Achievable Rates for a LAN-Limited Distributed Receiver in Gaussian Interference

Christian D. Chapman, Hans Mittelman, Adam R. Margetts, Daniel W. Bliss IEEE, Fellow

Abstract—A base node seeks to receive a broadcast with its own observation in addition to side-information provided via a local area network (LAN) from several ‘helper nodes,’ potentially occluded by an external interferer. Ideally, helpers would convey their precise observations to the base but the LAN has limited capacity, so helpers must compress and forward. Bounds are found on the capacity of such a system with channel state information available at the helpers and in interference environments where the receivers experience correlated noise. The bounds are tight when the LAN allows for either a great or small amount of communications from helpers to base. The bounds also demonstrate that such systems are robust in mitigating low-rank interferers. Trials of the bounds reveal that while the achievable rates are stable in varying the ratio of scattered-path to line-of-sight signal power seen at the helpers, strategies for achieving these rates must change.

I. INTRODUCTION

We consider a problem where a group of single-antenna listeners seek to receive a broadcast message, but none can observe the broadcast with much clarity. This is perhaps due to weak channels or the presence of interferers. The listeners wish to pool their information to form a better view of the broadcaster, but cannot do so directly because connectivity between receive nodes is limited. In our formulation of the problem, the listeners identify one of the receive nodes as a base and the rest as helpers. Each helper then summarizes its observation to the base within the constraints of the local link. The base then attempts to recover the broadcast message from this data. A rough sketch of the system is shown in Figure 1.

A. Contributions

We perform an information theoretic analysis of the system and the rate at which data can be conveyed from source to base. The following results are presented:

• Upper and lower bounds on the system’s achievable rate, and regimes where these bounds are tight (Section IV).
• Performance of the bounds in interference scenarios of interest (Section V-A). These systems are mostly able to mitigate the distortion from a low-rank interferer.
• Performance and behavior of the strategies in various scattering environments (Section V-B). The scattering environment is seen to not affect performance on average.
• Strengthening of a previous result, where the same rate is achieved with weaker conditions (Remark 6).

II. BACKGROUND

Broadcasting to distributed receivers has been studied in many contexts, and in most conceivable situations is an instance of communications over a single-input multiple-output (SIMO) channel. A variety of SIMO channels have been extensively analyzed, but are slightly different than the problem considered here. Our work is in the context of information theory as discussed by Shannon (Reference [1]) and Cover (Reference [2]).

Results presented here are a significant and direct extension of the work in [3]. The differences are as follows: each bound in [3] is presented with greater generality and extended to the case where the base node has its own observation, proofs of achievability of the bounds are added, plots account for the system’s interference mitigation capability whereas ones in [3] do not, and an original, tigher lower bound is presented in Theorem 3 along with Remark 6, which relates Theorem 3 to other literature.

The system studied here is a special case of the relay structure in Reference [4]. The source-to-relay link is the channel from the source to all the helpers. The relay-to-destination link
is the LAN, available for design within constraints (each helper can only use its own observation, the LAN only supports a limited sum of data rates) and independent from the source-to-relay channel so that no self interference is experienced. Studies such as References [5], [6], [7] detail using a collection of receivers as beam forming relays to a destination node in a network with structure similar to the system considered here. In our situation, each helper-to-base link is orthogonal to other links, so results about beam forming are not pertinent.

Performance of specific coding schemes for this system have been studied in References [8], [9], [10]. In particular, Reference [10] presents a scheme that can perform to within 1.5 dB of allowing the base and receivers to communicate without constraints. In contrast, the results here build towards characterizing achievable rates of the system rather than designing and analyzing the performance of specific codes.

The problem we consider is similar to many-help-one problems such as the ‘Central Estimating Officer (CEO) problem’ posed by Berger in Reference [11], where a CEO node seeks to estimate some message by listening to a set of ‘agent’ nodes which communicate to the base at some rate (much like the helpers discussed here). Many variations of this problem have been studied, for instance in Reference [12] for many agents. The focus of the CEO problem and most of its derivatives are to estimate a source seen by the agents with distortion (often mean squared error). We focus on finding rates for lossless communications.

The noisy Slepian-Wolf problem (Reference [13]) can be interpreted as communications in the opposite direction of our system: distributed, cooperative helpers have a message for a base, but their modulations must be sent over a noisy channel.

References [14], [15] analyze the capacity region of a system with nearly identical structure to the one considered here. In our model, the base has its own reception while in References [14], [15] it only has side information provided by the helpers. In these studies, achievable rates are shown for general discrete channels and AWGN channels without interference. The rates are achieved by maximizing over source and helper message distributions with the appropriate Markov conditions. Additionally, Reference [14] demonstrates that the achievable rate for the AWGN channel is in fact a the capacity when helpers are disallowed from decoding the transmitter’s broadcast. We establish an extension of this theorem to Gaussian interference channels in Theorem 3.

III. PROBLEM SETUP

Throughout the paper we use notation from Table I. At each time slot \( t \in \mathbb{N} \), a single-antenna broadcaster emits a random complex-valued signal, \( X_t \) with a power constraint \( \mathbb{E}[|X_t|^2] = 1 \). Over a period of \( t \) channel uses, a sequence of \( t \) values are broadcast, \( X = (X_1, \ldots, X_t) \). Because the time slot extension is not useful outside of proofs, in formulae outside the Appendix we focus on a single time slot where a value \( X \in \mathbb{C} \) is broadcast.

The signal is observed by \( N + 1 \) single-antenna receive nodes through a static flat fading channel and additive white Gaussian noise. Enumerating these receivers 0 through \( N \), the “0th” receiver is identified as the base, and the other \( N \) receivers are called helpers. The goal of the system is for the broadcaster to convey as much average-information-per-channel-use as possible from the broadcaster to the base.

For \( n \in \{0, \ldots, N\} \), the \( n \)th receiver observes

\[
Y_n = h_nX + W_n
\]

where

- \( h \in \mathbb{C}^{(N+1) \times 1} \) is a deterministic channel
- \( W = (W_0, \ldots, W_N) \sim \mathcal{CN}(0, \Sigma) \) is noise, with some covariance matrix \( \Sigma \)

The vector of helper’s observations is denoted as

\[
\mathbf{Y} = (Y_0, Y_1, \ldots, Y_N) \in \mathbb{C}^{(N+1) \times 1}.
\]

It is instructive to see that if the noise covariance \( \Sigma \) is diagonal then there is no interference, and receivers experience totally independent noise. In contrast, if we can perform the decomposition \( \Sigma = aa^\dagger \) then receivers experience no noise other than interference.

Each helper can only use its own observation and does not have any access to the other helpers’ receptions. Helper \( n \) for \( n \in \{1, \ldots, N\} \) encodes its observations into an \( r_n \)-average-bitrate summary and forwards it to the base over a Local Area Network (LAN). To reconstruct the signal the base uses its own full-precision observation and side information it receives from the helpers. The vector of the helper’s rates is denoted

\[
r = (r_1, \ldots, r_N) \in \mathbb{C}^{(N+1) \times 1}.
\]

We assume that a LAN has been established amongst the receivers, and that each helper-to-base link is lossless and orthogonal to other links in the system so there is no self-interference. Any LAN will not support an arbitrarily large communication rate between the helpers and the base, which we model by asserting that all feasible rate vectors \( r \) belong to a set \( \mathcal{R}_{\text{LAN}}(L) \) with a sum-capacity parameter \( L > 0 \)

\[
\mathcal{R}_{\text{LAN}}(L) \triangleq \left\{ r \left| \sum_{n=1}^{N} r_n \leq L, \ r_n \geq 0 \right. \right\}.
\]

We refer to the condition that \( r \in \mathcal{R}_{\text{LAN}}(L) \) as the LAN constraint. Although the LAN constraint is referenced throughout the paper, no results other than graphs and Remark 4 depend on the form of \( \mathcal{R}_{\text{LAN}} \). All other theorems and remarks hold for arbitrary sets of feasible rates.

Helper \( n \) for \( n \in \{1, \ldots, N\} \) employs a quantizer block \( Q_n \) to produce its coarse \( r_n \)-bitrate summary of its observations. \( Q_n \) is a channel with Shannon capacity at most \( r_n \), where the channel does not contain the signal or other helper’s receptions as a hidden variable. The \( n \)th receiver’s quantizer output is a finite-alphabet random variable \( U_n \), and the vector of helper messages received by the base at each channel use is denoted

\[
\mathbf{U} = (U_1, \ldots, U_N).
\]

This description of the quantizers precisely is equivalent to the two statements:
The probability measure of \( (X, Y, U) \) is separable as
\[
P(X, Y, U) = P_X \cdot P_{Y|X} \cdot \prod_{n=1}^{N} P_{U_n|Y_n} \tag{6}
\]

- Messages can be encoded to within their respective helper-to-base rates
\[
I(Y_n; U_n) \leq r_n, \quad n = 1, \ldots, N \tag{7}
\]

\( h \) and \( \Sigma \) are assumed to be static throughout the transmission of each message, and have been estimated by the receivers a priori and are known at both the broadcast and receive nodes. We assume that the transmitter and the distribution of \( X \) are available for design, and that this distribution is known to all the receive nodes. Both the transmitter and the receivers also have knowledge of the set of feasible rates, \( \mathcal{R}_{LAN}(L) \). A block diagram of the system is shown in Figure 2.

The environment determines:
- channel fades \( h \)
- noise covariance \( \Sigma \)
- maximum LAN throughput \( L \)

Available for design are:
- the distribution of the source \( X \)
- the rates for the base to collect from each helper \( r \)
- the behavior of quantizers \( Q_1, \ldots, Q_K \)
- the combining method at the base

We are interested in finding the maximum rate at which data can be sent from transmitter and received at the base with negligible error probability.

### IV. BOUNDS ON COMMUNICATIONS RATES

A rough initial bound on the system’s capacity follows by expanding the mutual information from source to base:
\[
I(X; \hat{X}) \leq I(X; Y_0, U) \leq I(X; U) + I(X; Y_0) \leq L + \log \left[ 1 + \|h_0\|^2 / \Sigma_{1,1} \right].
\tag{10}
\]

If the LAN constraint were not imposed and the base had access to the helpers’ receptions in full precision, the receive side would be equivalent to a multi-antenna receiver. By the noisy channel coding theorem (Reference [1]), we can compute the capacity of this channel by maximizing \( I(X; Y) \) over distributions on \( X \). This derivation is presented in more detail in Reference [16], and gives a familiar upper bound:
\[
\log \left[ \frac{\|\Sigma + hh^H\|^2}{|\Sigma|} \right].
\tag{11}
\]

Taking the minimum of this bound and (10) yields an upper bound for our original system.

**Remark 1.** Any rate of reliable communications \( R \) for the system must satisfy
\[
R \leq \min \left\{ \log \left[ \frac{\|\Sigma + hh^H\|^2}{|\Sigma|} \right], L + \log \left[ 1 + \|h_0\|^2 / \Sigma_{1,1} \right] \right\}.
\tag{12}
\]

**Proof:** Justified by the preceding discussion.

**A. Achievable Rate through Dirty Paper Coding**

Treating each helper node as a user seeking to receive its own message, the link from the broadcaster to helpers is a Gaussian broadcast channel. The capacity region of this channel has been characterized in Reference [17], and is succinctly stated in Reference [18]. Sharing power between messages to each receiver at the transmitter, using dirty paper coding (Reference [19]) to mitigate interference from one receiver’s message to the others is the optimal strategy. We denote the region as \( \mathcal{R}_{DPC} \).

**Remark 2.** The following rate is achievable
\[
R_{DPC}(L) = \max_{r \in \mathcal{R}_{DPC} \cap \mathcal{R}_{LAN}(L)} \sum_{n=0}^{N} r_n.
\tag{13}
\]

**Proof:** Fix any rate vector \( r \in \mathcal{R}_{DPC} \cap \mathcal{R}_{LAN}(L) \), have each receiver decode its own message, and have helpers forward to the base. By Reference [17] we are guaranteed a strategy for reliable communications of each receiver’s message, and the LAN constraint is satisfied by construction of \( \mathcal{R}_{DPC} \cap \mathcal{R}_{LAN}(L) \).

This rate can be improved by considering broadcast channels which also sends a common message to each subset of helpers.
the subscribers. We suspect that this strategy yields the capacity, but to the author’s knowledge this has not been studied.

**Remark 3.** When the base does not have its own reception (i.e. \(I(X;Y_0) = 0\)) and

\[
L = \min_{A > 0} \log \left( \frac{|A + hh^H|}{|A|} \right)
\]

then the system’s capacity is \(L\).

**Proof:** By assumption, the upper bound in (12) equals \(L\). We have from the main result of Reference [20] that the maximum sum-rate of the vector Gaussian broadcast channel is the right side of (14), so \(\mathcal{R}_{\text{DPC}}\) must contain a point with sum-rate \(L\). By construction, this point is also in \(\mathcal{R}_{\text{LAN}(L)}\), so by Remark 2 a broadcaster-to-base rate \(L\) is achievable.

Unfortunately the technique used to achieve this rate requires a strong amount of cooperation from the broadcaster, and implementation of dirty paper coding is a current field of study (see Reference [21]).

### B. Achievable Rate through Gaussian Distortion

Rate-distortion theory shows that in order to encode a Gaussian source \(Y \sim \mathcal{N}(0, \sigma^2)\) with minimum rate such that distortion does not exceed some maximum allowable mean squared error \(D > 0\), then the minimum rate where this is feasible is

\[
\mathcal{R}(D) = \log(\sigma^2) - \log(D).
\]

To achieve this rate, the encoding operation must emulate the following test channel (Reference [2]):

\[
\begin{align*}
\alpha / \beta & \quad \beta \\
Y & \rightarrow \otimes \rightarrow \oplus \rightarrow \otimes \rightarrow Z = \alpha Y + \beta W \\
W & \sim \mathcal{C}\mathcal{N}(0, 1)
\end{align*}
\]

\[\otimes\] is a scalar multiply, \(\alpha = \sqrt{1 - D/\sigma^2}\) and \(\beta = \sqrt{D}\). If perfect Rate-distortion encoders could be realized at each helper, each helper side information would provide the helper’s observation with some added Gaussian noise, the amount depending on the encoding rate. Quantizers like this can be realized in limit with code length using dithered lattice quantization techniques presented in Reference [22]. Decodings of messages from a helper transmitting at rate \(r_n\) would have the form

\[
Z_n(r_n) \triangleq Y_n + W_{Q,n}(r_n) \in \mathbb{C}^{N \times 1}
\]

where \(W_{Q,n}(r_n)\) is independent Gaussian distortion added from quantization whose variance is some function of the rate \(r_n\). This function will be derived in the following paragraph. We denote the distortion vector as

\[
W_Q(r) \triangleq (W_{Q,1}(r_1), \ldots, W_{Q,N}(r_N)) \in \mathbb{C}^{N \times 1}
\]

and the vector of \(Z_n\)s corresponding to this choice of \(W_{Q,n}\) as:

\[
Z(r) \triangleq Y + W_Q(r).
\]

The subscript \(Q\) is a decoration to distinguish quantization distortion terms \(W_Q, W_{Q,n}\). We refer to this system as a Gaussian distortion system. A block diagram of it is shown in Figure 3.

If the \(n^{th}\) helper is to forward information to the base at rate \(r_n\), then the amount of distortion in the helper’s encoding under this strategy can be determined by setting (15) equal to \(r_n\) and solving for \(D\), with \(\sigma^2\) equal to the helper observation’s variance, \(||h_n||^2 + \Sigma_{n,n}\). Scaling the output of the test channel the helper emulates (Equation (16)) by \(\alpha/\beta\) causes it to equal the helper’s observation plus independent Gaussian noise with variance \((\beta/\alpha)^2 = D/(1 - D/\sigma^2)\). This means that in such a system, the distortion from the helpers is equivalent to adding independent additive Gaussian noise with variance:

\[
\text{Var}(W_{Q,n}(r_n)) = \frac{||h_n||^2 + \Sigma_{n,n}}{2^{r_n} - 1}.
\]

By (1) and (20), all the noise and distortion on the signal present in the helpers’ messages to the base are summarized by the vector \(W + W_Q\) with covariance matrix

\[
D(r) = \Sigma + \text{diag} \left( \frac{||h_1||^2 + \Sigma_{1,1}}{2^{r_1} - 1}, \ldots, \frac{||h_N||^2 + \Sigma_{N,N}}{2^{r_N} - 1} \right).
\]

Since the noise and distortion is Gaussian, the following rate is achievable:

\[
R_G(r) = \max I(X; Z(r)) = \log \left[ \frac{||D(r) + hh^H||}{||D(r)||} \right].
\]

With this intuition we can state the following:

**Theorem 1.** For a distributed receive system as described in Section III with noise covariance matrix \(\Sigma\), LAN constraint \(L\) and fixed average helper quantization rates \(r \in \mathcal{R}_{\text{LAN}(L)}\), then a rate \(R_G(r)\) is achievable.

A proof of this is given in Appendix A.

Maximizing \(R_G\) over \(\mathcal{R}_{\text{LAN}(L)}\) gives a lower bound on the system capacity:

\[
R_{G,\text{max}}(L) \triangleq \max_{r \in \mathcal{R}_{\text{LAN}(L)}} R_G(r).
\]

Unfortunately this expression cannot be simplified much further outside of a few special cases. Maximization of \(R_G\) over \(\mathcal{R}_{\text{LAN}(L)}\) can be performed efficiently using quasi-convex optimization techniques (See Reference [23]).

**Remark 4.** The max of \(R_G(r)\) follows the maximum of \(\mathbf{h}^T D(r)^{-1} \mathbf{h}\), which is quasi-concave in \(r\).

**Proof:** The maximum of \(R_G\) follows the maximum of \(\mathbf{h}^T D(r)^{-1} \mathbf{h}\) by the matrix determinant lemma.

It suffices to show that the restriction of this functional on the intersection of any line with \(\mathbb{R}^{N \times 1}_+\) is quasi-concave (see Reference [23]). Further, a 1-dimensional function is quasi-concave if anywhere its derivative is 0, its second derivative is below 0.

Fix any \(a, b \in \mathbb{R}^{N \times 1}_+\), denote \(D_t \triangleq D(\cdot)|_{t \alpha + b}\) for any \(t\) where \(t \alpha + b \in \mathbb{R}_+^{N \times 1}\). Define \(A_t \triangleq \frac{d}{dt} D_t\), and \(f(t) \triangleq -h^T D_t^{-1} h\). Then

\[
\frac{d}{dt} f(t) = A_t D_t^{-1} h
\]
compress their encodings with no loss down to rates $r$, and the encodings $U$ satisfy the conditions that for all subsets $S \subseteq \{1, \ldots, N\}$, then:
\[
H(U_S|U_{SC}, Y_0) < \sum_{n \in S} r_n.
\] (27)

Note that $Y_0$ is always included in the conditioning, because $Y_0$ is available at the base node in full precision. In the Gaussian distortion setting described directly preceding (17), then for large enough $n$ the value $H(U_S|U_{SC}, Y_0)$ equals $I(Y_S; Z_S|Z_{SC}(\rho), Y_0)$. Maximizing over the LAN constraint (Equation (4)), the following rate is achievable:
\[
R_{DC}(L) \triangleq \max_{r \in R_{LAN}(L)} \max_{\rho \in R_{DC}(r)} R_G(\rho) \quad (28)
\]

where
\[
\mathcal{R}_{DC}(r) = \left\{ \rho : \forall S \subseteq \{1, \ldots, N\}, \quad I(Y_S; Z_S|Z_{SC}(\rho), Y_0) < \sum_{n \in S} r_n \right\}. \quad (29)
\]

$Z(\cdot)$ in (29) is defined as in (19). We can then state the following:

**Theorem 2.** For a distributed receive system as described in Section III with noise covariance matrix $\Sigma$ and LAN constraint $L$ then a rate $R_{DC}(L)$ is achievable.

The base can unambiguously decompress the helper’s compressed encodings with low probability of error if and only if $\rho$ is chosen such that (27) is satisfied. However, in analogue to Corollary 1 from Reference [14], the rate in 2 can be improved by expanding $R_{DC}(r)$ to include some of the $\rho$ where the base cannot perform unambiguous decompression. This helps because even if the encoding rates in $\rho$ are chosen outside of $\mathcal{R}_{DC}(r)$ so that helpers cannot convey $U$ to the base unambiguously, some extra correlation with $X$ is retained through this distortion.

**Theorem 3.** For a distributed receive system as described in Section III with noise covariance matrix $\Sigma$ and LAN constraint $L$ then the following rate is achievable:
\[
R_{DC}(L) \triangleq \max_{r \in R_{LAN}(L), \lambda \in R_{DC}(r)} \max_{\rho \in R_{DC}(\lambda)} R_G(\rho) - \lambda \quad (30)
\]

where
\[
\mathcal{R}_{DC}(\lambda) = \left\{ \rho : \forall S \subseteq \{1, \ldots, N\}, \quad I(Y_S; Z_S|Z_{SC}(\rho), Y_0) < \sum_{n \in S} r_n + \lambda \right\}. \quad (31)
\]

Proofs of Theorems 2 and 3 are shown in Appendix A.

**Remark 6.** (Due to Reference [14]) The capacity of the system is $R_{DC}(L)$ under the following restrictions:

- $\Sigma$ is diagonal (no interference).
- The base does not have its own full-precision observation of the broadcast ($h_0 = 0$)
function of:

- The broadcaster must transmit a Gaussian signal
- Helper messages are independent of the transmitter's codebook

This is demonstrated in Appendix B. We suspect that Remark 6 holds without the first assumption, but this has not been proven. The last three assumptions are necessary:

- Since the base has full code knowledge, it is possible for the transmitter to send a direct message to the base (possibly with dirty paper coding), which is not accounted for in the compress-and-forward strategy used for $R_{\text{PC}}(L)$.
- The Gaussian broadcast assumption is needed because of a counterexample given in Reference [14].
- Codebook independence is necessary because $R_{\text{PC}}(L)$ is strictly less than the upper bound in (12), but by Remark 3 this upper bound is achieved in some regimes.

Theorem 3 and Remark 6 strengthen Corollary 1 and Theorem 5 from Reference [14] since we show that the same rate can be achieved with less cooperation between transmitter and receivers. Reference [14] uses a ‘nomadicity’ assumption which asserts that the mapping from transmitter messages to codewords is not present at the helpers, whereas Theorem 3 shows (the same argument applying to the general discrete case) that indeed the same rate is achievable when helpers have no knowledge at all of the codewords the transmitter is using.

Like the Gaussian distortion achievable rate, the distributed compression technique does not require cooperation between the transmitter and helpers. It does, however require a priori sharing of codes from the helpers to the base so that the helpers can perform distributed compression.

V. ERGODIC BOUNDS

In this section, the bounds are averaged over random channels in various regimes. Each bound tested is a deterministic function of:

- Number of helpers $N$
- LAN constraint $L$
- Noise covariance matrix $\Sigma$

- Channel $h$ (assumed to be static and precisely estimated a priori per each channel use).

For each channel drawn, the bounds were optimized over this parameter space using Artelys KNITRO a fast nonlinear optimization tool (Reference [25]).

In all graphs, the rate from (12) is called ‘Upper Bound,’ the rate from (22) is ‘Gaussian Distortion’ and the rate from Theorem 3 is ‘Distributed Compression.’ Bounds from (28) and Remark 2 are not plotted.

A. Performance in the Presence of an Interferer

One of the benefits of receiver diversity is increased ability to compensate for degradation caused by interferers. In this section we investigate the extent to which interference mitigation is possible through distributed reception with a single interferer occluding all the nodes. We assume the transmitter has no knowledge of the interferer. This necessarily invalidates the use of dirty paper coding (Reference [19]) and other transmit-side interference mitigation strategies. We model an interferer as additive Gaussian noise term present at each of the receive nodes and independent of all other aspects of the system, where the scale of the term at each receiver varies depending on the interferer’s presence there. With this model, the covariance matrix of noise terms associated with a particular interferer seen by the receivers is a matrix $A = aa^\dagger$, where $a \in \mathbb{C}^{N+1}$. In all trials run, a single interferer was assumed to be present at all nodes with power $\alpha$ and random phase so that

$$\Sigma = I_{N+1 \times N+1} + \alpha \cdot aa^\dagger.$$  \hspace{1cm} (32)

with $a_n = e^{j2\pi \theta_n}$, $\theta_n \sim \text{Unif}[0, 1]$ where Unif is the uniform distribution.

Figure 4 shows that even when a strong rank-one interferer is present at all the nodes, achievable communications rates are comparable to the case without an interferer. With or without an observation at the base and at low or high interference, above $L = 25$ bits shared across the helpers, the achievable rates studied become close to the upper bound. In all regimes tested the Gaussian distortion bound is close in
performance to the more complicated distributed compression bound especially when the base has its own observation.

B. Path Diversity versus Performance

It is reasonable to expect that for the best performance, distributed receivers should adjust their operation depending on their scattering environment. If the receivers observe the signal through a single line-of-sight path, all else being equal, the noise on each receiver will have similar statistics. In contrast, in multipath environments the level of signal attenuation may vary more across receivers and in turn so will the noise statistics.

The level of attenuation at a receiver is modeled here as being Rician distributed over communication periods, as is done in other literature (see Reference [26]). The Rician distribution follows the magnitude of a circularly-symmetric complex Gaussian with nonzero mean and can be parameterized by two nonzero values: a scale \( \Omega > 0 \) representing the average receive SNR and a shape \( K > 0 \) (called a K-factor in other literature) denoting the ratio of signal power received from direct paths to the amount of power received from scattered paths. By construction, if \( K = 0 \), then a Rician distribution is equivalent to a Rayleigh distribution with mean \( \Omega \). In contrast if \( K \to \infty \) then the distribution approaches a point mass at \( \Omega \) (Reference [26]).

All data sets generated showed that the type of scattering environment is uncorrelated with the average rate achievable by the bounds (Figure 6). Despite the invariance of the bounds to \( K \), the distribution of the amount of bits each helper sends to the base does change. Figure 5 shows that no matter other parameters, in high scattering it is most helpful for the base to draw the majority of its information from the highest SNR helper, almost completely ignoring low-SNR helpers. The sharing of LAN capacity amongst helpers even out as \( K \) increases and the SNR becomes less varied across receivers.

When the base has its own observation, the average distribution of bits on helpers is more or less identical at -20 and 20 dB INR at any scattering environment. However when the base does not have its own observation, at low \( K \)-factor the base goes from not using the lowest SNR helpers at all at -20 dB INR to drawing a reasonable amount of information from them at 20 dB INR. The authors hypothesize that in practice these helper messages would go towards providing an estimate of the interferer, and that when the base has its own full precision observation available, that is used as an interference estimate instead and listening to low SNR helpers is less beneficial. However, we note that the situation may change if the base node can be selected from all the receive nodes after the channel is estimated.

VI. CONCLUSION

In total, one upper bound and three lower bounds on capacity were shown. A simple upper bound (Equation (12)) was derived, achievable when the LAN constraint is stringent (small \( L \), and achievable in limit as the LAN constraint is relaxed (\( L \to \infty \)). An achievable rate using dirty paper coding is shown (Equation (13)). Another achievable rate (Equation (22)) was found by considering quantizers which add Gaussian distortion. It alleviates the requirement of coordination between the transmitter and helpers. A third lower bound (Equation (28)) was derived by using a distributed compression technique on the helper observations from the bound in (22). It outperforms the bound in (22), but requires intensive a priori coordination between helpers.

It was seen from plots in Section V that distributed compression does not drastically change performance unless there are many helpers, and that leaving SNR fixed, performance is mostly unchanged in high-scattering versus low-scattering environments. In high-scattering environments some helpers’ observations are virtually ignored by the base, while in line-of-sight environments, helpers inform the base roughly equally, but performance is the same. Adding interference diminishes the improvement gained by having broadcaster cooperation.

It remains to extend these results to situations where any of the receive nodes can be selected as the base, and situations where all nodes wish to receive the broadcast. Practical implementations of a bound-achieving system are also pending investigation.
APPENDIX A

ACHIEVABILITY OF BOUNDS

Here we prove theorems 1, 2 and 3. First some lemmas are needed.

Lemma 1. For a random series of helper observations $Y^t$, and lattice encodings $U^t$ with distortion vectors $W_{Q,n}^t$ as constructed in the Helper encoder setup subsections of the proofs for Theorems 1 or 3, where $\text{Var}(W_{Q,n}^t) = \sigma_n^2 \ell$, for any $S \subseteq \{1, \ldots, N\}$ then for any $\epsilon > 0$, there is some blocklength $t$ large enough that

$$\frac{1}{t} H(U_n^t) - \frac{1}{t} I(Y_n^t; Z_n^t(r)) < \epsilon$$

(33)

where $Z(r) = Y^t + W^t$, with $W^t$ $Y^t$ and

$$W^t \sim \mathcal{CN}(0, \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)).$$

(34)

Proof: We prove this by induction. By Theorem 1 in Reference [27], for any $n = 1, \ldots, N$

$$H(U_n^t) = I(Y_n^t; Y_n^t - W_{Q,n}^t).$$

(35)

Further, by Theorem 3 in Reference [22]

$$\frac{1}{t} I(Y_n^t; Y_n^t - W_{Q,n}^t) \rightarrow I(Y_n; Y_n + W_n) = I(Y_n; Z_n(r_n)).$$

(36)

So the statement holds for $\{n\}$ and arbitrarily small $\epsilon' > 0$. The statement is also vacuously true for $\emptyset$, so it holds for $\emptyset \cup \{n\}$.

Now assume there are two disjoint sets $A, B$ for which the claim holds for arbitrarily small $\epsilon' > 0$. Then there is a deterministic function at the decoder $f_{A UB}$ where $f_{A UB}(U_{A UB}^t) = Y_{A UB}^t - W_{Q,AB}^t$, so

$$H(U_{A UB}^t) \geq I(f(U_{A UB}^t); Y_{A UB}^t)$$

(37)

$$\geq I(Y_{A UB}^t; Z_{A UB}^t(r)).$$

(38)

The first line follows by the data processing inequality and the second follows by the property that for fixed variance, the Gaussian distribution minimizes mutual information between a random variable and its version with additive independent noise (Reference [22]). Similarly,

$$H(U_{A UB}^t) = H(U_A^t) + H(U_B^t) - I(U_A^t; U_B^t)$$

(39)

$$\leq I(Y_A^t; Z_A^t(r)) + I(Y_B^t; Z_B^t(r)) - \epsilon.$$ (40)

By construction, at the decoder there are deterministic functions $f_A, f_B$ where $f_A(U_A^t) = Y_A^t - W_{Q,A}^t$, and $f_B(U_B^t) = Y_B^t - W_{Q,B}^t$ so

$$I(U_A^t; U_B^t) \geq I(f(AUB^t); f(BUB^t))$$

(41)

$$\geq I(Z_A^t(r); Z_B^t(r)) - \epsilon'.$$

(42)

where the first line follows by using the data processing inequality twice, and the second follows by the same Gaussian property used above. Combining (38), (40) and (42) then

$$0 \leq H(U_{A UB}^t) - I(Y_{A UB}^t; Z_{A UB}^t(r)) \leq 3\epsilon'.$$

(43)

Since $S \subseteq \{1, \ldots, N\}$, the inductive step will only need to be used up to $N$ times. Letting $0 < \epsilon' < \epsilon \cdot 3^{-N}$, the initial statement holds.

A small result is also needed to show that the base having its own full precision observation is approximately equivalent to the base not having its own observation, but also receiving a high-bitrate helper message. This facilitates code construction in the theorems, because it handles the asymmetry of including a full-precision observation at the decoder.

Lemma 2. If $\forall S \subseteq \{1, \ldots, N\}$ then

$$I(Y_S; Z_S(r)|Z_{SC}(r), Y_0) < \sum_{m \in S} r_m.$$ (44)

then some value $r_0$ can be chosen sufficiently large so that for any $S \subseteq \{1, \ldots, N\}$, both

$$I(Y_S; Z_S(r), Z_0(r_0)|Z_{SC}(r)) < \sum_{m \in S} r_m + r_0.$$ (45)

and

$$I(Y_S; Z_S(r)|Z_{SC}(r), Z_0(r_0)) < \sum_{m \in S} r_m.$$ (46)

Proof: Since all the conditional mutual information terms in the statement are continuous in $r$, it is enough to show the statement for $r \to 0$. In this case, with no loss of generality we can say there is some $\delta > 0$ where:

$$I(Y_S; Z_S(r)|Z_{SC}(r), Y_0) < \sum_{m \in S} r_m - \delta.$$ (47)

Fix $r_0$ large enough so that for any $S \subseteq \{1, \ldots, N\}$,

$$\|I(Y_S; Z_S(r)|Z_{SC}(r), Y_0) - \ldots - I(Y_S; Z_S(r)|Z_{SC}(r), Z_0(r_0))\| \leq \delta.$$ (48)

There is guaranteed to be such an $r_0$ because the above expression is a continuous function of the covariance matrix of $(Y_0, Y, Z_0(r_0), Z(r))$ component-wise, and as $r_0$ is made large, the covariance matrices involved in the second term converge component-wise to those of the first since $\|Q_n(r)\| \to 0$ as $r_0 \to \infty$. By (47) and (48) the statement holds for any $S \subseteq \{1, \ldots, N\}$, and it remains to show that it also holds for each $S \cup \{0\}$.

Note that for any $S \subseteq \{1, \ldots, N\}$

$$I(Y_S^t, Y_0^t; Z_S^t(r), Z_0^t(r)|Z_{SC}^t(r))$$

(49)

$$= I(Y_S^t; Z_S^t(r)|Z_{SC}^t(r)) + I(Y_0^t; Z_0^t(r)|Z_{SC}^t(r))$$

(50)

$$\leq I(Y_S^t; Z_S^t(r)|Z_{SC}^t(r)) + r_0$$

(51)

$$\leq \sum_{m \in S} r_m + r_0$$

(52)

where (51) follows because condition reduces mutual information, and (52) follows from what was shown previously. Thus, the statement holds.

Theorem 1. For a distributed receive system as described in Section III with noise covariance matrix $\Sigma$, LAN constraint $L$ and fixed average helper quantization rates $r \in R_{LAN}(L)$, then a rate $R_{Q}(r)$ is achievable.
Proof: If for some $n \in \{1, \ldots, N\}$ then $r_n = 0$, the system is equivalent to the case where the $n$th helper is not present, so without loss of generality assert that $r_n > \varepsilon$ for $n = 1, \ldots, N$. Fix some rate $R < R_G(r)$ and a block length $T = t^2 \in \mathbb{N}$.

**Operation:**

**Transmitter setup:** Generate a codebook $\mathcal{X} = \{X_1, \ldots, X_{2^R}\} \subseteq \mathbb{C}^{t \times 1}$ where all the vectors' components are drawn iid from $\mathcal{CN}(0, 1)$. Distribute $\mathcal{X}$ to the transmitter and base.

**Helper encoder setup:** For each helper $n$, $n \in \{1, \ldots, N\}$, generate $t$ distortion vectors \{$W_{Q,n}^t\}_{l=1}^t \subseteq \mathbb{C}^{t \times 1}$, each independent and uniform in the base region of the Voronoi partition $\mathcal{P}_n$ of a regular, white $n$-dimensional lattice $\mathcal{L}_n$ scaled to have a normalized-second-moment:

$$G_l^l(\mathcal{L}_n, \mathcal{P}_n) = \frac{\|h_n\|_2^2 + \Sigma_{n,n}}{2^n - \varepsilon - 1}. \tag{53}$$

These terms are detailed at the beginning of Reference [22]. Share \{$W_{Q,n}^t\}_{n=1}^N$ with the corresponding helper and the base.

At the base do the same, generating $t$ distortion vectors \{$\tilde{W}_{Q,n}^t\}_{n=1}^N \subseteq \mathbb{C}^{t \times 1}$ with $r_0$ chosen large enough so that the statement in Lemma 2 holds. These are generated for the purpose of the base quantizing its own receptions. The base does not need to quantize, but doing so simplifies the proof.

**Transmission:** To send a message $M = (m_1, \ldots, m_t) \in \prod_{t=1}^t \{1, \ldots, 2^{RT}\}$, have the transmitter broadcast:

$$X = (X_{m_1}, \ldots, X_{m_t}) \in \mathcal{X}^t \subseteq \mathbb{C}^{T \times 1}. \tag{54}$$

**Helper encoding and forwarding:** For the transmission period, receiver $n$ ($n \in \{0, \ldots, N\}$) observes a sequence of length $T$ which we split up into $t$ sequences of length $t$:

$$Y_n^t \triangleq (Y_{n,t-t(t-1)+1}, \ldots, Y_{n,t-t(t-1)+t}) \in \mathbb{C}^{t \times 1}, \quad t \in \{1, \ldots, t\} \tag{55}$$

Form a set of quantizations \{$U_n^t\}_{t=1}^t \subseteq \mathcal{L}_n^t$ by finding the point in $\mathcal{L}_n$ corresponding to the region in $\mathcal{P}_n$ in which $Y_n^t - W_{Q,n}^t$ resides. The properties of such $U_n^t$ are the subject of References [27], [22]. By Theorem 1 in Reference [27],

$$H(U_n^t) = I(Y_n^t; Y_n^t - W_{Q,n}^t). \tag{56}$$

Further, by Theorem 3 in Reference [22]

$$\frac{1}{t} I(Y_n^t; Y_n^t - W_{Q,n}^t) \rightarrow_I I(Y_n; Y_n + W_n) = r_n - \varepsilon \tag{57}$$

where $W_n \sim \mathcal{CN}(0, G_n^t(\mathcal{L}_n^t, \mathcal{P}_n^t))$ (in agreement with notation in Section IV-B). Thus the encoded messages \{$U_n^t\}_{t=1}^t$ are within the LAN constraint for large enough blocklength $T$. Forward \{$U_n^t\}_{t=1}^t$ to the base.

**Decoding:** Take $A^t_r(X, U)$ to be the set of

$$((x^1, \ldots, x^t), (u^1, \ldots, u^t)) \in \mathcal{X}^t \times \prod_{n=0}^N \mathcal{L}_n^t \tag{58}$$

which are jointly-$\varepsilon$-weakly-typical with respect to the joint distribution of:

$$(X, (U^1, \ldots, U^t)) \tag{59}$$

where $U^t = (U^t_0, \ldots, U^t_N)$.

Weak- and joint-typicality are defined in Reference [2].

At the base, find $\hat{x} = (\hat{x}^1, \ldots, \hat{x}^t) \in \mathcal{X}^t$ where $(\hat{x}, (U^1, \ldots, U^t)) \in A^t_r(X, U)$. Declare error events $E_0$ if $X$ is not found to be typical with $(U^1, \ldots, U^t)$, and $E_1$ if there is some $\hat{x} \in \mathcal{X}^t$ where $\hat{x} \neq X$ and $(\hat{x}, (U^1, \ldots, U^t)) \in A^t_r(X, U)$.

**Error analysis:** By typicality and the law of large numbers, $P(E_0) \rightarrow 0$ as $t \rightarrow \infty$. Also,

$$P(E_1) \leq \sum_{\hat{x} \in \mathcal{X}^t, \hat{x} \neq X} P\left(\{(u_0^1, u_1^1, \ldots, u_N^1) : \hat{x} \in A^t_r(X, U)\}\right) \tag{61}$$

$$\leq \sum_{\hat{x} \in \mathcal{X}^t, \hat{x} \neq X} 2^{-t(I(X; U^1, \ldots, U^t) - 3\varepsilon)} \tag{62}$$

$$< 2^{-t(I(X; U^1, \ldots, U^t) - 3\varepsilon - tR)} \tag{63}$$

$$< 2^{-t(t(I(X; (r-1)\varepsilon)) - 3\varepsilon - tR)} \tag{64}$$

Equation (64) follows from Equation (63) because both have the same mean and covariance matrix, and the Gaussian distribution maximizes entropy under these constraints (Reference [28]). So if $R$ is chosen less than $I(X; Z(r - 1 \cdot \varepsilon)) - 3\varepsilon$ then $P(E_0 \cup E_1) \rightarrow 0$ as $T \rightarrow \infty$. For small enough $\varepsilon$, by lower semi-continuity of mutual information $I(x; Z(r - 1 \cdot \varepsilon)) - 3\varepsilon$ can be made arbitrarily close to $R_G(r)$.

Roughly the same strategy is used in the demonstration of achievability of the distributed compression system discussed in Section IV-C up to the helper’s operation, where a joint-compression stage is added.

**Theorem 2.** For a distributed receive system as described in Section III with noise covariance matrix $\Sigma$ and LAN constraint $L$ then a rate $R_{DC(L)}$ is achievable.

Proof: Apply Theorem 3 with $\lambda = 0$ (proof shown below).

**Theorem 3.** For a distributed receive system as described in Section III with noise covariance matrix $\Sigma$ and LAN constraint $L$ then $R_{DC(L)}$ is achievable.

**Proof:** It is enough to show that any component in the maximization from (30) is achievable. Fix $\lambda \in \mathbb{R}$, a helper rate vector $r \in \mathcal{R}_{LAN}(L)$, and a compression rate vector $\rho \in \mathcal{R}_{DC}(r)$. If for some $n \in \{1, \ldots, N\}$ then $r_n = 0$ or $\rho_n = 0$, the system is equivalent to the case where the $n$th helper is not present, so without loss of generality assert that $r_n, \rho_n > \varepsilon$ for $n = 1, \ldots, N$. Fix some rate $R < R_G(r - 1 \cdot \varepsilon) - \lambda$ and a block length $T = t^2 \in \mathbb{N}$.

**Operation:**

**Transmitter setup:** Generate a codebook $\mathcal{X} = \{X_1, \ldots, X_{2^R}\} \subseteq \mathbb{C}^{t \times 1}$ where all the vectors' components are drawn iid from $\mathcal{CN}(0, 1)$. Distribute $\mathcal{X}$ to the transmitter and base.

Proof: Apply Theorem 3 with $\lambda = 0$ (proof shown below).
Helper encoder setup: For each helper $n$, $n \in \{1, \ldots, N\}$, generate $t$ distortion vectors $\{\hat{W}_{Q_n}^t\}_{\ell = 1}^t \subseteq \mathbb{C}^{t \times 1}$, each independent and uniform in the base region of the Voronoi partition $\mathcal{P}_n$ of a regular, white $n$-dimensional lattice $\mathcal{L}_n$ scaled to have a normalized second-moment:

$$G^t_n(\mathcal{L}_n; \mathcal{P}_n) = \frac{\|h_n\|^2 + \sum_{n=1}^N \sigma_{n,n}}{2 \rho_n - \varepsilon - 1}. \quad (65)$$

These terms are detailed at the beginning of Reference [22]. Share $\{\hat{W}_{Q_n}^t\}_{\ell = 1}^t$ with the corresponding helper and the base.

At the base do the same, generating $t$ distortion vectors $\{\hat{W}_{Q_0}^t\}_{\ell = 1}^t \subseteq \mathbb{C}^{t \times 1}$ with $\rho_0$ chosen large enough so that the statement in Lemma 2 holds. These are generated for the purpose of the base quantizing its own receptions. The base does not need to quantize, but doing so simplifies the proof.

At receiver $n$, $n \in \{0, 1, \ldots, N\}$, independently choose a random mapping, $\text{Index}_n: \mathcal{U}_n \rightarrow \{1, \ldots, 2^{tr_n}\}$:

$$\text{Index}_n \sim \text{Unif} \left( \{ \phi: \phi: \mathcal{U}_n \rightarrow \{1, \ldots, 2^{tr_n}\} \} \right) \quad (66)$$

where $\text{Unif}(\cdot)$ is the uniform distribution on its set, and $\mathcal{U}_n$ is the finite alphabet of a particular random variable $U_n$ which will be described in the Helper encoding truncation stage. $\text{Index}_n$ represents the binning scheme used by receiver $n$.

Have each receiver distribute its chosen $\text{Index}_n$ to the base.

Transmission: To send a message $\mathcal{M} = (m_1, \ldots, m_t) \in \prod_{t=1}^t \{1, \ldots, 2^{tR}\}$, have the transmitter broadcast:

$$\mathcal{X} = (X_1, \ldots, X_t) \in \mathcal{X}^t \subseteq \mathbb{C}^{t \times 1}. \quad (67)$$

Helper encoding: For one transmission period, receiver $n$ ($n \in \{0, \ldots, N\}$) observes a sequence of length $t$ which we split up into $t$ sequences of length $t$:

$$Y_n^t \triangleq (Y_{n,t(\ell-1)+1}, \ldots, Y_{n,t\ell(t-1)+t}) \in \mathbb{C}^{t \times 1}, \quad \ell \in \{1, \ldots, t\} \quad (68)$$

Form a set of quantizations $\{\hat{U}_n^t\}_{\ell = 1}^t \subseteq \mathcal{L}_n^t$ by finding the point in $\mathcal{L}_n^t$ corresponding to the region in $\mathcal{P}_n^t$ in which $Y_n^t$ lies. The properties of such $\hat{U}_n^t$ are the subject of References [27], [22]. By Lemma 1 and the chain rule for entropy and mutual information, then for large enough $t$ for any $S \subseteq \{0, \ldots, N\}$,

$$\frac{1}{t} H(\hat{U}_n^t|U_n^t) \rightarrow I(Z_S(\rho); Y_S Z_{SC}(\rho)) < \sum_{n \in S} \frac{r_n}{n} \quad (69)$$

where convergence follows from (35) and Theorem 3.1 in Reference [29], and the inequality comes from choice of $\rho$.

Helper encoding truncation: We now face a small problem: we wish to convey $\{\hat{U}_n^t\}_{\ell = 1}^t$, $n = 1, \ldots, N$ to the base, but to do so requires joint compression. Binning for joint compression depends on the presumption that variables to be compressed have finite alphabets, but the $U_n^t$ are distributed on a countably infinite lattice. In this step, appropriate finite alphabet variables $\hat{U}_n^t \in \mathcal{U}_n$ are constructed.

Take $A$ to be any sub-vector of $(X_{m_1}, U_0^t, \ldots, U_N^t)$. The entropy of $A$ is an absolutely convergent sum of positive numbers:

$$H(A) = \sum_{a \in \text{Dom}(A)} -\log(P_A(a)) P_A(a) < \infty. \quad (70)$$

Then for each $\ell \in \{1, \ldots, t\}$ and $n \in \{0, \ldots, N\}$, by definition of countable summation, nonempty finite sets

$$\mathcal{U}_0 \subseteq \mathcal{L}_0, \ldots, \mathcal{U}_N \subseteq \mathcal{L}_N, \quad (71)$$

can be chosen so that that the random variables:

$$U_n^t \triangleq \{ \hat{U}_n^t, \cdots, \hat{U}_n^t \} \quad n \in \{0, \ldots, N\} \quad (72)$$

satisfy (for any $S \subseteq \{0, \ldots, N\}$, $b \in \{0, 1\}$, $\varepsilon' > 0$)

$$0 < H(b \cdot X_m, \hat{U}_S^t) - H(b \cdot X_m, \hat{U}_S^t) < \varepsilon'. \quad (73)$$

By forming linear combinations of appropriate forms of Equation 73,

$$\|H(U_S^t|U_{SC}) - H(U_S^t|U_{SC}^t)\| < \varepsilon \quad (74)$$

so by Equation 69 for large enough $t$ and sufficiently large $\mathcal{U}_n$s,

$$\frac{1}{t} H(U_S^t|U_{SC}^t) < \sum_{n \in S} \frac{r_n}{n} \quad (75)$$

By the same reasoning on the sum $\sum_n P_A(a)$, assure $\mathcal{U}_n$s are also large enough that

$$P(U_n^t = \infty) \leq \frac{1}{(N + 1) t^2}. \quad (76)$$

Helper joint-compression and forwarding: At receiver $n$ for $n = 0, 1, \ldots, N$, form random variables

$$V_n^t \triangleq \text{Index}_n^t(U_n^t), \ell \in \{1, \ldots, t\}, \quad (77)$$

where

$$\text{Index}_n^t(U_n^t) = \begin{cases} 1 & \text{Index}_n(U_n^t) = \infty \\ \text{Index}_n(U_n^t) & \text{otherwise} \end{cases} \quad (78)$$

At each helper forward $\{V_n^t\}_{\ell = 1}^t$ to the base. Note that because $|\text{Range}(\text{Index}_n^t)| \leq 2^{tr_n}$, $H(V_n^t) \leq tr_n$ so each $V_n^t$ can be sent within the LAN constraint. Denote $\mathbf{V} \triangleq (V_0, V_1, \ldots, V_n)$.

Decoding: Take $A^t_n(\mathbf{X}, \mathbf{U})$ to be the set of

$$((x^t, \ldots, x^t), (u^t, \ldots, u^t)) \in \mathcal{X}^t \times \left(\prod_{n=0}^N \mathcal{U}_n\right)^t \quad (79)$$

which are jointly-weakly-typical with respect to the joint distribution of:

$$(X_n^t, U_0^t, \ldots, U_n^t) \quad (80)$$

where

$$U^t_n = (U_0^t, \ldots, U_n^t). \quad (81)$$

Weak- and joint-typicality are defined in Reference [2].

Define a family of sets $\mathcal{B}$ indexed by vectors

$$\nu = (\nu_0^t, \ldots, \nu_N^t) \in \mathcal{V} \triangleq \left(\prod_{n=0}^N \text{Range}(	ext{Index}_n)\right)^t \quad (82)$$

with $\nu^t = (\nu_0^t, \ldots, \nu_N^t) \in \prod_{n=0}^N \text{Range}(\text{Index}_n)$ where $\mathcal{B} \triangleq \{B_\nu|\nu \in \mathcal{V}\}$ and

$$B_\nu \triangleq \{u|\text{Index}_n^t(u_n^t) = \nu_n^t\} \quad (83)$$

forall $n \in \{0, \ldots, N\}, \ell \in \{1, \ldots, t\}$ \subseteq \left(\prod_{n=0}^N \mathcal{U}_n\right)^t \quad (83)
Each $B_v$ is the set of helper encodings represented by the compressed messages $v$.

At the base, find $X \in \mathcal{X}$ for which there is some $\hat{U} \in B_V$ where $(X, \hat{U}) \in A^T(X, U)$. Declare the message associated with $X$ to be the broadcast.

**Error analysis:**

We have the following error events:

- $E_{\text{Trunc}, t, n}; U^t_t \rightarrow \infty$
- $E_{JT}; X$ is not typical with any $u \in B_V$
- $E_{\hat{m}, S};$ For $S \subseteq \{0, 1, \ldots, N\}$ and $\hat{m} \neq M$, then there is some $u = (\hat{u}, u_{SC}) \in B_V$ where $\forall n \in S \hat{u}_n = U_n$, $\forall n \notin S \hat{u}_n \neq U_n$, and $(\hat{x}_n, \hat{u}) \in A^T(X, U)$. $E_{\hat{m}, S}$ denotes the situation where the system cannot uniquely decompress encodings from receivers $S$, causing the broadcast message to look like $\hat{m}$.

Take $E_{\text{Trunc}} = \cup_{t=1}^{\infty} \cup_{n=0}^{N} E_{\text{Trunc}, t, n}$ By choice of alphabets in Equation (76),

$$P(E_{\text{Trunc}}) \leq \sum_{n,t} P(E_{\text{Trunc}, t, n}) \leq \frac{N t}{t^2} \rightarrow 0 \text{ as } t \rightarrow \infty. \quad (85)$$

By the strong law of large numbers and construction of the typical set, as $t$ becomes large, eventually $P(E_{JT}) < \varepsilon$.

Without loss of generality assume $M = 1$. Take $A^T(U)$ to be the collection of jointly-typical sequences in $\left(\prod_{n=1}^{N} U_n\right)^t$ up to the joint distribution on each element in the vector $U = (U^1, \ldots, U^N)$. For each $S \subseteq \{0, 1, \ldots, N\}$,

$$\sum_{\hat{m} \neq 1} P(E_{\hat{m}, S} \cap E_{\text{Trunc}} \cap E_{JT}) \leq 2^{TR} \sum_{\hat{u} \in B_V \cap A^T(U)} P\left(\hat{X}, \hat{u} \in A^T(X, U)\right) \quad (86)$$

where $\hat{X} \sim \mathcal{CN}(0, 1)^T$, $\hat{X} \perp U$. Then:

$$\sum_{\hat{m} \neq 1} P(E_{\hat{m}, S} \cap E_{\text{Trunc}} \cap E_{JT}) \leq \ldots \quad (87)$$

Taking the log,

$$\log \left(\sum_{\hat{m} \neq 1} P(E_{\hat{m}, S} \cap E_{\text{Trunc}} \cap E_{JT})\right) \leq \ldots \quad (88)$$

Since $R$ was chosen so that $R \leq R_G(\rho - 1 \cdot \varepsilon) - \lambda = I(X; Z(\rho - 1 \cdot \varepsilon)) - \lambda$, then (90) approaches $-\infty$ as $t \rightarrow \infty$ (Thereby the left side of Equation (86) approaches 0) if for any $S \subseteq \{0, \ldots, N\}$ then:

$$I(Y_{SC}; Z_S(\rho - 1 \cdot \varepsilon)|Z_{SC}(\rho - 1 \cdot \varepsilon)) < \lambda + \sum_{m \in S} r_m - 7\varepsilon. \quad (91)$$

By assumption that $\rho \in R_G^\lambda$, and for small enough $\varepsilon$, then (91) holds for each $S$. Since all error events approach 0, then a rate of $R_G(\rho - 1 \cdot \varepsilon) - \lambda$ is achievable. For small enough $\varepsilon$, by lower semi-continuity of mutual information $R_G(\rho - 1 \cdot \varepsilon) - \lambda$ can be made arbitrarily close to $R_G(\rho) - \lambda$.

**APPENDIX B**

**PROOF OF CONDITIONAL CAPACITY**

**Remark 6.** (Due to Reference [14]) The capacity of the system is $R_{\text{FR}}^L(L)$ under the following restrictions:

- $\Sigma$ is diagonal (no interference).
- The base does not have its own full-precision observation of the broadcast ($I_0 = 0$)
- The broadcaster must transmit a Gaussian signal
- Helper messages are independent of the transmitter’s codebook $\mathcal{X}$.

**Proof:** By Theorem 5 in Reference [14], the capacity of the system under the assumed restrictions is:

$$\max_{V \in \mathcal{V}} C_{r, V} \quad (92)$$

$$C_{r, V} \triangleq \min_{S \subseteq \{1, \ldots, N\}} \left\{ \sum_{m \in S} [r_m - I(V_m; Y_m|X)] + I(V_{SC}; X) \right\}. \quad (93)$$

In (93), $V$ is the collection of random vectors $V = (V_1, \ldots, V_N)$ whose components are of the form $V_n = Y_n + W_n$ with independently distributed $W_n$:

$$\hat{W}_n \sim \mathcal{CN}\left(0, \sigma^2_n \frac{2^{-2v_n}}{1 - 2^{-2v_n}}\right) \quad (94)$$

for any $v_n \geq 0$. Since both (94) as a function of $v_n$ and (20) as a function of $r_n$ are injective on $(0, \infty)$, then for any $V \in \mathcal{V}$ there is some $\rho$ (in the context of Equation (31); possibly either inside or outside $\cup_{\lambda \in R_G^\lambda} R_{\text{DC}}^L(L)$) which will yield a variable $Z(\rho)$ with identical distribution to $V$.

Fixing helper rates $r$, for any $S \subseteq \{1, \ldots, N\}$, we can instead write $C_{r, V}$ as:

$$C_{r, V} = \max \left\{ c : \forall S \subseteq \{1, \ldots, N\} \right\} \quad (95)$$

$$c \leq I(V_{SC}; X) + \sum_{m \in S} r_m - I(V_m; Y_m|X)$$

1 Variable names have been altered from Reference [14], and constants have been adapted for complex variables, but the form is the same.
its own observation.

is identical to the set of rates shown in Theorem 3.

\[ \forall n, V_n \text{ given } (X, Y_n) \text{ is conditionally independent of } V_{j \neq n} \text{ and } Y_{j \neq n}. \]

maximizing over \( r \) this is identical to the set of rates shown in Theorem 3. \( \blacksquare \)

ACKNOWLEDGMENTS

We would like to thank Shankar Ragi for aid in producing data sets for this paper. We would also like to thank Bryan Paul for his comments and corrections.

Support by Hans Mittelmann offered by AFOSR under grant FA 9550-15-1-0351

REFERENCES