science Foundation. The material in this paper was presented in part at the IEEE Information Theory Workshop, Volos, Greece, June 2009.

O. Simeone is with the CWCSPR, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: osvaldo.simeone@njit.edu).

O. Somekh and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: orens@princeton.edu; poor@princeton.edu).

E. Erkip is with the Department of ECE, Polytechnic Institute of New York University, Brooklyn, NY 11201 USA (e-mail: elza@poly.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion, Haifa, 32000, Israel.

Communicated by H. Bölcskei, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2011.2145830

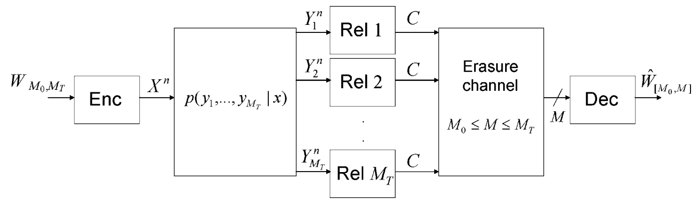

Fig. 1. Single transmitter (encoder) communicates to a remote receiver (decoder) via M_T relays connected to the destination through unreliable finite-capacity links. The number of functioning links M is unknown to source and relays (*uninformed source and relays*) and satisfies $M_0 \le M \le M_T$. Links are on or off for the entire duration of the codeword (nonergodic scenario). Agents may also be oblivious to the codebook used by the source (*oblivious agents*) as in [17].

or amplify-and-forward achievable strategies. This assumption is of particular relevance for nomadic applications, in which no signalling is in place to exchange information regarding modulation and coding used at the source, or in networks with inexpensive relays whose processing cannot adapt to the specific source operation.

In this paper, we consider the multirelay channel of Fig. 1. Unlike the works considered above, we assume that the links between the relays and the destination are unreliable, suffering from *nonergodic failures*, and the current state of the links is unknown to the encoders (source and relays). This assumption complicates significantly the problem and calls for the adoption of robust coding technique at both the source and the relays.

A related model with unreliable (nonergodic) connectivity was studied in [12] and [13] in the context of distributed source compression. In these works, a number of agents measure samples of a source process, independent and identically distributed, i.i.d., over time, via independent white Gaussian noise channels. A description of the measurements is provided via separate encoding by each agent to the destination over finite-capacity links, so as to enable the receiver to reconstruct an estimate of the source (the CEO problem [18]). Unbeknownst to the agents, the links to the destination may not be functioning, and robust distributed compression strategies must be devised to cope with the different possible connectivity conditions. Notice that, unlike the model of interest in the current work, the goal of [12] and [13] is to reconstruct a given fixed source and not to design the source coding strategy for reliable communications. Therefore, our model combines both issues related to robust source coding as in [12] and [13], but also of channel coding.

The basic idea behind our approach to the analysis of the system in Fig. 1 is to exploit the synergy between the BC coding approach of [6] at the source, which allows for variable-data delivery to the destination depending on the current connectivity conditions, and the robust distributed compression strategies of [12] and [13]. It is noted that a related idea was put forth in [3]–[5], in which the BC coding approach was combined with successive-description compression techniques for transmission of a Gaussian source over a slowly fading channel without channel state information.

The organization of the paper and its main contributions are as follows. We formalize the problem in Section II and derive upper bounds on the average achievable rate, in the sense of [6] and [7], for the system in Fig. 1 in Section III. Average achievable rates based on BC coding and robust compression are derived in Section IV, which are shown via numerical results in Section V to perform close to the derived outer bounds and to provide relevant gains with respect to conventional strategies. Finally, we study the case of nonoblivious agents in Section VI.

Notation: The notation $[a,b]$ with integers a,b represents the interval $[a,a+1,\ldots,b]$, with the convention that if $a>b$ then $[a,b]=\emptyset$. Similarly, the subscript notation $X_{[a,b]}$ denotes the vector $[X_a,\ldots,X_b]$ with the same convention that, if $a>b$, $X_{[a,b]}=\emptyset$. In general, lower-case letters represent instances of the random variables denoted by the corresponding upper-case letters. Moreover, using standard notation, we will sometimes use superscripts to denote index bounds in sequences as in $x^i=[x_1,\ldots,x_i]$. The use of the superscript will be made clear by the context. Probability distributions are identified by their arguments, e.g., $p_X(x)=\Pr[X=x]\triangleq p(x)$. We use standard definitions for information measures as defined in [19].

II. System Model

We consider the decentralized communication scenario of Fig. 1, in which a source communicates to a destination via M_T "agents" or relays, connected to the receiver via orthogonal finite-capacity (backhaul) links of capacity C. No direct connection from the source to the destination is available. The channel from source to relays is memoryless and either discrete or Gaussian. For the former case, the signal $Y_{i,j}\in\mathcal{Y}$ received by the agent $i\in[1,M_T]$ at time instant $j\in[1,n]$ is the output of a *symmetric* memoryless channel defined by the conditional distribution $p(y_1,\ldots,y_{M_T}|x)$, with input $x\in\mathcal{X}$ and block length n. Symmetric here means that the observations $Y_{i,j}$ for different i are statistically exchangeable, see, e.g., [10]–[12]. For the Gaussian case, we similarly have the input-output relationship

$$Y_{i,j}=X_j+Z_{i,j} \tag{1}$$

with X_j being the jth transmitted symbol and the noise $Z_{i,j}\sim\mathcal{N}(0,1)$ being i.i.d. over both i and j. We assume an average input power constraint of P: $1/n\sum_{j=1}^{n}x_j^2\le P$. In describing the model below, we will use the notation for the discrete model, but it is understood that the extension to the Gaussian model (1) is immediate.

To account for a nomadic scenario and/or to simplify the operations at the relays, we assume, as in [17], that the relays are not informed about the codebooks used by the transmitter (**oblivious agents**). As formalized in Section II-A, this condition can be modelled by assuming that the channel codebook is generated at the source based on a random key F that is not available at the relays, but is available at the destination. For reference, Section VI also considers the case of nonoblivious agents.

The model described above coincides with the one studied in [17] . Here, however, we are interested in investigating the scenario in which the backhaul links from the relays to the destination are affected by **nonergodic failures**. Specifically, following [12], we assume that only a number $M \leq M_T$ of links are functioning in a given coding block, while the remaining $M_T - M$ are *erased*, e.g., in outage, for the entire duration of the current transmission, i.e., this is a nonergodic scenario. As in [12], we assume that the number of functioning links M is always guaranteed to be larger than a minimum value $M_0 \leq M_T$ so that $M_0 \leq M \leq M_T$ with probability one. Moreover, we define the probability that $M = m$ as p_m and collect the probabilities p_m in a vector $\mathbf{p} = (p_{M_0}, \ldots, p_{M_T})$. Notice that, according to the discussion above, this probability is zero for $m < M_0$. We remark that, by the symmetry of $p(y_1, \ldots, y_M \,|\, x)$ (discrete model) and (1) (Gaussian model), the system configuration for a given M depends only on the number M of active links active and not on which links are active. Finally, in keeping with the models for distributed source coding of [12] and [13], we are interested in scenarios in which *no instantaneous information regarding the current state of the unreliable links, i.e., the value of* M, *is available a priori to the source and the agents* (**uninformed source and agents**). More precisely, the only information that is available at the source and the relays is the probability mass function $\mathbf{p}$, which represents the *a priori* state of knowledge of the source and agents on the state of the unreliable links. This assumption is appropriate in scenarios where feedback is not available or the outage events are difficult to predict, e.g., when the channel has a short coherence time.

We are interested in *average* achievable rates, where the average is taken with respect to the *a priori* connectivity probability vector $\mathbf{p}$. Specifically, we consider a *degraded message* structure in which the overall source message of rate T_{M_T} [bits/channel use] is split into submessages $(W_{M_0}, \ldots, W_{M_T}) \triangleq W_{[M_0, M_T]}$ of rates $R_{M_0}, \ldots, R_{M_T}$, respectively, i.e., $W_m \in [1, 2^{nR_m}]$. When $M = m$ links are active, with $m \in [M_0, M_T]$, the receiver decodes messages $W_{[M_0, m]}$ of total rate $T_m = \sum_{i=M_0}^{m} R_i$. Notice that the more links are active the more bits (and messages) are decoded. An *average rate*

$$R = \sum_{m=M_0}^{M_T} p_m T_m \tag{2}$$

is said to be achievable if all rates T_m are simultaneously achievable, in the sense that, when $M = m$ links are active, the given code guarantees decoding of $W_{[M_0, m]}$ with vanishingly small probability of error for all $m \in [M_0, M_T]$. We remark that, as in [6], the average rate (2) does not have the operational significance of an ergodic rate, since the channel is nonergodic. It is instead a measure of the rate that could be accrued with repeated and independent transmission blocks, or of the expected rate. The analysis presented below can also accommodate different criteria, such as the outage capacity, in which a zero rate is tolerated with a given probability. We refer also to [7] and [8] for further discussion on capacity definitions for "nonergodic" scenarios. The setting is formalized in the following.

A. Formal Setting

Denoting by n the size of each coding block, a code for the channel in Fig. 1 with *oblivious*[1] relays is defined by the following elements:

- The *encoder* performs a (stochastic) mapping $\phi_F^{(E)}$ from the messages $W_{[M_0, M_T]}$ to a codeword x^n, namely $x^n = \phi_F^{(E)}(W_{[M_0, M_T]})$ with
$$\phi_F^{(E)} : [1, 2^{nR_{M_0}}] \times \cdots \times [1, 2^{nR_{M_T}}] \rightarrow \mathcal{X}^n. \tag{3}$$
The codebook $\phi_F^{(E)}$ is indexed by a random key $F \in \mathcal{F} = [1, |\mathcal{X}|^{n2^{nT_{M_T}}}]$, which runs over all possible codebooks of size $2^{nT_{M_T}}$. The key $F \in \mathcal{F}$ is chosen randomly at the beginning of the communication session, and is revealed to the destination, but *not* to the relays. Notice that this model, in which coding is stochastic due to the random key F, is merely a way of formalizing the fact that the relays have no prior knowledge of the codebook, and it does not entail any real overhead. To elaborate, as in [17], the probability $P_F(f)$ of choosing a codebook $\phi_F^{(E)}$ indexed by key $F \in \mathcal{F}$ depends on a measure $p_{X^n}(x^n)$ over the space of the codewords as
$$P_F(f) = \prod_{w_{M_0}, \ldots, w_{M_T}} p_{X^n}\left(\phi_f^{(E)}(w_{M_0}, \ldots, w_{M_T})\right) \tag{4}$$
where the product is taken over the message sets $w_m \in [1, 2^{nR_m}]$ for $m \in [M_0, M_T]$. Moreover, the measure $p_{X^n}(x^n)$ is assumed to factor as $p_{X^n}(x^n) = \prod_{i=1}^{n} p_X(x_i)$ for a given single-letter probability distribution $p_X(\cdot)$. Reference [17] shows that, in the absence of information regarding F, i.e., at the relays, the signal transmitted by the source X^n is distributed i.i.d. according to a distribution $p_{X^n}(x^n) = \prod_{j=1}^{n} p_X(x_j)$ and, similarly, the received signals Y_i^n appear i.i.d., see [17, Lemma 1]. The source does not know the current number M of active links.
- Each ith *relay* ($i \in [1, M_T]$), unaware of the codebook F (oblivious relays), maps the received sequence $y_i \in \mathcal{Y}^n$ into an index $s_i \in [1, 2^{nC}]$ via a given mapping $s_i = \phi^{(i)}(y_i^n)$ as $\phi^{(i)} : \mathcal{Y}^n \rightarrow [1, 2^{nC}]$. Notice that this mapping does not depend on the current state of the link or the number M of active links. We remark that in Section VI, we consider a scenario in which the relays are aware of the source codebook.
- When $M = m$ *links are active*, the decoder decodes messages $W_{[M_0, m]} = (W_{M_0}, \ldots, W_m)$ based on its knowledge of the codebook key F and the received indices s_i over the m active links. These can be assumed by symmetry to be $s_1, \ldots, s_m$. The decoding function can be written as
$$\phi_F^{(D)} : [1, 2^{nC}]^m \rightarrow [1, 2^{nR_{M_0}}] \times \cdots \times [1, 2^{nR_m}]. \tag{5}$$

[1]The nonoblivious case will be treated in Section VI.

The probability of error when $M = m$ links are active, which is averaged over F, is defined as

$$P_{e,m}^{n} = \Pr\left[\phi_{F}^{(D)}\left(S_{[1,m]}\right) \neq W_{[M_0,m]}\right]. \tag{6}$$

An average rate R (2) is achievable if there exists a sequence of codes such that all rates $T_m = \sum_{j=M_0}^{m} R_j$ for $m \in [M_0, M_T]$ are achievable, i.e., $\max_m P_{e,m}^{n} \to 0$ as $n \to \infty$. The *average capacity* $\mathrm{C}_{\mathrm{avg}}$ is the supremum of all average achievable rates (2).[2]

III. Reference Results

In this section, we start the study of the system presented above by deriving an upper bound on the capacity $\mathrm{C}_{\mathrm{avg}}$. It is emphasized that the upper bound is valid under the given assumption of oblivious relays. Moreover, it is noted that, as in [17], for the Gaussian model, *we restrict the input distribution to be Gaussian* with no claim of optimality. We refer to [17] for further discussion on the suboptimality of the Gaussian distribution. The upper bound below also motivates the use of a BC coding approach at the source and of *compress-and-forward* (CF) at the relays, as exploited in the transmission strategies considered in the rest of the paper, see Remark 3.1.

Proposition 3.1: (Cooperative Relays): The following is an upper bound on the average capacity $\mathrm{C}_{\mathrm{avg}}$ for the discrete model:

$$\mathrm{C}_{\mathrm{avg}} \leq \max \sum_{m=M_0}^{M_T} p_m \left(\sum_{j=M_0}^{m} R_j\right) \tag{7}$$

where the rates

$$R_{M_0} = I\left(U_{M_0}; V_{M_0}\right) \tag{8a}$$
$$R_m = I(U_m; V_m \,|\, U_{m-1}) \text{ for } m \in [M_0, M_T - 1] \tag{8b}$$
$$R_{M_T} = I\left(X; V_{M_T} \,|\, U_{M_T-1}\right) \tag{8c}$$

are calculated with respect to a joint distribution

$$p\left(u_{[M_0,M_T-1]}, x, y_{[1,M_T]}, v_{[M_0,M_T]}\right) = p\left(u_{[M_0,M_T-1]}, x\right) p\left(y_{[1,M_T]} \,|\, x\right) p\left(v_{[M_0,M_T]} \,|\, y_{[1,M_T]}\right) \tag{9}$$

and the maximization is taken with respect to the marginals $p(u_{[M_0,M_T-1]}, x)$ and $p(v_{[M_0,M_T]} \,|\, y_{[1,M_T]})$ that factor as

$$p\left(u_{[M_0,M_T-1]}, x\right) = \prod_{m=M_0}^{M_T-1} p(u_m \,|\, u_{m-1}) p\left(x \,|\, u_{M_T-1}\right) \tag{10a}$$

$$p\left(v_{[M_0,M_T]} \,|\, y_{[1,M_T]}\right) = \prod_{m=M_0}^{M_T} p\left(v_m \,|\, y_{[1,m]}\right)$$

and satisfy

$$mC \geq I\left(V_m; Y_{[1,m]}\right). \tag{11}$$

Moreover, for the Gaussian model, (7) is an upper bound, under the constraint that the input distribution is Gaussian, with

$$R_m = \frac{1}{2}\log_2\left(1 + \frac{m\beta_m P}{1 + m\sigma_m^2 + mP\sum_{k=m+1}^{M_T}\beta_k}\right) \tag{12}$$

for $m \in [M_0, M_T]$, where the maximization is taken with respect to parameters $\beta_{M_0}, \ldots, \beta_{M_T} \geq 0$ with $\beta_{M_0} + \cdots + \beta_{M_T} = 1$ and $\sigma_m^2 = (1/m + P)/(2^{2mC} - 1)$.

Remark 3.1: As discussed in the sketch of the proof below, the upper bounds of Proposition 3.1 are obtained by assuming that all of the M relays that are connected to the corresponding active links can fully cooperate in processing their received signals. Notice that this implies that they are also informed of which links are active. The upper bounds can then be interpreted as stating that, under this assumption, the best way to operate at the source is to use a standard BC code characterized by auxiliary random variables U_m, $m \in [M_0, M_T - 1]$, for the discrete case or powers $\beta_m P (m \in [M_0, M_T])$ for the Gaussian case. Such variables or powers correspond to the transmission of message W_m to be decoded at the receiver when $M = m$. Notice that the variables U_m satisfy the Markov chain condition (10a), or equivalently $U_1 - U_2 - \cdots - U_{M_T-1} - X$ as for a regular degraded broadcast channel [24]. Moreover, the result in Proposition 3.1 also proves that fully cooperative relays can employ CF techniques without loss of optimality to communicate to the receiver: The auxiliary variables V_m account for the quantization codebook used by the $M = m$ cooperating relays when the active number of links is $M = m$ and parameter σ_m^2 is the corresponding compression noise power for the Gaussian case. In fact, from standard rate-distortion considerations, (11) is easily interpreted in this sense as necessary and sufficient to guarantee successful compression for all m (see the proof below). Notice that the optimality of CF in this context is a consequence of the obliviousness assumption, as detailed in the proof in Appendix A. Finally, we remark that the optimization problem in (7) for the Gaussian case (12) corresponds to the maximization of the weighted sum-rate of a broadcast channel, which can be solved using standard techniques [21]. We will use this intuition regarding the appropriateness of BC coding and CF when designing transmission strategies in the next section. We also notice that a similar conclusion regarding optimality of BC coding for a different setting was presented in [7].

Proof: (Sketch): Here we provide a simple proof for the bound (7), (12) for the Gaussian channel, as the derivation is more direct. The discrete model is discussed later and proved in Appendix A. Assume that the relays are perfectly cooperating so that, when $M = m$ links are active, they can be seen as

[2]The average capacity can be seen as identifying the hyperplane tangent to the region of all achievable rates $T_{M_0}, \ldots, T_M$ in the direction specified by vector $\mathbf{p}$.

a unique compound agent with m measurements $Y^n_{[1,m]}$, since the signals are statistically equivalent, there is no loss in generality in this choice of Y^n_j. It is easy to see that the compound agent can be equivalently considered as having scalar measurements: $Y_i^{(m)} = 1/m\sum_{j=1}^{m} Y_{j,i}$, because there is no performance loss in projecting the received signal over the signal space, since the noise in (1) is uncorrelated over the agents. As aforementioned, we limit the analysis, as in [17], to random coding with Gaussian inputs at the source. From [17], it is known that the optimal operation at the compound agent, which is clearly aware of the capacity mC toward the destination, is to quantize to a rate mC bits/source symbol the received signal via a Gaussian test channel $V_m = Y^{(m)} + Q_m$ with $Q_m \sim \mathcal{N}(0, \sigma_m^2)$ independent of $Y^{(m)}$. From standard arguments in rate-distortion theory, in order to have vanishing probability of error in the quantization process, as the block size n increases, we can set $mC = I(V_m; Y^{(m)})$ [cf. (11)], thus obtaining $\sigma_m^2 = (1/m + P)/(2^{2mC} - 1)$. As a result, since the source is not informed about the current value of M, the equivalent channel can be seen as a *degraded Gaussian broadcast channel*, in which the $M_T - M_0 + 1$ destinations observe received signals V_m with equivalent noise variances $1/m + \sigma_m^2$ for $m \in [M_0, M_T]$. Notice that such variances are clearly decreasing with m. Recalling the capacity region for the Gaussian broadcast channel, bound (12) then easily follows. As a final remark, it is noted that (7) and (12) for the Gaussian channel can also be obtained from the corresponding discrete result of Proposition 3.1, proved in Appendix A, by setting auxiliary variables $U'_i \sim \mathcal{N}(0, \beta_i P)$ independent for $i \in [M_0, M_T]$, $U_m = \sum_{i=M_0}^{m} U'_i$ and $X = \sum_{i=M_0}^{M_T} U'_i$, and V_m as discussed earlier.

IV. Achievable Rates

In the following, motivated by the upper bound of Proposition 3.1 , we propose achievable schemes based on the BC coding strategy of [6] and CF at the relays. In [6], a BC strategy was proposed to deal with uncertain fading conditions. The basic idea is that of treating all the possible channel fading states that might occur as distinct users, thus effectively converting the fading channel into a degraded broadcast channel. In this paper, the same principle is leveraged to operate over the channel at hand, which presents *unknown connectivity conditions* from the relays to the receiver. Specifically, the source transmits a superposition of $M_T - M_0 + 1$ codewords of rates R_m for $m \in [M_0, M_T]$. When $M = m$, the receiver decodes $W_{[M_0,m]}$. As far as the operation at the relays is concerned, due to the fact that codebook information, i.e., the key F of Section II-A, is not available at the relays, here we will assume that CF relaying is implemented. Notice that this coincides exactly with the strategy that was proved to be optimal for the setting of Proposition 3.1, see Remark 3.1. The different techniques proposed in the following differ in the way the CF strategy is implemented in terms of compression at the agents and decompression/decoding at the receiver, and entail increasing levels of complexity.

A. Broadcast Coding and Single-Description Compression (BC-SD)

In this section, we consider a transmission strategy based on BC coding and single-description (SD) compression at the relays. In other words, each relay sends over the backhaul link a single index (description), which is a function of the received signal. Moreover, we consider first separate decompression/decoding at the decoder, and then a potentially more effective, but more complex, joint decompression/decoding approach. A performance comparison that shows the performance-complexity tradeoff of these schemes is provided via numerical results in Section V for the Gaussian model.

1) Separate Decompression/Decoding (BC-SD-S): Here we propose a strategy based on separate (S) decompression/decoding. The compression/decompression scheme is inspired by the technique used in [12], see also [10], for robust distributed source coding in a CEO problem. The technique works by performing random binning at the agents, as is standard in distributed compression, see, e.g., [20]. Moreover, the binning rate is selected so that the receiver can recover with high probability the compressed signals on the M active links irrespective of the realized value of M as long as $M \geq M_0$, as guaranteed by assumption. In other words, design of the compression scheme targets the *worst-case scenario* of $M = M_0$. Notice that, should more than M_0 links be active ($M > M_0$), the corresponding compressed signals would also be recoverable at the receiver, since, by design of the binning rate, any subset of M_0 descriptions can be decompressed [12]. After decompression is performed, the receiver uses all the M signals obtained from the relays to decode the codewords up to the Mth layer, that is, the layers with rates R_m with $M_0 \leq m \leq M$.

Proposition 4.1: (BC-SD-S): The average rate (2) is achievable for the discrete model with

$$R_{M_0} \leq I\left(U_{M_0}; V_{[1,M_0]}\right) \tag{13a}$$

$$R_m \leq I\left(U_m; V_{[1,m]} \,|\, U_{m-1}\right) \text{ for } m \in [M_0+1, M_T-1] \tag{13b}$$

$$R_M \leq I\left(X; V_{[1,M_T]} | U_{M_T-1}\right). \tag{13c}$$

where the variables at hand satisfy the joint distribution

$$p(u_{[M_0,M_T-1]}, x, v_{[1,M_T]}, y_{[1,M_T]}) = \prod_{m=M_0}^{M_T-1} p(u_m \,|\, u_{m-1}) p(x \,|\, u_{M_T-1}) p\left(y_{[1,M_T]} \,|\, x\right) \prod_{i=1}^{M_T} p(v_i | y_i) \tag{14}$$

with $p(v_i \,|\, y_i)$ being the same for every $i \in [1, M_T]$, and the condition

$$C \geq \frac{1}{M_0}\left[H\left(V_{[1,M_0]}\right) - M_0 H(V_i \,|\, Y_i)\right]. \tag{15}$$

Moreover, for the Gaussian model, the average rate (2) is achievable with

$$R_m \leq \frac{1}{2}\log_2\left(1+\frac{m\beta_m P}{1+\sigma^2+mP\sum_{k=m+1}^{M_T}\beta_k}\right) \quad (16)$$

and σ^2 satisfying

$$C \geq \frac{1}{2}\log_2\left[\left(1+\frac{M_0 P}{1+\sigma^2}\right)^{\frac{1}{M_0}}\left(1+\frac{1}{\sigma^2}\right)\right] \quad (17)$$

for any power allocation $\beta_{M_0},\ldots,\beta_{M_T} \geq 0$ with $\beta_{M_0}+\cdots+\beta_{M_T}=1$.

Proof: See Appendix B. ■

Remark 4.1: Similar to the discussion around Proposition 3.1, the auxiliary random variable U_m for the discrete case and power $\beta_m P$ for the Gaussian case, $m \in [M_0, M_T-1]$, represents the codebook used for the transmission of the mth layer to be decoded at the receiver when $M=m$. Moreover, the variable V_i represents the compression codebook v_i^n used at each agent i. Notice that by symmetry the same distribution $p(v_i\,|\,y_i)$ is selected for all $i \in [1, M_T]$. Conditions (15) for the discrete case and (17) for the Gaussian case are shown in [12] to guarantee that the decoder is able to decompress the signals corresponding to any set of M_0 agents, see also proof in Appendix B. We finally notice that the only difference between the achievable rate of Proposition 4.1 obtained with BC-SD-S and the upper bound of Proposition 3.1 is related to the variables V_m used for compression, and in the Gaussian case to the power of the equivalent compression noise [compare (16) with (12)]. Specifically, for the upper bound of Proposition 3.1, variables V_m are selected jointly by the m cooperative relays when the number of active links is $M=m$, whereas each V_i in the achievable rate of Proposition 4.1 is selected independently by each ith relay, irrespective of the value of M.

Remark 4.2: For $M_T = M_0$ (fully reliable links), the achievable rate of Proposition 4.1 coincides with the one presented in [17, Theorem 1].

Remark 4.3: With separate decompression/decoding, one could in principle target successful decompression for a value of $\tilde{M}$ larger than M_0. While this would in general imply $R_{M_0}=\cdots=R_{\tilde{M}-1}=0$, the average rate (2) could be better in some cases. A simple example of this is the case in which the number of guaranteed links is $M_0=0$. Under this assumption, BC-SD-S, in which the binning rate is designed for the worst-case scenario ($M=M_0=0$), clearly cannot achieve a nonzero average rate. In contrast, one could design the binning rate for a $\tilde{M}>0$ and be able to achieve a nonzero average rate.

2) Joint Decompression/Decoding (BC-SD-J): Here we look at a potentially more efficient, but also more complex, implementation of a system working with BC coding and SD compression. Specifically, rather than performing separate decompression and decoding as in the scheme considered above, here, inspired by [17], we look at joint (J) decompression/decoding for each layer of the broadcast codebook to be decoded. It is noted that, since no separate decompression is performed, with the joint approach, there is no need to select the compression rate so that decompression is always possible with vanishing error whenever $M \geq M_0$, or more generally $M \geq \tilde{M}$, see Remark 4.3. This is because here we can allow for errors in decompression as long as decoding is correct, see Remark 4.4. The performance gains of this scheme will be assessed in Section V.

Proposition 4.2: (BC-SD-J): The average rate (2) is achievable for the discrete model with (18)–(19) as shown at the bottom of the page, with variables factorizing as in (14), with equal $p(v_i|y_i)$ for every i. Moreover, for the Gaussian model, the average rate (2) is achievable with (20)-(21), as shown at the bottom of the next page, for any $\sigma^2>0$.

Proof: See Appendix C. ■

Remark 4.4: The strategy (BC-SD-J) that achieves the rate defined above is characterized by the same BC coding scheme considered for Proposition 4.1. The agents also operate via CF in the same way, by producing an SD obtained by compressing the received signals into codewords v_i^n, and then binning the corresponding codebooks to C bits/symbol. The only difference is that here the decoder performs decoding of the messages $W_{[M_0,m]}$ *jointly* with the decompression of the codewords v_i^n from the indices received over the m active links. This is different from BC-SD-S, in which the decoder first recovers the compression-codebook codewords v_i^n and then, *if decompression is successful*, performs decoding (Proposition 4.1). In other words, with the separate strategy BC-SD-S, an error is declared whenever the decoder fails at the first step (decompression of the codewords v_i^n) or, if the first step is successful, it fails at the second (decoding). In contrast, with the joint approach BC-SD-J, an error is declared *only when decoding fails*, irrespective of whether some of the compression codewords

$$R_m \leq \min_{k\in[0,m]} k(C - I_m(k)) + I\left(U_m; V_{[k+1,m]}|U_{m-1}\right) \text{ for } m \in [M_0, M_T-1] \quad (18a)$$

$$R_{M_T} \leq \min_{k\in[0,M_T]} k(C - I(V_i;Y_i\,|\,X)) + I\left(X; V_{[k+1,M_T]}|U_{M_T-1}\right) \quad (18b)$$

and

$$I_m(k) = \frac{1}{k}H\left(V_{[1,k]}|U_m, V_{[k+1,m]}\right) - H(V_i\,|\,Y_i) \quad (19)$$

might have been erroneously decompressed. It follows that, in BC-SD-J, there is no need to force the binning rate to guarantee error-free decompression for all $M \geq M_0$, or more generally $M \geq \tilde{M}$, see Remark 4.3, as in BC-SD-S, that is, to impose the constraint (15) on the quantization defined by $p(v_i|y_i)$. Because of the larger set of compression strategies $p(v_i|y_i)$ allowed by BC-SD-J, it follows that the latter scheme generally achieves larger achievable rate than BC-SD-S of Proposition 4.1.[3]

Remark 4.5: The rates R_m in Proposition 4.2 are defined in terms of a minimization over a parameter $k \in [0, m]$. In other words, rates R_m, must be smaller than or equal to the right-hand side of the inequalities at hand for every k. As detailed in Appendix C, the parameter k runs over all the possible error events at the decoder. Specifically, in the kth error event for the mth layer (R_m), the receiver decodes incorrectly the message W_m and k compression codewords v_i^n, where we can choose $i \in [1, k]$ without loss of generality. The rate expressions (18a) can then be interpreted accordingly by noticing that: (*i*) The term $I(U_m; V_{[k+1,m]} \,|\, U_{m-1})$ is the mutual information between the input U_m for the mth layer and the $(m-k)$ correctly decoded compression codewords represented by $V_{[k+1,m]}$; (*ii*) The second term $k(C - I_m(k))$ accounts for the extra capacity available on the k links over which decompression is not successful, when one removes the rate $kI_m(k)$ wasted to convey either channel noise or other information that is nuisance for the mth layer, i.e., the layers $m+1, \ldots, M_T$) (see also [17]).

Remark 4.6: If $M_0 = M_T$ (fully reliable links), the results of Proposition 4.2 reduce to the achievable rate derived in [17 , Corollary 1] for the discrete case and the capacity result of [17, Theorem 5] for the Gaussian case, assuming Gaussian inputs.[4]

B. Broadcast Coding and Multidescription Robust Compression (BC-MD)

In this section, we propose to couple the BC coding approach considered throughout the paper with multidescription (MD), rather than SD, compression at the agents. The idea follows the work in [13], which focused on the CEO problem, see also [11]. Accordingly, each relay shares the nC bits it can convey to the destination between multiple descriptions of the received signal to the decoder. The basic idea is that different descriptions are designed to be recoverable only if certain connectivity conditions are met, that is, if the number of functioning links M is sufficiently large. This adds flexibility and robustness to the compression strategy.

To simplify the presentation, here we focus on the two-agent case $(M_T = 2)$. Dealing with the more general setup requires a somewhat more cumbersome notation, but is conceptually a straightforward extension (see [11] for related discussion). Moreover, without loss of generality, we assume $M_0 = 0$ or $M_0 = 1$, since with $M_0 = M_T = 2$ the system coincides with the system with fully reliable links studied in [17]. The two agents send two descriptions: a basic one to be used at the receiver in case the number of active links turns out to be $M = M_0 = 1$ and a “refined” one that will be used only if $M = M_T = 2$. It is also noted that for the scheme at hand the only difference between the cases $M_0 = 0$ and $M_0 = 1$ is in the prior $\mathbf{p} = (p_0, p_1, p_2)$, where in the former case, unlike the latter, we have $p_0 > 0$.

Proposition 4.3: (BC-MD): For $M_T = 2$, $M_0 = 0$ or $M_0 = 1$, the average rate (2) is achievable for the discrete model for

$$R_1 \leq I(U; V_{1i}) \tag{22a}$$

$$R_2 \leq I(X; V_{11}, V_{12}, V_{21}, V_{22} \,|\, U) \tag{22b}$$

with joint distribution

$$p(u, x, v_{11}, v_{12}, v_{21}, v_{22}, y_1, y_2) = p(u, x) p(y_1, y_2 \,|\, x) \prod_{i=1}^{2} p(v_{1i}, v_{2i} | y_1) \tag{23}$$

where $p(v_{1i}, v_{2i}|y_1)$ is the same for $i = 1, 2$, satisfying the constraint

$$C \geq I(V_{1i}; Y_1) + \frac{1}{2} I(V_{21}, V_{22}; Y_1, Y_2 \,|\, V_{11}, V_{12}). \tag{24}$$

Moreover, for $M_T = 2$, $M_0 = 0$ or $M_0 = 1$ and the Gaussian model, the average rate (2) is achievable for

$$R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{\beta P}{1 + (1-\beta)P + \sigma_1^2 + \sigma_2^2} \right) \tag{25a}$$

$$R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{2(1-\beta)P}{1 + \sigma_2^2} \right) \tag{25b}$$

[3]More rigorously, this fact can be seen as follows. Choose a distribution $p(v_i|y_i)$ that satisfies (15), as required by BC-SD-S. This guarantees that decompression is always successful for all $M \geq M_0$ (a similar argument could be used for the case $M \geq \tilde{M}$ of Remark 4.3). It can be now seen with some algebra that, with this choice, the rate region (13) of BC-SD-S coincides with the rate region (18) of BC-SD-J (that is, the minimum in (18) is obtained for $k = 0$). However, BC-SD-J allows for more general distributions $p(v_i|y_i)$ than the ones satisfying (15) and thus it achieves potentially larger rates.

[4]The result in [17] is recovered by setting $r = 1/2 \log(1 + 1/\sigma^2)$.

$$R_m \leq \min_{k \in [0,m]} \left(k(C - I_m(k)) + \frac{1}{2} \log_2 \left(1 + \frac{(m-k)\beta_m P}{1 + \sigma^2 + (m-k) P \sum_{i=m+1}^{M_T} \beta_i} \right) \right) \tag{20}$$

and

$$I_m(k) = \frac{1}{2} \log_2 \left(1 + \frac{1}{\sigma^2} \right) + \frac{1}{2k} \log_2 \left(1 + \frac{k \left(\sum_{i=m+1}^{M_T} \beta_i P \right)^2}{\left(\sum_{i=m+1}^{M_T} \beta_i P \right)(1 + \sigma^2) - (k-1) \left(\sum_{i=m+1}^{M_T} \beta_i P \right)^2} \right) \tag{21}$$

with any power allocation $0 \leq \beta \leq 1$, and any σ_1^2 and σ_2^2 such that [see (26) at the bottom of the page].

Remark 4.7: In the MD scheme achieving the rate above, each transmitter divides its capacity C into two parts, say with a fraction $0 \leq \lambda \leq 1$ devoted to the first $(m=1)$ and $(1-\lambda)$ to the second $(m=2)$ description. Auxiliary variables V_{mi} in (22) represent the quantization codebooks corresponding to the mth description $(m=1,2)$ of the ith terminal $(i=1,2)$. As explained above, the binning rate of the mth description is selected so that the description is recoverable at the destination whenever $M=m$. To ensure this, it is sufficient to impose the condition $\lambda C \geq I(V_{1i};Y_1)$ for $m=1$ from standard rate-distortion theoretic arguments, and $2(1-\lambda)C \geq I(V_{21},V_{22};Y_1,Y_2 \mid V_{11},V_{12})$ for $m=2$, from distributed lossy distortion theory, see, e.g., [13]. Notice that the latter inequality exploits the fact that the first descriptions V_{11} and V_{12} have been correctly decompressed at the decoder when $M=2$, and thus provide side information (see also [25]). In the Gaussian model, variances σ_1^2 and σ_2^2 in (25)–(26) account for the compression noises for the first and second description, respectively, and (26) corresponds to (24). The auxiliary random variable U in the discrete model and powers $(\beta P, (1-\beta)P)$ represent, as in the rest of the paper, the BC code.

Remark 4.8: Following the discussion in the previous section, one could employ joint decompression/decoding for the second description, rather than separate processing as done in Proposition 4.3. We do not further pursue this option here, but the result can be derived similarly to the proof of Proposition 4.2.

Remark 4.9: Setting V_{21} and V_{22} to be constant for the discrete model or letting $\sigma_2^2 \to \infty$ for the Gaussian model, Proposition 4.3 reduces to Proposition 4.1 for $M_T=2$, $M_0=0$ or 1.

Proof: The proof for the discrete model follows easily from the discussion above and the capacity of a degraded Gaussian channel, see the proof of Proposition 4.1 for similar arguments. For the Gaussian model, the rate bounds (25) are obtained from (22), and (26) from (24), by setting $V_{2i}=Y_i+Q_{2i}$ and $V_{1i}=V_{2i}+Q_{1i}$ with $Q_{1i}\sim\mathcal{N}(0,\sigma_1^2)$ and $Q_{2i}\sim\mathcal{N}(0,\sigma_2^2)$ independent for $i=1,2$.[5] Moreover, we set $X=U+U'$ with $U\sim\mathcal{N}(0,\beta P)$ and $U'\sim\mathcal{N}(0,(1-\beta)P)$ independent. The proof is then concluded with some algebra. ■

V. Numerical Results

In this section, we provide some numerical examples to illustrate the performance of the proposed BC-based strategies for the system in Fig. 1. Consider a two-agent system $(M_T=2)$ with $M_0=1$ guaranteed functioning links. We compare the performance of the schemes described above, with single description (SD) and either separate (S) or joint (J) decompression/decoding, or multidescription (MD) compression. As a reference, we consider the upper bound (12) corresponding to cooperative relays, labelled as "cooperative." To assess the impact of non-ergodic link outage, we also show the performance of a system in which the link outages occur in an ergodic fashion so that the agents effectively see a link capacity equal to the average $E[C]=(1-p_1/2)C$, labelled "ergodic." This rate clearly sets another upper bound on the average capacity, and can be found from [17] to be

$$\mathrm{C}_{\mathrm{avg}} \leq \frac{1}{2} \log_2\left(1+2P\left(1-\frac{\sqrt{P^2+2^{4E[C]}(1+2P)}-P}{2^{4E[C]}}\right)\right). \tag{27}$$

Finally, the rate of a baseline single-layer (SL), or nonbroadcast, transmission in which the source sends only one information layer to be decoded in the worst case scenario $M_0=1$ and the relays perform SD compression, is shown for reference. The rate of this SL-SD scheme is easily seen to be

$$R_{SL-SD}=\frac{1}{2}\log_2\left(1+\frac{P}{1+\sigma^2}\right) \tag{28}$$

with $\sigma^2=(1+P)/(2^{2C}-1)$. It is recalled that the upper bounds considered here are valid only when the restriction on Gaussian inputs is taken into account: Better performance could be generally achieved by considering more general distributions as discussed in [17].

Fig. 2 shows the average rates of the proposed schemes for $P=15$ dB and $C=0.5$ versus the probability $p_2=1-p_1$ of having $M=2$ active links (rather than the minimum guaranteed $M_0=1$). The rates are optimized numerically over the parameters at hand, i.e., the compression noise variances σ_i^2 and power allocation β. It can be seen that the BC coding strategy provides relevant advantages over SL as long as the probability p_2 is sufficiently large, since it offers the possibility of exploiting better connectivity conditions when they arise. Moreover, MD compression clearly outperforms SD-based approaches for all values of p_2 for which BC coding is advantageous, due to the added flexibility in allocating part of the backhaul link rate for the case of full connectivity $(M=M_T)$. In particular, BC-MD performs very close to the upper bound of cooperative relays and for $p_2=1$ achieves the capacity for $M_0=M_T=2$ of [17], that is, (27) with $p_1=0$. It is also noted that joint decompression/decoding (BC-SD-J) enables slightly better rates than separate decompression/decoding for p_2 sufficiently large.

The achievable rates and bounds are shown in Fig. 3 versus the power P for $C=0.5$ and $p_1=0.1$. MD compression is again seen to have almost optimal performance, with a rate gain over SD that increases with P. For P sufficiently large $(P\to\infty)$, the capacity with ergodic failures (27) tends to $2(1-p_1/2)C$, and the achievable rate (28) with SL-SD to C, whereas

[5] This choice is referred to as the "joint decoder first" approach in [26],

$$C \geq \frac{1}{2}\log\left(1+\frac{P+1}{\sigma_1^2+\sigma_2^2}\right)+\frac{1}{4}\log\left(\frac{\left(\sigma_1^2+\sigma_2^2\right)^2\left(2P+\sigma_2^2+1\right)\left(\sigma_2^2+1\right)}{\left(2P+\sigma_1^2+\sigma_2^2+1\right)\left(\sigma_1^2+\sigma_2^2+1\right)\sigma_2^4}\right) \tag{26}$$

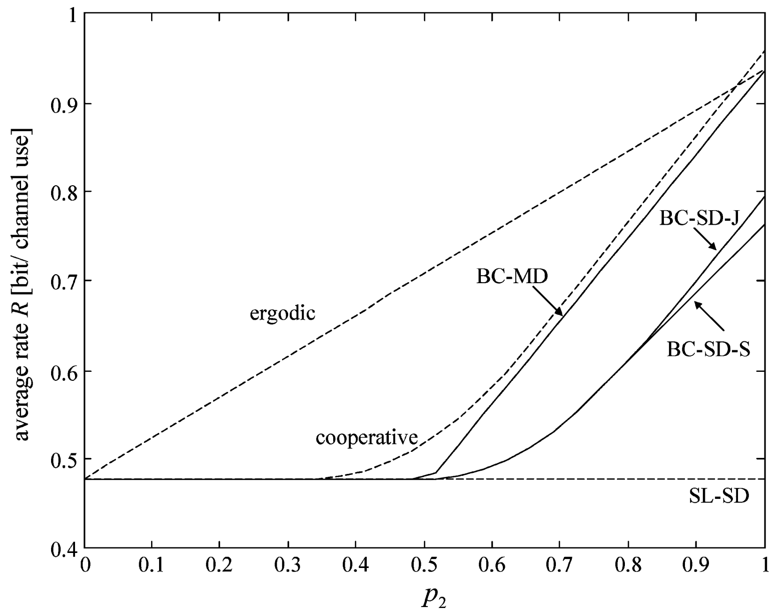

Fig. 2. Average achievable rates (2) for the proposed BC-based schemes with single description (SD) and either separate (S) or joint (J) decompression/ decoding, or multidescription (MD) compression, versus the probability $p_2 = 1 - p_1$ of having $M = 2$ active links. For reference, the upper bound (12) achievable with cooperative relay, the upper bound (27) corresponding to ergodic link failures and the rate (28) of single-layer (SL), or nonbroadcast, transmission with SD compression are also shown ($P = 15$ dB and $C = 0.5$).

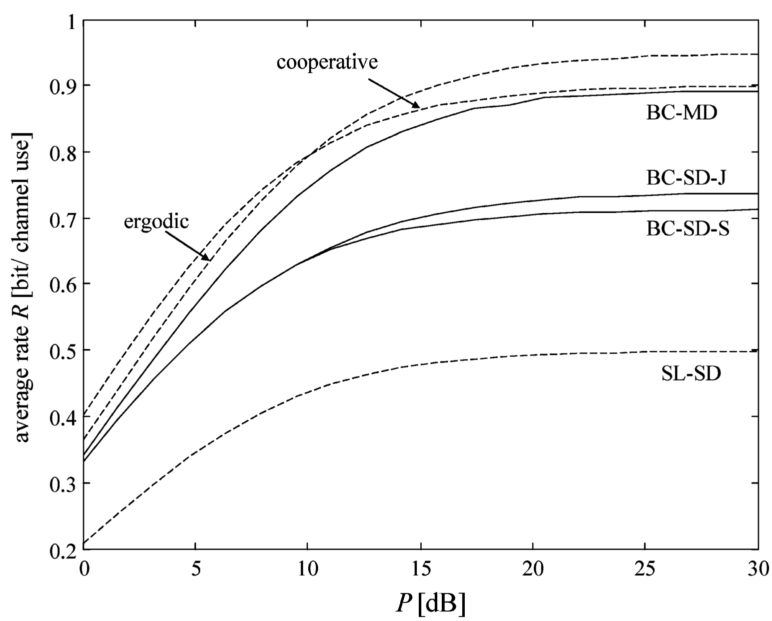

Fig. 3. Same rates as in Fig. 2 versus the power P for $C = 0.5$ and $p_1 = 0.1$.

the proposed schemes attain intermediate rates. Finally, we remark that, in general, joint decompression/decoding (BC-SD-J) is to be preferred to the separate approach (BC-SD-S) only for sufficiently large P.

Finally, Fig. 4 shows the rates at hand for $P = 10$ dB and $p_1 = 0.1$ versus the link capacity C. Here, for large C the capacity with ergodic failures (27) tends to $1/2 \log_2(1 + 2P) \simeq 2.2$ [17], while the baseline strategy SL-SD achieves the single-user rate $1/2 \log_2(1 + P) \simeq 1.72$. The proposed schemes attain intermediate rates between these two cases. Moreover, as C increases, all the proposed techniques perform very close to the upper bound of cooperative relays. This implies that, when the backhaul capacity C is large enough, there is no need for more sophisticated MD techniques.

VI. Nonoblivious Agents

In this section, we consider the model in which the agents are informed about the codebook used at the source, that is, equivalently, about the key F (nonoblivious agents, recall Section II-A). A similar model was considered in [17] for the case of fully reliable links, $M_0 = M_T$. We limit the analysis to the Gaussian model, but extension to the discrete model follows along the same lines.

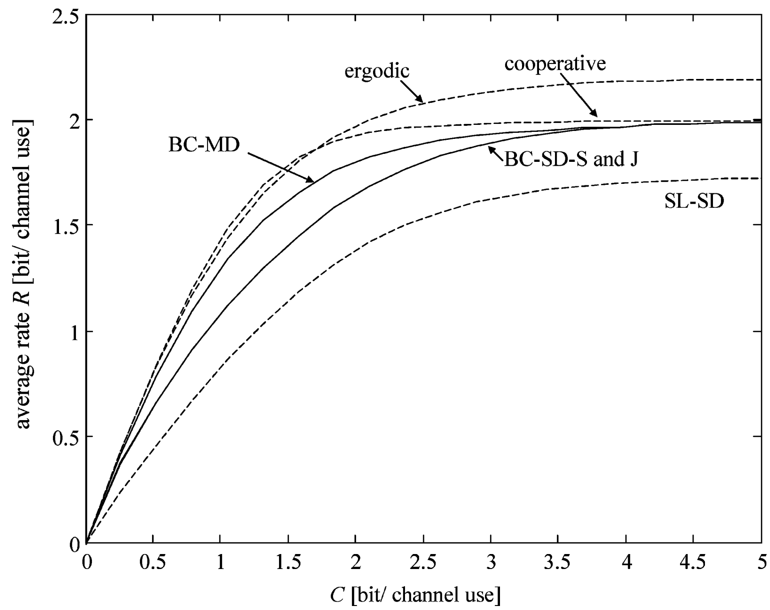

Fig. 4. Same rates as in Fig. 2 versus the link capacity C for $P = 10$ dB and $p_1 = 0.1$.

We first consider a simple upper bound on the capacity, i.e., the maximum average achievable rate, that is a direct consequence of cut-set arguments [19]. Specifically, it can be seen that the average capacity for the setup at hand is upper bounded by

$$\mathrm{C}_{\mathrm{avg}} \leq \sum_{m=M_0}^{M_T} p_m \min\left\{\frac{1}{2}\log_2(1+mP), mC\right\} \tag{29}$$

where the first term in the $\min\{\cdot,\cdot\}$ follows by considering the cut between source and agents (agents not connected to the destination cannot contribute to the rate) and the second depends on the cut from agents to destination.

As for an achievable strategy, we propose the following scheme that generalizes the BC-SD strategy considered in the previous section. We remark that an MD-based approach could also in principle be devised. However, this extension is conceptually rather straightforward and will not be further pursued here. In the proposed scheme, the source uses BC coding with Gaussian codebooks as considered throughout the rest of the paper. However, on top of the $M_T - M_0 + 1$ layers assumed in the schemes described in Section IV here the source superimposes a further layer carrying a common message, say W_0, with rate R_0, to be *decoded* by all agents (recall that in our model all agents are statistically equivalent) and then forwarded to the destination. We would like the destination to be able to recover such a message at all times, that is, as long as the number of active link M satisfies $M \geq M_0$. This is akin to the SD approach to compression studied in Section IV-A. Towards this goal, each agent reserves a rate of R/M_0 on its outgoing links to send an index computed as a random function of the decoded W_0. It can be easily seen that, even though the agents are unaware of which links are currently active (but only that $M \geq M_0$), the receiver will be able to recover W_0 with vanishing probability of error as $n \to \infty$ (this is a special case of the Slepian-Wolf problem). The extra layer carrying W_0 is decoded first by the agents and cancelled, and the rest of coding/decoding takes place as for the BC-SD-S scheme of Section IV-A with the caveat that now the remaining link capacity to forward compression indices is $C - R/M_0$.

Proposition 6.1: The average rate (2) is achievable in the presence of nonoblivious relays for the Gaussian model with

$$R_{M_0} \leq \tilde{R}_{M_0} + R_0 \tag{30a}$$
$$R_m \leq \tilde{R}_m \text{ for } m = M_0+1,\ldots,M_T \tag{30b}$$

and

$$R_0 = \frac{1}{2}\log_2\left(1 + \frac{\beta_0 P}{1+(1-\beta_0)P}\right) \tag{31}$$

with rates $\tilde{R}_m$ satisfying the inequalities in (16), and σ^2 satisfying

$$C - \frac{R_0}{M_0} \geq \frac{1}{2}\log_2\left[\left(1+\frac{M_0 P(1-\beta_0)}{1+\sigma^2}\right)^{\frac{1}{M_0}}\left(1+\frac{1}{\sigma^2}\right)\right] \tag{32}$$

for any power allocation $\beta_0, \beta_{M_0}, \ldots, \beta_M \geq 0$ with $\beta_0 + \beta_{M_0} + \cdots + \beta_M = 1$.

Proof: This result follows easily from the proof of Proposition 4.1 and the description of the scheme provided above. ■

Remark 6.1: The parameter β_0 in (31)–(32) represents the amount of power spent for transmission of message W_0. Moreover, if $\beta_0 = 0$, the rate of the proposition above reduces to the BC-SD-S scheme of Proposition 4.1.

To gain some insight into the performance of the scheme proposed above, Fig. 5 shows the average achievable rate for a two-agent system ($M_T = 2$) versus capacity C with $M_0 = 1$, for different values of p_1 (probability of $M = 1$) and $P = 10$ dB. The rates are compared with the upper bound (29) drawn for two representative values of p_1, namely 0 and 1. From (29) and the proposition above, it is noted that the cut-set bound for

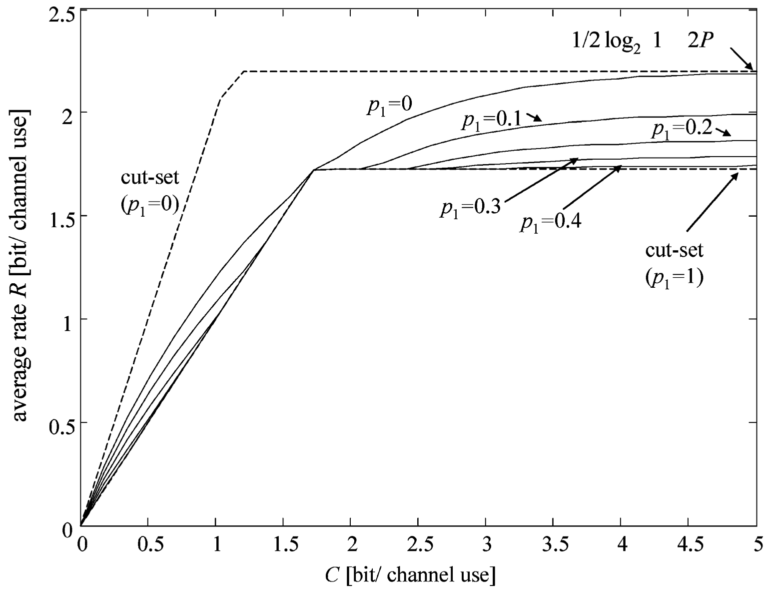

Fig. 5. Average achievable rate (30) with nonoblivious agents versus capacity C, for different values of p_1 and $P = 10$ dB. ($M_T = 2$ with $M_0 = 1$). Also shown is the cut-set bound (29) for $p_1 = 0$ and $p_1 = 1$.

$p_1 = 1$ coincides with the rate achievable by sending only the message W_0, that is, by setting $\beta_j = 0$ for $j \in [M_0, M_T]$ in (30). Therefore, for $p_1 = 1$, the proposed scheme is optimal for any value of C, and there is no need for compression of the received signal. Considering then the other limiting case, $p_1 = 0$, it is seen that the proposed scheme achieves the cut-set bound, and specifically the fully cooperative rate $1/2\log_2(1+2P)$, for C sufficiently large. Moreover, this result is achieved by setting $\beta_0 = 0$ or equivalently $R_0 = 0$, that is, by not exploiting the decoding capability of the agents. This fact is immediate if one notices that for C large and $p_1 = 0$, the two received signals can be sent by two agents with full reliability to the destination via quantization (see also [17]). Increasing p_1, the proposed scheme does not achieve the cut-set bound (not shown), even though the loss is rather limited. Furthermore, in general, for $p_1 < 1$ one can gain by using the backhaul links to send "soft" (quantization) information, along with the "hard" information on W_0, as is clear by comparing the performance with the cut-set bound with $p_1 = 1$ (see discussion above). In the example at hand, for $p_1 \gtrsim 0.4$ such gain vanishes.

VII. Concluding Remarks

In modern packet data networks serving delay-sensitive applications, link failures are often appropriately modelled as being unpredictable and nonergodic. The conventional design choice is to target worst-case scenarios by transmitting at a judiciously selected constant rate that guarantees an acceptable outage probability. However, it is often possible, and desirable, to deploy transmission strategies that are able to provide variable-rate data delivery depending on the current state of the involved links. Moreover, data communication networks typically include distributed nodes, whose operation is decentralized. In this paper, we have considered a baseline model for communication networks that include these two basic elements of nonergodic link failures and decentralized operation. Focusing on a multirelay network with one transmitter-receiver pair and unreliable orthogonal link between each relay and the destination, we have exploited the synergy between the broadcast coding approach of [6] and the distributed source coding techniques of [12] and [13] to propose a number of robust communication strategies. Via comparison with performance upper bounds, we have shown that the combination of broadcast encoding and robust multidescription compression is almost optimal for the model in which the relays are oblivious to the source codebooks. This work opens a number of possible avenues for future research, such as the extension to multiuser scenarios with more than one source.[6]

Appendix A
Proof of Proposition 3.1

The part of Proposition 3.1 regarding the Gaussian model was proved in the text. Here we concentrate on the statement regarding the discrete model. Given the symmetry of the signals received by the relays, one can assume without loss of generality that in state $m = M_0, \ldots, M_T$ relays indexed by $[1, m]$ are active, as discussed in Section II. Moreover, we define as S_m with a little abuse of notation, the signal sent by the compound relay to the destination in state m, which is a collection of the signals sent by the active relays $[1, m]$. Now notice that

$$H(S_m) \leq mC \tag{33}$$

and that S_{m-1} is a function of S_m. We also have from the Fano inequality and the definition of achievability that for all $m \in [M_0, M_T]$

$$H(W_m \,|\, S_m, F) \leq n\epsilon_n \tag{34}$$

where $\epsilon_n \rightarrow 0$ for $n \rightarrow \infty$. Notice that (34) accounts for the fact that decoding is based on the codebook key F. For any m,

[6]A scenario with interference and a single relay was recently studied in [28].

using (34), we have $H(W_m \,|\, S_m, F, W_{[M_0,m-1]}) \leq n\epsilon_n$, and thus the following conditions:

$$nR_m \leq H(W_m) = H\left(W_m \,|\, W_{[M_0,m-1]}, F\right) \tag{35}$$
$$= I\left(W_m; S_m \,|\, W_{[M_0,m-1]}, F\right) + H\left(W_m \,|\, S_m, W_{[M_0,m-1]}, F\right) \tag{36}$$
$$\leq I\left(W_m; S_m \,|\, W_{[M_0,m-1]}, F\right) + n\epsilon_n \tag{37}$$
$$= I\left(W_m, F; S_m \,|\, W_{[M_0,m-1]}, F\right) + n\epsilon_n \tag{38}$$
$$\leq I\left(W_m, F; S_m, Y_{[1,m]}^{n-1} \,|\, W_{[M_0,m-1]}, F\right) + n\epsilon_n \tag{39}$$

where (35) follows because of the independence of variables W_m and F. Define variables $V_{m,i} \triangleq (S_m, Y_{[1,m]}^{i-1})$ for $m \in [M_0, M_T]$ and $i \in [1,n]$ and notice that $V_{m-1,i}$ is a function of $V_{m,i}$. Continuing from (39) we obtain for $m \in [M_0+1, M_T-1]$

$$I\left(W_m, F; V_m^n | W_{[M_0,m-1]}, F\right) \tag{40a}$$
$$= \sum_{i=1}^{n} H\left(V_{m,i} | V_m^{i-1}, W_{[M_0,m-1]}, F\right) - H\left(V_{m,i} \,|\, V_m^{i-1}, W_{[M_0,m]}, F\right) \tag{40b}$$
$$\leq \sum_{i=1}^{n} H(V_{m,i} \,|\, U_{m-1,i}) - H\left(V_{m,i} \,|\, V_m^{i-1}, W_{[M_0,m]}, F, V_{m+1}^{i-1}\right) \tag{40c}$$
$$= \sum_{i=1}^{n} H(V_{m,i} \,|\, U_{m-1,i}) - H(V_{m,i} \,|\, U_{m-1,i}, U_{m,i}) \tag{40d}$$
$$= \sum_{i=1}^{n} I(U_{m,i}; V_{m,i} \,|\, U_{m-1,i}) \tag{40e}$$

where in (40c) we have used the fact that conditioning reduces the entropy and we have defined $U_{m,i} \triangleq (V_{m+1}^{i-1}, W_{[M_0,m]}, F)$. We consider separately the remaining cases $m = M_T$ and $m = M_0$. For $m = M_T$ from (39) we can write

$$I\left(W_{M_T}, F; V_{M_T}^n | W_{[1,M_T-1]}, F\right) \tag{41a}$$
$$= \sum_{i=1}^{n} I\left(W_{M_T}, F; V_{M_T,i} \,|\, V_{M_T}^{i-1}, W_{[1,M_T-1]}, F\right) \tag{41b}$$
$$\leq \sum_{i=1}^{n} I\left(X_i; V_{M_T,i} | U_{M_T-1,i}\right) \tag{41c}$$

where (41c) follows from the definition of $V_{m,i}$. Finally, for $m = M_0$, we can obtain similarly starting from (34)

$$nR_{M_0} \leq I\left(W_{M_0}, F; V_{M_0}^n\right) + n\epsilon_n \tag{42a}$$
$$\leq \sum_{i=1}^{n} H\left(V_{M_0,i}\right) - H\left(V_{M_0,i} | V_{M_0}^{i-1}, W_{M_0}, F, V_{M_0+1}^{i-1}\right) \tag{42b}$$
$$= \sum_{i=1}^{n} H\left(V_{M_0,i}\right) - H\left(V_{M_0,i} | W_{M_0}, F, V_{M_0+1}^{i-1}\right) \tag{42c}$$
$$\leq \sum_{i=1}^{n} I\left(U_{M_0,i}; V_{M_0,i}\right) \tag{42d}$$

where (42c) follows from the fact that $V_{m-1,i}$ is a function of $V_{m,i}$.

Overall, from the inequalities above, defining a time-sharing variable Q uniformly distributed in $[1,n]$ and independent of all other variables, we get for $m = M_0$ from (42)

$$R_{M_0} \leq I\left(U_{M_0,Q}; V_{M_0,Q} | Q\right) = I\left(U_{M_0,Q}, Q; V_{M_0,Q}, Q | Q\right) \leq I\left(U_{M_0,Q}, Q; V_{M_0,Q}, Q\right),$$

where the last inequality follows from the concavity of the mutual information with respect to the input distribution; for $m \in [M_0+1, M_T-1]$, from (40)

$$R_m \leq I(U_{m,Q}; V_{m,Q} \,|\, U_{m-1,Q}, Q) = I(U_{m,Q}, Q; V_{m,Q}, Q \,|\, U_{m-1,Q}, Q);$$

and finally for $m = M_T$ from (41)

$$R_{M_T} \leq I\left(X_Q; V_{M_T,Q}, Q \,|\, U_{M_T-1,Q}, Q\right).$$

Now, we define $U_m \triangleq (U_{m,Q}, Q) = (V_{m+1}^{Q-1}, W_{[M_0,m]}, F, Q)$, $X \triangleq X_Q$ and $V_m \triangleq (V_{m,Q}, Q) = (S_m, Y_{[1,m]}^{Q-1}, Q)$. With these definitions, we obtain (8). Moreover, the distribution of the variables at hand factorizes as (9). It remains to prove (11). This is obtained as follows:

$$nmC \geq H(S_m) = H(S_m) - H\left(S_m \,|\, Y_{[1,m]}^n\right)$$
$$= I\left(S_m; Y_{[1,m]}^n\right)$$
$$= \sum_{i=1}^{n} H\left(Y_{[1,m],i}\right) - H\left(Y_{[1,m],i} \,|\, S_m, Y_{[1,m]}^{i-1}\right)$$
$$= \sum_{i=1}^{n} I\left(V_{m,i}; Y_{[1,m],i}\right)$$
$$= nI\left(V_{m,Q}; Y_{[1,m],Q} \,|\, Q\right)$$
$$= nI\left(V_{m,Q}, Q; Y_{[1,m],Q}\right) - nI\left(Q; Y_{[1,m],Q}\right)$$
$$= nI\left(V_m; Y_{[1,m]}\right)$$

where in first line we have use the functional dependence S_m on $Y_{[1,m]}^n$, and the third line follows from the fact that, without knowledge of the key F, sequences Y_i^n are i.i.d. (see [17, Lemma 1]). This concludes the proof.

Appendix B
Proof of Proposition 4.1

We start with the discrete model.

Codeword generation and encoding (transmitter): The source employs random coding based on BC "inner" codebooks u_m^n for $m \in [M_0, M_T - 1]$ and codewords x^n generated according to the distribution (10a) [24]. Specifically, the transmitter randomly generates $2^{nR_{M_0}}$ auxiliary codewords

$u^n_{M_0}(w_{M_0})$ for $w_{M_0} \in [1, 2^{nR_{M_0}}]$, i.i.d. according to distribution $p(u_{M_0})$. Then, for each w_{M_0}, it generates $2^{nR_{M_0+1}}$ codewords $u^n_{M_0+1}(w_{M_0+1}, w_{M_0})$ with $w_{M_0+1} \in [1, 2^{nR_{M_0+1}}]$, independently for each symbol j according to distribution $p(u_{M_0+1} \,|\, u_{M_0,j}(w_{M_0}))$. The process continues similarly for all $w_{[M_0,M_T-1]} \in [1, 2^{nR_{M_0}}] \times \cdots \times [1, 2^{nR_{M_T-1}}]$, generating codewords $u^n_m(w_m, w_{[M_0,m-1]})$. Finally, the transmitter generates $2^{nR_{M_T}}$ codewords $x^n(w_{M_T}, w_{[M_0,M_T-1]})$ with $w_{M_T} \in [1, 2^{nR_{M_T}}]$, independently for each symbol j according to distribution $p(x \,|\, u_{M_T-1,j}(w_{M_T-1}, w_{[M_0,M_T-2]}))$. Encoding takes place via the mapping from messages to codewords implied by the notation used above.

Codeword generation, compression and binning (agents): Each ith agent compresses the received signal y^n_i using a randomly generated quantization codebook of $2^{n\hat{R}}$ codewords v^n_i. Quantization takes place via joint typicality according to the test channel $p(v_i \,|\, y_i)$. Specifically, each ith agent generates $2^{n\hat{R}}$ codewords v^n_i by choosing i.i.d. every symbol with probability $p(v_i)$ obtained by marginalizing (14). Compression at the agents takes place via joint typicality according to the test channel $p(v_i \,|\, y_i)$. Random binning is performed by associating (with uniform probability) a random index $s_i \in [1, 2^{nC}]$ to each codeword v^n_i, $s_i = \tilde{\phi}^{(i)}(v^n_i)$.

Separate Decompression/Decoding: When $M = m$, we can assume without loss of generality that the indices s_i for $i \in [1, m]$ are received. Based on the received s_i, the receiver first looks for an m-tuple of codewords v^n_i, $i \in [1, m]$, such that the codewords are jointly (strongly) typical (see, e.g., [19]) according to distribution $p(v_{[1,m]})$, obtained by marginalizing (14), and they belong to the received bins, $s_i = \tilde{\phi}^{(i)}(v^n_i)$. If such a tuple cannot be found or more than one such tuple is found, an error is declared. After decompression, the decoder looks sequentially for messages $w_{M_0}, w_{M_0+1}, \ldots, w_m$ such that the corresponding codewords $u^n_m(w_m, w_{[M_0,m-1]})$ are jointly typical with the decompressed vectors v^n_i [24] [19].

Analysis of the probability of error (sketch):

Compression/Decompression: It can be shown via standard rate-distortion theoretic arguments that quantization at each agent is successful with high probability (as n grows large) if $\hat{R} \geq I(V_i; Y_i)$ [19]. Moreover, as shown in [12], as long as condition (15) is satisfied, the receiver, upon reception of any subset of M_0 indices s_i, is able to decompress the corresponding M_0 quantized codewords v^n_i with vanishingly small probability as $n \to \infty$. An "ϵ-free" sketch of the proof of this fact is reported here for completeness. We need to show that the probability that a given an M_0-tuple of codewords $\bar{v}^n_{[1,M_0]}$, different from the correct one, happens to satisfy $s_i = \tilde{\phi}^{(i)}(\bar{v}^n_i)$ (i.e., to be within the received bins) and to be jointly typical vanishes with large n. The number of jointly typical M_0-tuples of sequences $v^n_{[1,M_0]}$ is approximately $2^{nH(V_{[1,M_0]})}$, while the number of M_0-tuples of sequences $v^n_{[1,M_0]}$ that are obtained by generating each sequence i.i.d. and independently from the others is about $2^{nM_0H(V_i)}$. Therefore, according to the codeword generation described above, the probability that a M_0-tuples of codewords $\bar{v}^n_{[1,M_0]}$, different from the correct one, happens to be jointly typical, is around $2^{n(H(V_{[1,M_0]})-M_0H(V_i))}$. The total number of codewords within each bin is $2^{n(\hat{R}-C)}$, so that the probability that an M_0-tuples of codewords $\bar{v}^n_{[1,M_0]}$ is jointly typical and within the received bins is upper bounded by

$$2^{nM_0(\hat{R}-C)} \cdot 2^{n\left(H\left(V_{[1,M_0]}\right)-M_0H(V_i)\right)}. \tag{43}$$

Recalling that $\hat{R} \geq I(V_i; Y_i)$, we obtain that condition (15) is sufficient to drive the probability (43) to zero.

Decoding: Decoding takes place based on codewords v^n_i decompressed with vanishing probability of error (as shown above) over the M active links. It is clear that, by symmetry, signals received when $M = m$ links are active can be considered to be $v^n_{[1,m]}$ and that we have the following Markov chain condition

$$X^n - V^n_{[1,M_T]} - V^n_{[1,M_T-1]} - \cdots - V^n_{[1,M_0]}. \tag{44}$$

Therefore, the channel between X^n and $V^n_{[1,m]}$ for $m \in [M_0, M]$ can be seen as a degraded broadcast channel and the rate region (13) easily follows from [24].

For the Gaussian channel, the same reasoning applies and we set the variables at hand as $U'_i \sim \mathcal{N}(0, \beta_i P)$ independent for $i \in [M_0, M_T]$, $U_m = \sum_{i=M_0}^{m} U'_i$ and $X = \sum_{i=M_0}^{M_T} U'_i$. Moreover, the test channel is selected as $V_i = Y_i + Q_i$, where $Q_i \sim \mathcal{N}(0, \sigma^2)$ independent of all other variables.

APPENDIX C
PROOF OF PROPOSITION 4.2

Codeword generation, encoding and compression take place as discussed in Appendix B. However, joint decompression/decoding is applied at the receiver.

Joint decompression/decoding: The receiver, when $M = m \in [M_0, M_T - 1]$ active links are available, say $[1, m]$ without loss of generality, performs joint decompression and decoding by looking for messages $w_{[M_0,m]}$ such that there exist codewords v^n_i for $i \in [1, m]$ for which $(u^n_{m-1}(w_{[M_0,m-1]}), u^n_m(w_{[M_0,m]}), v^n_1, \ldots, v^n_m)$ are jointly typical according to the distribution (14) and $s_i = \tilde{\phi}^{(i)}(v^n_i)$ (that is, codewords v^n_i belong to the bins indexed via the backhaul link by s_i) and no other $w_{[M_0,m]}$ satisfies these properties. A similar procedure applies for $M = M_T$.

Analysis of probability of error (sketch): When considering the mth layer ($m \in [M_0, M_T - 1]$), corresponding to message W_m of rate R_m, we can assume, in deriving the conditions for vanishing probability of error, that the layers $[M_0, m - 1]$ have been correctly decoded, due to the Markov condition (44). Moreover, errors in the quantization process have arbitrarily small probability due to the choice of rate $\hat{R}$. Therefore, given the transmission of a given set of messages $w_{[M_0,m]}$ and the compression to a set of codewords $v^n_1, \ldots, v^n_m$, an error event at the mth layer occurs if the decoder can find a $\bar{w}_m \in [1, 2^{nR_m}]$, with $\bar{w}_m \neq w_m$, and $k \in [0, m]$ sequences $\bar{v}^n_1, \ldots, \bar{v}^n_k \neq v^n_1, \ldots, v^n_k$ such that the set of sequences $(u^n_{m-1}(w_{[M_0,m-1]}), u^n_m(w_{[M_0,m-1]}, \bar{w}_m), \bar{v}^n_1, \ldots, \bar{v}^n_k, v^n_{k+1}, \ldots, v^n_m)$ are jointly typical according to the distribution (14) and $s_i = \tilde{\phi}^{(i)}(v^n_i) = \tilde{\phi}^{(i)}(\bar{v}^n_i)$ for $i \in [1, k]$ (i.e., codewords $\bar{v}^n_i$ are in the same bin as v^n_i). Notice that due to symmetry there is no loss of generality in considering only

$$
\begin{aligned}
R_m &\le k(C - I(V_i;Y_i) + H(V_i)) - H(U_{[m-1,m]}, V_{[1,m]}) + H(U_m \,|\, U_{m-1}) + H\left(V_{[k+1,m]}, U_{m-1}\right) \\
&= k(C + H(V_i \,|\, Y_i)) - H(U_{m-1}) - H\left(V_{[1,m]} \,|\, U_{[m-1,m]}\right) + H\left(V_{[k+1,m]}, U_{m-1}\right) \\
&= k(C + H(V_i \,|\, Y_i)) - H\left(V_{[1,m]} \,|\, U_{[m-1,m]}\right) + H\left(V_{[k+1,m]} \,|\, U_{m-1}\right) \\
&= k(C + H(V_i|Y_i) - \frac{1}{k} H\left(V_{[1,k]}|U_{[m-1,m]}, V_{[k+1,m]}\right)) + I\left(U_m; V_{[k+1,m]} \,|\, U_{m-1}\right).
\end{aligned}
$$

$$
h\left(\sum_{i=m+1}^{M_T} U_i' | \tilde{V}_{[k+1,m]}\right) = \frac{1}{2}\log\left(\sum_{i=m+1}^{M} \beta_i P - \frac{(m-k)\left(\sum_{i=m+1}^{M}\beta_i P\right)^2}{\sum_{i=m+1}^{M}\beta_i P + 1 + \sigma^2}\right)
$$

$$
h\left(\sum_{i=m+1}^{M_T} U_i' | \tilde{V}_{[1,m]}\right) = \frac{1}{2}\log\left(\sum_{i=m+1}^{M} \beta_i P - \frac{j\left(\sum_{i=m+1}^{M}\beta_i P\right)^2}{\sum_{i=m+1}^{M}\beta_i P + 1 + \sigma^2}\right).
$$

events with erroneous $\bar{v}_i^n$ as selected here. Similar to [17], the probability of any of the $2^{n[k(\hat{R}-C)+R_m]}$ set of sequences of the type above to be jointly typical is upper bounded by

$$
2^{n\left[H(U_{[m-1,m]}, V_{[1,m]}) - k\cdot H(V_i) - H(U_m \,|\, U_{m-1}) - H(V_{[k+1,m]}, U_{m-1})\right]}
$$

so that, to drive the probability of error to zero with large n, we need (recall that $\hat{R} \ge I(V_i;Y_i)$) (see the top equation at the top of the page).

One can then repeat similar arguments for the last layer $M = M_T$, and thus conclude the proof for the discrete case.

For the Gaussian case, we assume Gaussian channel codebooks by setting auxiliary variables $U_i' \sim \mathcal{N}(0, \beta_i P)$ independent for $i \in [M_0, M_T]$, $U_m = \sum_{i=M_0}^{m} U_i'$ and $X = \sum_{i=M_0}^{M_T} U_i'$. Moreover, we set Gaussian quantization codebooks $V_i = Y_i + Q_i$ with $Q_i \sim \mathcal{N}(0, \sigma^2)$ independent of all other variables. Extension of the rates derived above to the Gaussian case then requires to substitute differential entropies for entropies in (18)–(19) and evaluate the corresponding information measures for the distributions given earlier. We have that $h(V_{[1,k]} \,|\, U_m, V_{[k+1,m]}) = h(V_{[1,k]} \,|\, X, U_m, V_{[k+1,m]}) + I(X; V_{[1,k]} \,|\, U_m, V_{[k+1,m]})$, but $h(V_{[1,k]} \,|\, X, U_m, V_{[k+1,m]}) = k/2 \cdot \log(2\pi e(1+\sigma^2))$, since, conditioned on X, variables V_j are independent of V_i with $i \ne j$ and of all U_i. Moreover, we have

$$
\begin{aligned}
&I\left(X; V_{[1,k]}|U_m, V_{[k+1,m]}\right) \\
&\quad = h\left(X \,|\, U_m, V_{[k+1,m]}\right) - h(X \,|\, U_m, V_{[1,m]})
\end{aligned}
$$

and the first term reads, defining

$$
\tilde{V}_j = \sum_{i=m+1}^{M} U_i' + Z_j + Q_j
$$

(see the center equation at the top of the page), while the second is (see the bottom equation at the top of the page). This concludes the proof.

REFERENCES

[1] M. Alasti, K. Sayrafian-Pour, A. Ephremides, and N. Farvardin, "Multiple description coding in networks with congestion problem," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 891–902, Mar. 2001.

[2] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[3] D. Gündüz and E. Erkip, "Source and channel coding for quasi-static fading channels," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 28–Nov. 1 2005, pp. 18–22.

[4] S. Shamai (Shitz) and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[5] C. T. K. Ng, D. Gündüz, A. Goldsmith, and E. Erkip, "Minimum expected distortion in Gaussian layered broadcast coding with successive refinement," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, June 24–29, 2007, pp. 2226–2230, (ISIT 2007).

[6] C. Tian, A. Steiner, S. Shamai (Shitz), and S. N. Diggavi, "Successive refinement via broadcast: Optimizing expected distortion of a Gaussian source over a Gaussian fading channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2903–2918, Jul. 2008.

[7] M. Effros, A. Goldsmith, and Y. Liang, "Capacity definitions for general channels with receiver side information," submitted [arXiv:0804.4239].

[8] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile *ad hoc* networks," *IEEE Commun. Mag.*, vol. 46, no. 12, pp. 94–101, Dec. 2008.

[9] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[10] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions—Part I: (n,k) source-channel erasure codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[11] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions—Part II: An achievable rate-distortion region," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1377–1392, Apr. 2005.

[12] P. Ishwar, R. Puri, K. Ramchandran, and S. S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 765–775, Apr. 2005.

[13] J. Chen and T. Berger, "Robust distributed source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3385–3398, Aug. 2008.

[14] B. E. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2001.

[15] M. R. Aref, "Information flow in relay networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1981.

[16] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proc. 45th Ann. Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, 2007, pp. 129–135.

[17] A. Sanderovich, S. Shamai (Shitz), Y. Steinberg, and G. Kramer, “Communication via decentralized processing,” *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3008–3023, Jul. 2008.

[18] V. Prabhakaran, D. N. C. Tse, and K. Ramachandran, “Rate region of the quadratic Gaussian CEO problem,” in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jul. 2004, p. 119, (ISIT 2004).

[19] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 2006.

[20] T. Berger, “Multiterminal source coding,” presented at the CISM Summer School on the Inf. Theory Approach to Commun. Udine, Italy, Jul. 1977.

[21] D. N. Tse, “Optimal power allocation over parallel Gaussian broadcast channels,” in *Proc. IEEE Int. Symp. Inf. Theory*, Ulm, Germany, 1997, pp. 27–29, (ISIT 1997).

[22] T. Cover, “Broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[23] A. F. Dana, R. Gowaikar, P. Palanki, B. Hassibi, and M. Effros, “Capacity of wireless erasure networks,” *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 789–804, Mar. 2006.

[24] P. P. Bergmans, “Random coding theorem for broadcast channels with degraded components,” *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 197–207, Mar. 1973.

[25] M. Gastpar, “The Wyner-Ziv problem with multiple sources,” *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.

[26] H. Feng and M. Effros, “On the achievable region for multiple description source codes on Gaussian sources,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2003)*, Yokohama, Japan, Jun. 29–Jul. 4 2003, p. 195.

[27] O. Somekh, O. Simeone, H. V. Poor, and S. Shamai (Shitz), “The two-tap input-erasure Gaussian channel and its application to cellular communications,” in *Proc. 46th Ann. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Sep. 2008.

[28] O. Simeone, E. Erkip, and S. Shamai (Shitz), “Oblivious relaying for primitive interference relay channels,” in *Proc. 2010 Int. Zurich Semin. Commun.*, Zurich, Switzerland, Mar. 3–5, 2010.

Osvaldo Simeone (M’02) received the M.Sc. degree (with honors) and the Ph.D. degree in information engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively.

He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark, where he is an Assistant Professor. His current research interests concern the cross-layer analysis and design of wireless networks with emphasis on information-theoretic, signal processing, and queuing aspects. Specific topics of interest are: cognitive radio, cooperative communications, *ad hoc*, sensor, mesh and hybrid networks, distributed estimation, and synchronization.

Dr. Simeone is a corecipient of Best Paper Awards of the IEEE SPAWC 2007 and IEEE WRECOM 2007. He currently serves as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS.

Oren Somekh (S’90–M’06) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion-Israel Institute of Technology, Haifa, in 1989, 1991, and 2005, respectively.

During 1991–1996, he served in the Israel Defense Forces working as a research engineer. During 1998–2002, he was the Vice President of Research and Development and later, the Chief Technical Officer of Surf Communication Solutions Ltd., Yokneam, Israel. Soon after, he cofounded and managed Kolvirtech Ltd., Pardes-Hana, Israel, a consultant and outsourcing company. From April 2005 to October 2006, he was a Visiting Research Fellow with the Center for Wireless Communications and Signal Processing Research, New Jersey Institute of Technology (NJIT), Newark. From November 2006 to October 2008, he was a Visiting Research Fellow with the Electrical Engineering Department, Princeton University, Princeton, NJ. During 2009, he was a Visiting Research Fellow with the Electrical Engineering Department, Technion. He joined the Applied Sciences Group of Yahoo! Labs, Israel, in 2009. He is also a research collaborator with the Electrical Engineering Department, Princeton University, Princeton, NJ. He is the coinventor of several patents in the field of communications. His current main research interests are scientific aspects of Internet technologies, and information theoretical aspects of cooperative wireless networks.

Dr. Somekh is a recipient of the European Community Marie-Curie Outgoing International Fellowship. He was the corecipient of the first IEEE Information Theory Society ISIT Best Student Paper award (2007).

Elza Erkip (S’93–M’96–F’11) received the B.S. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA.

Currently, she is an Associate Professor of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY. In the past, she has held positions with Rice University, Houston, TX, and with Princeton University, Princeton, NJ. Her research interests are in information theory, communication theory, and wireless communications.

Dr. Erkip received the National Science Foundation CAREER Award in 2001, the IEEE Communications Society Rice Paper Prize in 2004, and the ICC Communication Theory Symposium Best Paper Award in 2007. She coauthored a paper that received the ISIT Student Paper Award in 2007. She was a Finalist for The New York Academy of Sciences Blavatnik Awards for Young Scientists in 2010. Currently, she is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and a Technical Co-Chair of WiOpt 2011. She was an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS during 2006–2009, a publications editor of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2006–2009, and a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE in 2007. She was the Co-Chair of the GLOBECOM Communication Theory Symposium in 2009, the Publications Chair of ITW Taormina in 2009, the MIMO Communications and Signal Processing Technical Area chair of the Asilomar Conference on Signals, Systems, and Computers in 2007, and the Technical Program Co-Chair of the Communication Theory Workshop in 2006.

H. Vincent Poor (S’72–M’77–SM’82–F’87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. His research interests are in the areas of stochastic analysis, statistical signal processing and information theory, and their applications in wireless networks and related fields. Among his publications in these areas are *Quickest Detection* (Cambridge University Press, 2009), coauthored with O. Hadjiliadis, and *Information Theoretic Security* (Now Publishers, 2009), coauthored with Y. Liang and S. Shamai.

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U.K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, in 2004–2007 as the Editor-in-Chief of these TRANSACTIONS, and in 2009 as General Co-Chair of the IEEE International Symposium on Information Theory, held in Seoul, South Korea. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society, the 2009 Edwin Howard Armstrong Achievement Award of the IEEE Communications Society, the 2010 IET Ambrose Fleming Medal for Achievement in Communications, and the 2011 IEEE Eric E. Sumner Award.

Shlomo Shamai (Shitz) (S’80–M’82–SM’88–F’94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1975, 1981, and 1986, respectively.

During 1975–1985, he was a Senior Research Engineer with the Communications Research Labs. Since 1986, he has been with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is the recipient of the 2011 Claude E. Shannon Award, the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a corecipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served on the Board of Governors of the Information Theory Society.