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

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

***Abstract*—We develop communications rate bounds for a single transmitter and a distributed array of receivers under the assumption of limited resources for a secondary local array network (LAN) that enables communications between the *helper* nodes of the distributed receive array and a *base* (fusion) node. We consider the increasingly important case of interference that causes spatially correlated noise. The bounds demonstrate that such systems have a moderate sensitivity to the effects of interference for well designed approaches. To improve usefulness to system designers, we develop multiple bounds under different assumed capabilities of the distributed reception system. We also compare scenarios in which the fusion node either does or does not have its own observation in addition to the helper's side information. We investigate spatially Ricean channels that vary from line-of-sight to Gaussian. Interestingly, 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, the distributed reception strategies for achieving these rates must change in response to the channel. The bounds are tight when the LAN allows for either large or small helpers-to-base communications rates.**

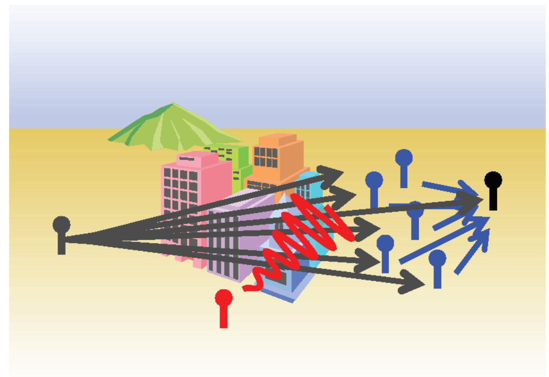


Fig. 1. A broadcast node speaks to many helper nodes, which forward limited amounts of information to a base node. Additive interference from neighboring systems is present in the link from broadcaster to helpers.

I. Introduction

We are at a remarkable point in time for radio systems. With the significant reductions in cost and improvements in performance, there has been an explosion in the number of radio systems with a wide range of applications including personal, machine-to-machine, vehicle-to-vehicle, and internet-of-things (IoT) communications. In many situations, radios cluster in physically close groups. Consequently, there is an opportunity to use groups of independent radios as distributed arrays to improve communications performance as indicated in Figure 1. Because of increasing spectral occupancy and usage, the likelihood of interference is increased. Under the constraints of limited LAN communications, there is a concern about the effectiveness of this distributed reception approach in an interference environment. In this paper, we develop theoretical bounds for a single source to a base node with the aid of a distributed receive array of helper nodes in which we incorporate the effects of interference in these bounds.

Our bounds are constructed to aid system designers in developing parameters and approaches for distributed reception. We consider multiple bounds which correspond to different assumptions about the capability of the source compression of the distributed array. We consider the effects of an overall LAN capability that is parameterized by a total number of bits that helpers can share with the base for each source channel usage. Furthermore, to investigate algorithmic needs as a function of environmental conditions we consider a spatially Ricean channel and a parametrically defined interference. Finally, we consider both the conditions in which the base node does and does not have its own noisy observation of the source signal. Our bound development in this paper provides significant extensions to and clarifications of the results found in References [1], [2] and our preliminary efforts in Reference [3].

This work was sponsored in part by Defense Advanced Research Projects Agency under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

A. Contributions

We perform an analysis of the achievable rates of the system in the presence of an interferer. The following contributions are provided:

- Upper and lower bounds on the system's achievable communication rate in correlated Gaussian noise, and regimes where these bounds are tight (Section IV).
- Performance of the bounds in interference practical scenarios (Section V-A). We demonstrate that these systems are able to mitigate distortion from a strong interferer using a relatively unsophisticated strategy.
- Performance and behavior of the strategies in various scattering environments (Section V-B). The scattering environment is seen to not affect average performance.
- A strengthening of a previous achievability result for channels with uncorrelated Gaussian noise, where the same communication rate is achieved using less cooperation between nodes (Remark 6).

The following bounds are presented:

- A cut-set upper bound, tight in some regimes (beginning of Section IV)
- An achievable rate where the transmitter directly broadcasts to a helper (Section IV-A)
- An achievable rate where helpers individually compress their observations (Section IV-B)
- Achievable rates where helpers jointly compress their observations (Section IV-C)

II. Background

Broadcasting to distributed receivers has been studied in a variety of 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 done in the context of information theory towards finding achievable communications rates whereas most existing studies are in terms of minimizing bit error rates or distortion.

Results presented here are a significant extension and generalization of the work in [3], where there is no treatment of the system's performance in the presence of an interferer. A lower bound tighter than all those in [3] is presented in Theorem 3, and proofs are provided for all the results.

References [1], [2] consider the distributed receive problem, but their results do not apply directly to correlated Gaussian noise channels. The studies provide an example demonstrating the sub-optimality of a Gaussian broadcaster without interference, and suggest that non-Gaussian signalling techniques are needed for these topologies. In contrast, we demonstrate in Section V that an achievable rate using Gaussian signalling comes quite close to the system's upper bound on average over an ensemble of channels in many practical regimes.

The structure of the system considered here has an important difference from most relay channels in the literature. The link from transmitter to helpers can be seen as a source-to-relay link, and the LAN can be seen as a relay-to-destination link, but what distinguishes the present topology from ones in most relay studies is that here the LAN link itself is a graphical network with noisy inputs. This is not usual in the context of relay networks, including those studied in Reference [4]. 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], but do not provide general analytic expressions for rates as we do. 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. Results in our study build towards characterizing achievable rates of the system rather than designing and analyzing the performance of specific codes.

The topology we consider is similar but different in detail to many-help-one problems such as the 'Central Estimating Officer (CEO) problem' posed by Berger in Reference [11]. In the CEO problem, a CEO node seeks to estimate a source by listening to a set of 'agent' nodes which communicate to the base at fixed rates. Many variations of this problem have been studied, for instance in Reference [12] in limit with 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 [13] can be interpreted as communications in the opposite direction of our system: distributed, cooperative helpers have a message for a base, but their cooperation ability is limited and their modulations must be sent over a noisy channel.

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 $E[X_t^2] = 1$. Over a period of t channel uses, a sequence of t values are broadcast, $X = (X_1, \dots, X_t)$. 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 "0^{th}" 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 $n \in [0 : N]$, the n^{th} receiver observes

$$Y_n = h_n X + W_n \tag{1}$$

where

- $\boldsymbol{h} \triangleq (h_0, \dots, h_N)^T \in \mathbb{C}^{(N+1)\times 1}$ is a deterministic channel
- $\boldsymbol{W} \triangleq (W_0, \dots, W_N) \sim \mathcal{CN}(0, \boldsymbol{\Sigma})$ is noise, with some covariance matrix $\boldsymbol{\Sigma}$

The vector of helper's receptions is denoted as

$$\boldsymbol{Y} \triangleq (Y_0, Y_1, \dots, Y_N) \in \mathbb{C}^{(N+1)\times 1}. \tag{2}$$

It is instructive to see that if the noise covariance $\boldsymbol{\Sigma}$ is diagonal then there is no interference, and receivers experience totally independent noise. In contrast, if we can perform the decomposition $\boldsymbol{\Sigma} = \boldsymbol{a}\boldsymbol{a}^\dagger$ 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 $n \in [1 : N]$ 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

$$\boldsymbol{r} \triangleq (r_1, \dots, r_N) \in \mathbb{C}^{(N+1)\times 1}. \tag{3}$$

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 $\boldsymbol{r}$ belong to a set $\mathcal{R}_{\text{LAN}}(L)$ with a sum-capacity parameter $L > 0$

$$\mathcal{R}_{\text{LAN}}(L) \triangleq \left\{ \boldsymbol{r} \,\middle|\, \sum_{n=1}^{N} r_n \leq L,\ r_n \geq 0 \right\}. \tag{4}$$

We refer to the condition that $\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L)$ 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 $n \in [1 : N]$ employs a quantizer block Q_n to produce its coarse r_n-bitrate summary of its observations. Q_n is any channelwhose output depends only on Y_n and possibly a shared randomness independent of the noise and message, and has average-Shannon-entropy-per-timeslot below r_n. Q_n's output is a discrete random variable we label U_n, and the vector of helper messages received by the base at each channel use is denoted

$$\boldsymbol{U} \triangleq (U_1, \ldots, U_N). \tag{5}$$

Equivalently, the quantizers are blocks that produce $\boldsymbol{U}$ where:

- The probability distribution of $(X, \boldsymbol{Y}, \boldsymbol{U})$ is separable as

$$P_{(X,\boldsymbol{Y},\boldsymbol{U})} = P_X \cdot P_{\boldsymbol{Y}|X} \cdot \prod_{n=1}^{N} P_{U_n|Y_n}. \tag{6}$$

- Quantizer messages are within the LAN constraint

$$H(U_i) \leq r_i, i \in [1 : K]. \tag{7}$$

Channel $\boldsymbol{h}$ and noise covariance $\boldsymbol{\Sigma}$ are assumed to be static throughout the transmission of each message, and have been estimated by the receivers a priori and are known to both the broadcast and receive nodes. Both the transmitter and the receivers are assumed to have knowledge of the set of feasible rates, $\mathcal{R}_{\text{LAN}}(L)$. A block diagram of the system is shown in Figure 2.

The environment determines:

- channel fades $\boldsymbol{h}$
- noise covariance $\boldsymbol{\Sigma}$
- maximum LAN throughput L

Available for design are:

- the distribution of the source X
- the rates for the base to collect from each helper $\boldsymbol{r}$
- 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 cut-set upper bound on the system's capacity comes by considering a relaxation where all helpers are informed of each other's observations. Expanding the mutual information from source to base:

$$\begin{aligned} I(X;\hat{X}) &\leq I(X;Y_0,\boldsymbol{U}) && (8)\\ &\leq I(X;\boldsymbol{U}|Y_0) + I(X;Y_0) && (9)\\ &\leq H(\boldsymbol{U}) + I(X;Y_0) && (10)\\ &\leq L + \log_2\left[1 + \|h_0\|^2/\boldsymbol{\Sigma}_{1,1}\right]. && (11) \end{aligned}$$

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 Gaussian receiver. By noisy channel coding [14], the capacity of this channel is $\max_{P_X:\text{Var}(X)<1} I(X;\boldsymbol{Y})$ which, by a derivation given in [15], simplifies to the familiar upper bound:

$$\log_2\left[\frac{|\boldsymbol{\Sigma} + \boldsymbol{h}\boldsymbol{h}^\dagger|}{|\boldsymbol{\Sigma}|}\right]. \tag{12}$$

Taking the minimum of this bound and (11) yields an upper bound for the original system.

Remark 1. *Any rate of reliable communications R for the*

TABLE I
Notation and Terminology

$\mathbb{R}_+$	Nonnegative real numbers
$\boldsymbol{a} \succeq 0$	Vector $\boldsymbol{a}$ has non-negative elements
$\mathbf{A} \succeq 0$	Matrix $\mathbf{A}$ is positive-semi-definite
$\mathbf{1}_N$	N-row vector containing 1s
$\mathbf{A}^\dagger, \boldsymbol{a}^\dagger$	Conjugate transpose of matrix $\mathbf{A}$, vector $\boldsymbol{a}$
$\boldsymbol{a}_S; S \subseteq T$	Vector of elements indexed by S
$S\backslash T$	$S \cap T^C$
$\mathcal{CN}(0, \boldsymbol{\Sigma})$	Complex circularly symmetric normal distribution with zero mean and Hermitian covariance matrix $\boldsymbol{\Sigma}$
$\mathbf{A}_{j,k}$	Element on row j, column k of a matrix $\mathbf{A}$
$\text{diag}(\boldsymbol{a})$	Diagonal matrix where the $(i,i)^{\text{th}}$ diagonal element is the i^{th} element of vector $\boldsymbol{a}$
$\mathbf{I}_{n\times n}$	$n \times n$ identity matrix
$\|\mathbf{A}\|$	Determinant of a matrix $\mathbf{A}$
$\|\|x\|\|$	Scalar norm/magnitude
$[1 : n]$	The set of integers from 1 to n, inclusive

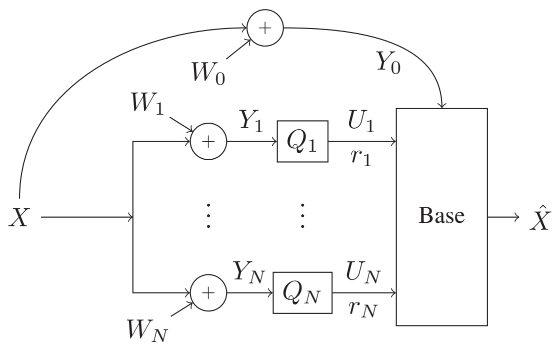


Fig. 2. The system in consideration. A message is broadcast as X with average power 1. The signal is received by a base and N helper nodes, occluded by correlated AWGN $\boldsymbol{W}$. Helper n for $n \in [1 : N]$ quantizes Y_n through Q_n to produce an r_n-bit summary U_n. The quantized observations $\boldsymbol{U}$ are combined at the base along with the base's full-precision reception Y_0 to produce $\hat{X}$, an estimate of X.

system must satisfy

$$R \leq \min\left\{\log_2\left[\frac{|\boldsymbol{\Sigma} + \boldsymbol{h}\boldsymbol{h}^\dagger|}{|\boldsymbol{\Sigma}|}\right], \quad L + \log_2\left[1 + \|h_0\|^2/\boldsymbol{\Sigma}_{1,1}\right]\right\}. \tag{13}$$

Proof: Justified by the preceding discussion. ∎

A. Achievable Rate by Decoding and Forwarding

Treating each helper node as a user seeking to receive its own message, the link from the broadcaster to helpers is a scalar Gaussian broadcast channel. The capacity region of this channel has been characterized in Reference [16], and in particular its sum-rate is given in Reference [17]. In the scalar case this sum rate reduces to:

$$\max_n \log_2\left[\frac{\boldsymbol{\Sigma}_{n,n} + |h_n|^2}{\boldsymbol{\Sigma}_{n,n}}\right], \tag{14}$$

where only the point-to-point channel between transmitter and the best receiver is used.

Remark 2. *The following rate is achievable*

$$R_{BC}(L) = \max\left\{\log_2\left[\frac{\boldsymbol{\Sigma}_{0,0} + |h_0|^2}{\boldsymbol{\Sigma}_{0,0}}\right], \quad \min\left\{L, \max_n \log_2\left[\frac{\boldsymbol{\Sigma}_{n,n} + |h_n|^2}{\boldsymbol{\Sigma}_{n,n}}\right]\right\}\right\} \tag{15}$$

Proof: By having the broadcaster send a message to the n^{th} receiver, a rate

$$\log_2\left[\frac{\boldsymbol{\Sigma}_{n,n} + |h_n|^2}{\boldsymbol{\Sigma}_{n,n}}\right] \tag{16}$$

is achievable from transmitter-to-receiver. If $n = 0$ (i.e. the transmitter broadcasts to the base) then (16) is achievable from transmitter-to-base. Otherwise (16) is achievable from transmitter-to-base if it is less than L by forwarding the receiver's decoding over the LAN. ∎

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

$$L \leq \max_n \log_2\left[\frac{\boldsymbol{\Sigma}_{n,n} + |h_n|^2}{\boldsymbol{\Sigma}_{n,n}}\right] \tag{17}$$

then the system's capacity is L.

Proof: By assumption, the upper bound in (13) equals L. This is achievable by Remark 2. ∎

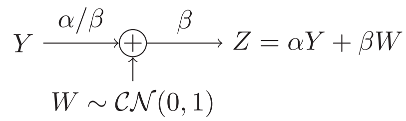
The strategy to achieve this rate requires a strong amount of cooperation from the broadcaster, since the decoding receivers must all share codebooks with the transmitter.

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 $\mathcal{D} > 0$, then the minimum rate where this is feasible is

$$\mathcal{R}(\mathcal{D}) = \log_2(\sigma^2) - \log_2(\mathcal{D}). \tag{18}$$

To achieve this rate, the encoding operation must emulate the following test channel (Reference [18]):



(19)

where $\alpha = \sqrt{1 - \mathcal{D}/\sigma^2}$ and $\beta = \sqrt{\mathcal{D}}$. 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 [19]. Decodings of messages from a helper transmitting at rate r_n would have the form

$$Z_n(r_n) \triangleq Y_n + W_{Q,n}(r_n) \in \mathbb{C}^{N \times 1} \tag{20}$$

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

$$\boldsymbol{W}_Q(\boldsymbol{r}) \triangleq (W_{Q,1}(r_1), \ldots, W_{Q,N}(r_n)) \in \mathbb{C}^{N \times 1}, \tag{21}$$

and the vector of Z_ns corresponding to this choice of $W_{Q,n}$ as:

$$\boldsymbol{Z}(\boldsymbol{r}) \triangleq \boldsymbol{Y} + \boldsymbol{W}_Q(\boldsymbol{r}). \tag{22}$$

The subscript Q is a decoration to distinguish quantization distortion terms $\boldsymbol{W}_Q, W_{Q,n}$ from environment noise, $\boldsymbol{W}, W_n$. 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 (18) equal to r_n and solving for $\mathcal{D}$, with σ^2 equal to the helper observation's variance, $\|h_n\|^2 + \boldsymbol{\Sigma}_{n,n}$. Scaling the output of the test channel the helper emulates (Equation (19)) by $1/\alpha$ causes it to equal the helper's observation plus independent Gaussian noise with variance $(\beta/\alpha)^2 = \mathcal{D}/(1 - \mathcal{D}/\sigma^2)$. This means that in such a system, the distortion from the helpers is equivalent to adding independent additive Gaussian noise with variance:

$$\mathrm{Var}(W_{Q,n}(r_n)) = \frac{\|h_n\|^2 + \boldsymbol{\Sigma}_{n,n}}{2^{r_n} - 1}. \tag{23}$$

By (1) and (23), all the noise and distortion on the signal present in the helpers' messages to the base are summarized by the vector $\boldsymbol{W} + \boldsymbol{W}_Q$ with covariance matrix

$$\mathbf{D}(\boldsymbol{r}) \triangleq \boldsymbol{\Sigma} + \mathrm{diag}\left(0, \frac{\|h_1\|^2 + \boldsymbol{\Sigma}_{1,1}}{2^{r_1} - 1}, \ldots, \frac{\|h_N\|^2 + \boldsymbol{\Sigma}_{N,N}}{2^{r_N} - 1}\right). \tag{24}$$

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

$$R_G(\boldsymbol{r}) \triangleq I(X; \boldsymbol{Z}(\boldsymbol{r})) = \log_2\left[\frac{|\mathbf{D}(\boldsymbol{r}) + \boldsymbol{h}\boldsymbol{h}^\dagger|}{|\mathbf{D}(\boldsymbol{r})|}\right]. \tag{25}$$

With this intuition we can state the following:

Theorem 1. *For a distributed receive system as described in Section III with noise covariance matrix $\boldsymbol{\Sigma}$, LAN constraint*

L and fixed average helper quantization rates $\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L)$, then a rate $R_G(\boldsymbol{r})$ is achievable.

A proof of this is given in Appendix B.

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

$$R_{G,\max}(L) \triangleq \max_{\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L)} R_G(\boldsymbol{r}). \tag{26}$$

Unfortunately this expression cannot be simplified much further outside of a few special cases. Maximization of R_G over $\mathcal{R}_{\text{LAN}}(\cdot)$ can be performed efficiently using quasi-convex optimization techniques (See Reference [20]).

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

Proof: The maximum of R_G follows the maximum of $-\boldsymbol{h}^\dagger \mathbf{D}(\boldsymbol{r})^{-1}\boldsymbol{h}$ 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 [20]). Further, a 1-dimensional function is quasi-concave if anywhere its derivative is 0, its second derivative is below 0.

Fix any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^{N\times 1}$, denote $\mathbf{D}_t \triangleq \mathbf{D}(\cdot)|_{t\boldsymbol{a}+\boldsymbol{b}}$ for any t where $t\boldsymbol{a}+\boldsymbol{b} \in \mathbb{R}_+^{N\times 1}$. Define $\mathbf{A}_t \triangleq \frac{d}{dt}\mathbf{D}_t$, and $f(t) \triangleq -\boldsymbol{h}^\dagger \mathbf{D}_t^{-1}\boldsymbol{h}$. Then

$$\frac{d}{dt} f(t) = \mathbf{A}_t \mathbf{D}_t^{-1}\boldsymbol{h} \tag{27}$$

$$\frac{d^2}{dt^2} f(t) = (\mathbf{D}_t^{-1}\boldsymbol{h})^\dagger \left[\frac{d}{dt}\mathbf{A}_t - 2\mathbf{A}_t\mathbf{D}_t^{-1}\mathbf{A}_t\right](\mathbf{D}_t^{-1}\boldsymbol{h}). \tag{28}$$

If $\frac{d}{dt}f(t) = 0$, then $\mathbf{D}_t^{-1}\boldsymbol{h}$ is in the null space of $\mathbf{A}_t$ so the $-2\mathbf{A}_t\mathbf{D}_t^{-1}\mathbf{A}$ term in (28) vanishes. $\frac{d}{dt}\mathbf{A}_t$ is negative definite in $\mathbb{R}_+^{N\times 1}$ so $\frac{d^2}{dt}f(t) < 0$ there, establishing f's quasi-concavity and thus that of $-\boldsymbol{h}^\dagger \mathbf{D}_t^{-1}\boldsymbol{h}^\dagger$. Since $-\boldsymbol{h}^\dagger \mathbf{D}_t^{-1}\boldsymbol{h}^\dagger$ is quasi-concave along all of $\mathbb{R}_+^{N\times 1}$, then it is also quasi-concave in any convex restriction of that domain. Simplexes are convex sets, establishing the claim. ■

$R_{G,\max}(L)$ is tight with the upper bound in (13), in limit with LAN throughput.

Remark 5. *$R_{G,\max}(L) \to C$ as $L \to \infty$.*

Proof: $R_{G\max}(L) \geq R_G(L/N \cdot \mathbf{1})$, and if $L \to \infty$ then

$$R_G(\ L/N \cdot \mathbf{1}\ ) \to \log_2\left[\frac{|\boldsymbol{\Sigma} + \boldsymbol{h}\boldsymbol{h}^\dagger|}{|\boldsymbol{\Sigma}|}\right] \tag{29}$$

by continuity of $\log_2$, sums, and the matrix determinant (in terms of matrix coordinates). The limit in (29) is greater than or equal to (13), establishing the remark. ■

This strategy has the benefit of not needing much coordination between receive nodes. The transmitter does not have to share codes with helpers, because here the helpers do not perform any decoding. Each helper needs only its encoding rate and its own channel state to form its messages properly.

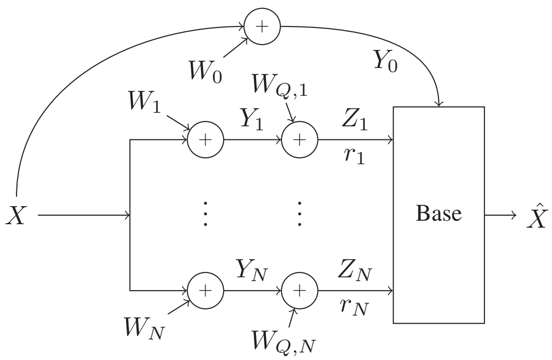


Fig. 3. Gaussian distortion System: Quantizer distortion is modeled as additive white Gaussian noise, where enough noise is added so that the capacity across quantizers is some given rate vector $\boldsymbol{r}$.

C. Achievable Rate through Distributed Compression

Helper n for $n \in [1:N]$ in a realization of the Gaussian distortion system described in Section IV-B, on average and with long codes, will produce an r_n-bit encoding of its observation that can be decoded to produce the helper observation with additional approximately Gaussian noise. Since all the helper's quantizations contain the same signal component (and possibly the same interference), they are correlated and and can be compressed before forwarding to allow for less noise to be introduced in quantization.

The Slepian-Wolf theorem [21] shows that if the LAN is such that helpers can encode at rates $\boldsymbol{\rho} \triangleq (\rho_1, \ldots, \rho_N)$, then is possible for the helpers to further jointly-compress their encodings with no loss down to rates $\boldsymbol{r}$, as long as $\boldsymbol{r}$ and the encodings $\boldsymbol{U}$ satisfy the conditions that for all subsets $S \subseteq [1:N]$, then:

$$H(\boldsymbol{U}_S|\boldsymbol{U}_{S^C}, Y_0) < \sum_{n\in S} r_n. \tag{30}$$

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 (20), then for large enough n the value $H(\boldsymbol{U}_S|\boldsymbol{U}_{S^C}, Y_0)$ equals $I(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{\rho})|\boldsymbol{Z}_{S^C}(\boldsymbol{\rho}), Y_0)$. Maximizing over the LAN constraint (Equation (4)), the following rate is achievable:

$$R_{DC}(L) \triangleq \max_{\boldsymbol{r}\,\in\mathcal{R}_{\text{LAN}}(L)}\ \max_{\boldsymbol{\rho}\,\in\mathcal{R}_{DC}(\boldsymbol{r})} R_G(\boldsymbol{\rho}) \tag{31}$$

where

$$\mathcal{R}_{DC}(\boldsymbol{r}) = \Big\{\boldsymbol{\rho} : \forall S \subseteq [1:N], \quad I\left(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{\rho})|\boldsymbol{Z}_{S^C}(\boldsymbol{\rho}), Y_0\right) < \sum_{n\in S} r_n\Big\}. \tag{32}$$

$\boldsymbol{Z}(\cdot)$ in (32) is defined as in (22). We can then state the following:

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

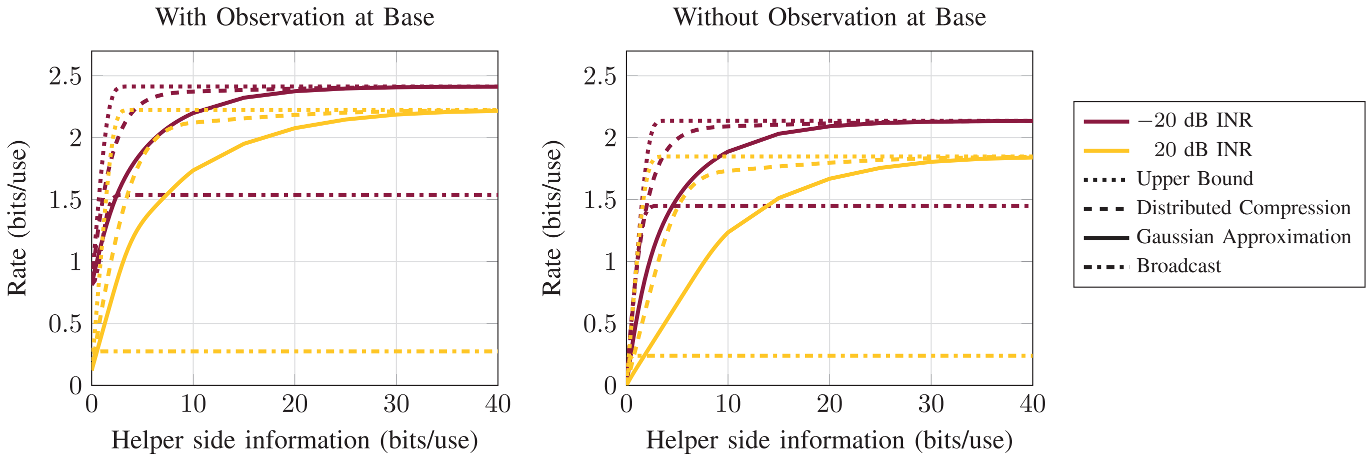


Fig. 4. Bound performance versus total cooperation ability between helpers. Single-user decode-and-forward becomes infeasible in interference (Broadcast curve drops dramatically from -20 dB to 20 dB INR). In interference, more thorough use of receive diversity is needed to attain reasonable communication rates. Distributed compression only offers great benefit over Gaussian compress-and-forward when there are few receive nodes and collaboration between them is severely limited (compare the bound difference between Distributed Compression and Gaussian Approximation in 20 dB INR with an observation at base to without an observation at base). 4 helpers with 0 dB average SNR, a LAN supporting $L = 5$ bits total from helpers to base, averaged over 1000 $\mathcal{CN}(0, I)$ channels

The base can unambiguously decompress the helper's compressed encodings with low probability of error if and only if $\boldsymbol{\rho}$ is chosen such that (30) is satisfied. However, in analogue to Corollary 1 from Reference [1], the rate in 2 can be improved by expanding $\mathcal{R}_{DC}(\boldsymbol{r})$ to include some of the $\boldsymbol{\rho}$ where the base cannot perform unambiguous decompression. This helps because even if the encoding rates in $\boldsymbol{\rho}$ are chosen outside of $\mathcal{R}_{DC}(\boldsymbol{r})$ so that helpers cannot convey $\boldsymbol{U}$ 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 $\boldsymbol{\Sigma}$ and LAN constraint L then the following rate is achievable:*

$$R_{\overline{DC}}(L) \triangleq \max_{\substack{\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L),\\ \lambda \in \mathbb{R}}} \; \max_{\boldsymbol{\rho} \in \mathcal{R}^{\lambda}_{DC}(\boldsymbol{r})} R_G(\boldsymbol{\rho}) - \lambda \tag{33}$$

where

$$\begin{aligned} \mathcal{R}^{\lambda}_{DC}(\boldsymbol{r}) \triangleq \Bigg\{ &\boldsymbol{\rho} : \forall S \subseteq [1:N], \\ &I\left(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{\rho}) | \boldsymbol{Z}_{S^C}(\boldsymbol{\rho}), Y_0\right) < \sum_{n \in S} r_n + \lambda \Bigg\}. \end{aligned} \tag{34}$$

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

Remark 6. *(Due to Reference [1]) The capacity of the system is $R_{\overline{DC}}(L)$ under the following restrictions:*

- *$\boldsymbol{\Sigma}$ is diagonal (no interference).*
- *The base does not have its own full-precision observation of the broadcast ($h_0 = 0$)*
- *The broadcaster must transmit a Gaussian signal*
- *Helper messages are independent of the transmitter's codebook*

This is demonstrated in Appendix C. 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, which is not accounted for in the compress-and-forward strategy used for $R_{\overline{DC}}(L)$.
- The Gaussian broadcast assumption is needed because of a counterexample given in Reference [1].
- Codebook independence is necessary because $R_{\overline{DC}}(L)$ is strictly less than the upper bound in (13), 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 [1] since we show that the same rate can be achieved with less cooperation between transmitter and receivers. Reference [1] 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 $\boldsymbol{\Sigma}$
- Channel $\boldsymbol{h}$ (assumed to be static and precisely estimated a priori per each channel use).

In all graphs, the rate from (13) is called 'Upper Bound,' the rate from (25) is 'Gaussian Distortion,' the rate from Theorem

3 is 'Distributed Compression' and the rate from Remark 2 is 'Broadcast.' A discussion of the optimizations performed for the Gaussian Distortion and Distributed Compression bounds is in Appendix A.

A. Performance in the Presence of an Interferer

Receive diversity helps a system tolerate interference. In this section we investigate the extent to which a distributed receiver is able to perform interference mitigation. We assume the transmitter has no knowlege of the interferer's operation beyond its output statistics, and that the interferer is uninformed of the transmitter's codebook. This assumption prohibits using dirty paper coding (Reference [22]) and other transmit-side interference mitigation strategies.

We model an interferer as additive Gaussian 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 $\mathbf{A} = \boldsymbol{a}\boldsymbol{a}^\dagger$, where $\boldsymbol{a} \in \mathbb{C}^{N+1}$. In all trials run, a single interferer was assumed to be present at all nodes with power α and random phase so that

$$\boldsymbol{\Sigma} = \mathbf{I}_{N+1\times N+1} + \alpha \cdot \boldsymbol{a}\boldsymbol{a}^\dagger. \tag{35}$$

with $\boldsymbol{a}_n = e^{j2\pi\cdot\theta_n},\ \theta_n \sim \mathsf{Unif}[0,1]$ where Unif is the uniform distribution.

Figure 4 shows the difference in achievable rates with and without interfernece for different levels of receive-side connectivity. Even when a strong rank-one interferer is present at all the nodes, achievable communications rates are comparable to the case without an interferer.

Reference [1] provides an example of a system such as the one in the present work where a Gaussian transmitter is very sub-optimal. In contrast, we see in Figure 4 that over an average of many random channels in a range of settings, achievable rates using Gaussian signalling come quite close to an upper bound on the system. Even the relatively simple Gaussian distortion bound is close in performance to the optimum in all regimes but one with strong interference and little cumulative helper-to-base information. In this regime the LAN constraint limits the helper's ability to provide the base with the diversity of observations necessary to perform good interference mitigation.

B. Path Diversity versus Performance

It is reasonable to expect that the best strategy for a distributed receiver to use will change depending on its environment. If the receivers observe the signal through a single line-of-sight path, all else being equal, the signal and noise at each receiver will have similar statistics. In contrast, in environments with many scatterers, the channel statistics will vary more across spatially distributed receivers.

The level of attenuation at a receiver can modeled with a Rician distribution [23]. 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 Ω. In contrast if $K \to \infty$ then the distribution approaches a point mass at Ω (Reference [24]).

The type of scattering environment does not greatly affect the average rate achievable by any bound other than the Broadcast bound (Figure 6). This is because the Broadcast bound's only utilization of receive diversity is in channel variances which is small when the channel is dominated by line-of-sight receptions. Despite the invariance of most bounds to K, the profile of helper-to-base bits 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 while mostly ignoring low-SNR helpers' observations. The imbalance is less pronounced in interference and when SNR is less varied across receivers (high K), since in these regimes, gaining diverse observations for combining is more beneficial than using the strongest helper's observation.

When the base collects its own full-precision observation, the helper rate profile is unaffected by intererence and scattering, where most helper information is provided by the highest SNR helper. In practice, in this situation the base might use its observation as an estimate of the interferer and subtract it from a compression of the strongest helper's observation.

VI. Conclusion

In total, one upper bound and three lower bounds on capacity were shown. A simple upper bound (Equation (13)) 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 a broadcast channel is shown (Equation (15)). Another achievable rate (Equation (25)) 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 (31)) was derived by using a distributed compression technique on the helper observations from the bound in (25). It outperforms the bound in (25), but requires intensive a priori coordination between helpers.

Results in Section V demonstrate that distributed compression does not drastically improve performance unless the LAN is severely limited in the presence of an interferer. Leaving SNR fixed, performance is mostly unchanged in high-scattering versus low-scattering environments, although the profile of helper-to-base communication changes. 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. In the presence of interference, more diversity is needed, so the base must utilize low-SNR helpers' observations more to maintian the same level of performance.

It remains to extend these results to situations where an interferer is adversarial, or where all nodes wish to receive the

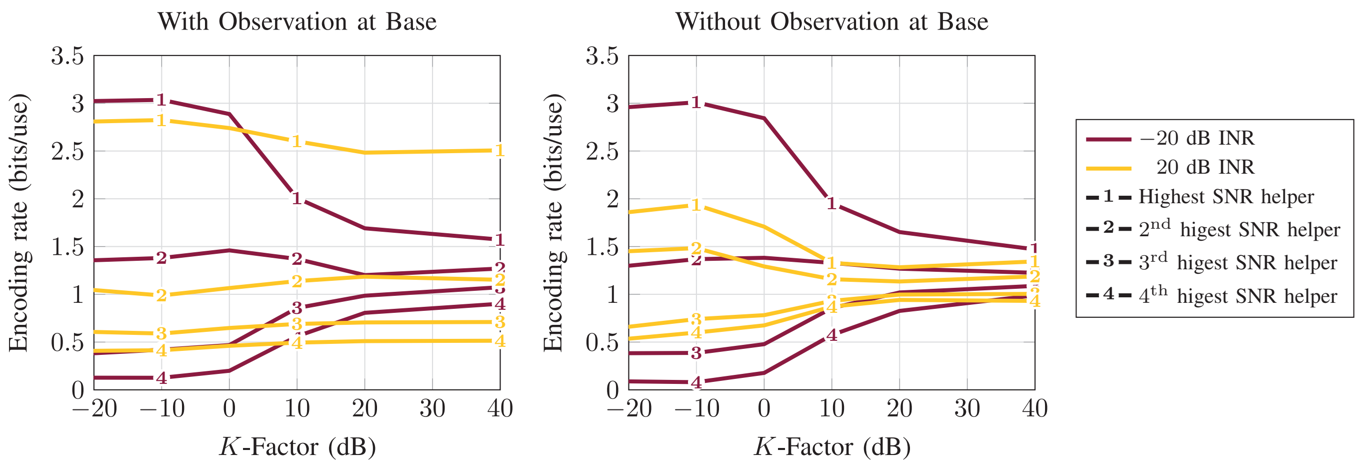


Fig. 5. Average rate with which each helper forwards to base in distributed compression (Theorem 3) versus K, scattered-to-direct-path received signal power ratio. As the proportion of line-of-sight path power increases, the need for receive diversity increases and the base draws information from more helpers. This does not occur when the base has its own observation, when presumably it is used as an estimate of the interferer to be mitigated from a compression of the highest-SNR helper's observation. 4 helpers with 0 dB average SNR, a LAN with $L = 5$ bits total from helpers to base, averaged over 1000 $\mathcal{CN}(0, I)$ channels.

broadcast rather than just the base. Practical implementations of systems that can operate near the achievable rates shown here are also pending investigation.

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

Appendix A
Optimization Method

Evaluation of the Gaussian distortion and distributed compression bounds in each environment requires optimization of a quasi-convex objective (Remark 4). In the case of distributed compression, the space of acceptable parameters is not convex: consider compressing the encodings for two highly-correlated sources. By the Slepian-Wolf theorem, if the encodings are high-bitrate then perturbing the encodings' bitrate does not greatly affect the region of feasible compression rates, since the encodings have high redundancy. On the other hand, perturbing the bitrate of low-bitrate encodings greatly affects the region of feasible compression rates, since uncorrelated distortion from encoding dominates source redundancies. Then a compression to an average between minimum feasible compression rates for low-bitrate encodings and ones for high-bitrate encodings is infeasible, and the distance from the feasible set becomes large as the difference between high and low bitrates is made large. This is the mechanism that governs $\mathcal{R}_{\overline{DC}}(\boldsymbol{r}) \subseteq \mathbb{R}_+^K$, hence the domain's non-convexity.

To overcome this, an iterative interior point method was used to find the maximum: each constraint $f(x) < 0$ was replaced with a stricter constraint, $f(x) + \beta < 0$ for some $\beta > 0$ and minimization was performed. For the next iteration, β in the inequalities was replaced with a smaller $\beta' < \beta$ and another minimization was run using the previous optimum as an initial guess. The process was repeated until all the inequalities became close to the initial $f(x) < 0$. Each individual minimization was performed using sequential least-squares programming (SLSQP) through SciPy [25].

Appendix B
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* $\boldsymbol{Y}^t$*, and lattice encodings* $\boldsymbol{U}^t$ *with distortion vectors* $\tilde{\boldsymbol{W}}_Q^t$ *as constructed in the* Helper encoder setup *subsections of the proofs for Theorems 1 or 3, where* $\mathrm{Var}(\tilde{\boldsymbol{W}}_{Q,n,\ell}^t) = \sigma_{Q,n}^2$ *(*$\ell \in [1;t]$*), for any* $S \subseteq [1:N]$ *then for any* $\varepsilon > 0$*, there is some blocklength* t *large enough that*

$$\left\| \frac{1}{t} H(\boldsymbol{U}_S^t) - \frac{1}{t} I(\boldsymbol{Y}_S^t; \boldsymbol{Z}_S^t(\boldsymbol{r})) \right\| < \varepsilon \tag{36}$$

where $\boldsymbol{Z}^t(\boldsymbol{r}) = \boldsymbol{Y}^t + \boldsymbol{W}^t$*, with* $\boldsymbol{W}^t \perp \boldsymbol{Y}^t$ *and*

$$\boldsymbol{W}^t \sim \mathcal{CN}\left(0, diag\left(\sigma_{Q,1}^2, \ldots, \sigma_{Q,N}^2\right)\right). \tag{37}$$

Proof: We prove this by induction. By Theorem 1 in Reference [26], for any $n \in [1:N]$

$$H(\boldsymbol{U}_n^t) = I(\boldsymbol{Y}_n^t; \boldsymbol{Y}_n^t - \tilde{\boldsymbol{W}}_{Q,n}^t). \tag{38}$$

Further, by Theorem 3 in Reference [19]

$$\frac{1}{t} I\left(Y_n^t; Y_n^t - \tilde{W}_{Q,n}^t\right) \underset{t\uparrow\infty}{\rightarrow} I(Y_n;\, Y_n + W_n) = I(Y_n; Z_n(r_n)). \tag{39}$$

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

Now assume there are two disjoint sets A, B for which the claim holds for arbitrarily small $\varepsilon' > 0$. Then there is a deterministic function at the decoder $f_{A\cup B}$ where $f_{A\cup B}(\boldsymbol{U}^t_{A\cup B}) = \boldsymbol{Y}^t_{A\cup B} - \tilde{\boldsymbol{W}}^t_{Q,A\cup B}$, so

$$H(\boldsymbol{U}^t_{A\cup B}) \geq I(f(\boldsymbol{U}^t_{A\cup B}); \boldsymbol{Y}^t_{A\cup B}) \quad (40)$$
$$\geq I(\boldsymbol{Y}^t_{A\cup B}; \boldsymbol{Z}^t_{A\cup B}(\boldsymbol{r})). \quad (41)$$

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 [18]). Similarly,

$$H(\boldsymbol{U}^t_{A\cup B}) = H(\boldsymbol{U}^t_A) + H(\boldsymbol{U}^t_B) - I(\boldsymbol{U}^t_A; \boldsymbol{U}^t_B) \quad (42)$$
$$\leq I(\boldsymbol{Y}^t_A; \boldsymbol{Z}^t_A(\boldsymbol{r})) + I(\boldsymbol{Y}^t_B; \boldsymbol{Z}^t_B(\boldsymbol{r})) - \ldots$$
$$I(\boldsymbol{U}^t_A; \boldsymbol{U}^t_B) + 2\varepsilon. \quad (43)$$

By construction, at the decoder there are deterministic functions f_A, f_B where $f_A(\boldsymbol{U}^t_A) = \boldsymbol{Y}^t_A - \tilde{\boldsymbol{W}}^t_{Q,A}$, and $f_B(\boldsymbol{U}^t_B) = \boldsymbol{Y}^t_B - \tilde{\boldsymbol{W}}^t_{Q,B}$ so

$$I(\boldsymbol{U}^t_A; \boldsymbol{U}^t_B) \geq I(f_{A\cup B}(\boldsymbol{U}^t_A); f(\boldsymbol{U}^t_B)) \quad (44)$$
$$\geq I(\boldsymbol{Z}^t_A(\boldsymbol{r}); \boldsymbol{Z}^t_B(\boldsymbol{r})) - \varepsilon' \quad (45)$$

where the first line follows by using the data processing inequality twice, and the second follows by the same Gaussian property used above. Combining (41), (43) and (45) then

$$0 \leq H(\boldsymbol{U}^t_{A\cup B}) - I(\boldsymbol{Y}^t_{A\cup B}; \boldsymbol{Z}^t_{A\cup B}(\boldsymbol{r})) \leq 3\varepsilon'. \quad (46)$$

Since $S \subseteq [1:N]$, the inductive step will only need to be used up to N times. Letting $0 < \varepsilon' < \varepsilon \cdot 3^{-N}$, 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* $\forall S \subseteq [1:N]$ *then*

$$I(\boldsymbol{Y}_S, \boldsymbol{Z}_S(\boldsymbol{r})|\boldsymbol{Z}_{S^C}(\boldsymbol{r}), Y_0) < \sum_{m\in S} r_m, \quad (47)$$

then some value r_0 *can be chosen sufficiently large so that for any* $S \subseteq [1:N]$, *both*

$$I(\boldsymbol{Y}_S, \boldsymbol{Z}_S(\boldsymbol{r}), Z_0(r_0)|\boldsymbol{Z}_{S^C}(\boldsymbol{r})) < \sum_{m\in S} r_m + r_0. \quad (48)$$

and

$$I(\boldsymbol{Y}_S, \boldsymbol{Z}_S(\boldsymbol{r})|\boldsymbol{Z}_{S^C}(\boldsymbol{r}), Z_0(r_0)) < \sum_{m\in S} r_m. \quad (49)$$

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

$$I(\boldsymbol{Y}_S, \boldsymbol{Z}_S(\boldsymbol{r})|\boldsymbol{Z}_{S^C}(\boldsymbol{r}), Y_0) < \sum_{m\in S} r_m - \delta. \quad (50)$$

Fix r_0 large enough so that for any $S \subseteq [1:N]$,

$$\begin{aligned}&\|I(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{r})|\boldsymbol{Z}_{S^C}(\boldsymbol{r}), Y_0)\\ &- I(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{r})|\boldsymbol{Z}_{S^C}(\boldsymbol{r}), Z_0(r_0))\| \leq \delta.\end{aligned} \quad (51)$$

There is guaranteed to be such an r_0 because the above expression is a continuous function of the covariance matrix of $(Y_0, \boldsymbol{Y}, Z_0(r_0), \boldsymbol{Z}(\boldsymbol{r}))$ component-wise, and as r_0 is made large, the covariance matrices invovled in the second term converge component-wise to those of the first since $\frac{\|h_0\|^2 + \boldsymbol{\Sigma}_{1,1}}{2^{r_0}-1} \to 0$ as $r_0 \to \infty$. By (50) and (51) the statement holds for any $S \subseteq [1:N]$, and it remains to show that it also holds for each $S \cup \{0\}$.

Note that for any $S \subseteq [1:N]$

$$I(\boldsymbol{Y}^t_S, \boldsymbol{Y}^t_0; \boldsymbol{Z}^t_S(\boldsymbol{r}), \boldsymbol{Z}^t_0(\boldsymbol{r})|\boldsymbol{Z}^t_{(S\cup\{0\})^C}(\boldsymbol{r})) \quad (52)$$
$$= I(\boldsymbol{Y}^t_S; \boldsymbol{Z}^t_S(\boldsymbol{r})|\boldsymbol{Z}^t_{(S\cup\{0\})^C}(\boldsymbol{r})) + I(\boldsymbol{Y}^t_0; \boldsymbol{Z}^t_0(\boldsymbol{r})|\boldsymbol{Z}^t_{[1:N]}(\boldsymbol{r})) \quad (53)$$
$$\leq I(\boldsymbol{Y}^t_S; \boldsymbol{Z}^t_S(\boldsymbol{r})|\boldsymbol{Z}^t_{(S\cup\{0\})^C}(\boldsymbol{r})) + r_0 \quad (54)$$
$$\leq \sum_{m\in S} r_m + r_0 \quad (55)$$

where (54) follows because condition reduces mutual information, and (55) 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* $\boldsymbol{\Sigma}$, *LAN constraint* L *and fixed average helper quantization rates* $\boldsymbol{r} \in \mathcal{R}_{\mathrm{LAN}}(L)$, *then a rate* $R_G(\boldsymbol{r})$ *is achievable.*

Proof: If for some $n \in [1:N]$ 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 $n \in [1:N]$. Fix some rate $R < R_G(\boldsymbol{r})$ and a block length $T = t^2 \in \mathbb{N}$.

Operation:

Transmitter setup: Generate a codebook:

$$\mathcal{X} = \{X_1, \ldots, X_{2^{tR}}\} \subset \mathbb{C}^{t\times 1} \quad (56)$$

where all the vectors' components are drawn iid from $\mathcal{CN}(0,1)$. Reveal $\mathcal{X}$ to the transmitter and base.

Helper encoder setup: For each helper n, $n \in [1:N]$, generate t distortion vectors $\{\tilde{W}^\ell_{Q,n}\}^t_{\ell=1} \subseteq \mathbb{C}^{t\times 1}$, each independent and uniform in the base region of the Voroni partition $\mathcal{P}^t_n$ of a regular, white $t-$dimensional lattice $\mathcal{L}^t_n$ scaled to have a normalized-second-moment:

$$G^t(\mathcal{L}^t_n, \mathcal{P}^t_n) = \frac{\|h_n\|^2 + \boldsymbol{\Sigma}_{n,n}}{2^{r_n - \varepsilon} - 1}. \quad (57)$$

These terms are detailed at the beginning of Reference [19]. Reveal $\{\tilde{W}^\ell_{Q,n}\}^t_{\ell=1}$ to the base and helper n.

At the base do the same, generating t distortion vectors $\{\tilde{W}^\ell_{Q,0}\}^t_{\ell=1} \subseteq \mathbb{C}^{t\times 1}$ 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.

Transmission: To send a message $M = (m_1, \ldots, m_t) \in \prod^t_{\ell=1}[1:2^{tR}]$, have the transmitter broadcast:

$$\boldsymbol{X} = (X_{m_1}, \ldots, X_{m_t}) \in \mathcal{X}^t \subseteq \mathbb{C}^{T\times 1}. \quad (58)$$

Helper encoding and forwarding: For one transmission period, receiver n ($n \in [0:N]$) observes a sequence of length T which we split up into t sequences of length t:

$$Y_n^\ell \triangleq (Y_{n,\ell\cdot(t-1)+1}, \ldots, Y_{n,\ell\cdot(t-1)+t}) \in \mathbb{C}^{t\times 1}, \quad \ell \in [1 : t] \tag{59}$$

Form a set of quantizations $\{U_n^\ell\}_{\ell=1}^T \subseteq \mathcal{L}_n^t$ by finding the point in $\mathcal{L}_n^t$ corresponding to the region in $\mathcal{P}_n^t$ in which $Y_n^\ell - \tilde{W}_{Q,n}^\ell$ resides. The properties of such U_n^ℓ are the subject of References [26], [19]. By Theorem 1 in Reference [26],

$$H(U_n^\ell) = I(Y_n^\ell; Y_n^\ell - \tilde{W}_{Q,n}^\ell) \tag{60}$$

Further, by Theorem 3 in Reference [19]

$$\frac{1}{t} I\left(Y_n^\ell; Y_n^\ell - \tilde{W}_{Q,n}^\ell\right) \underset{t\uparrow\infty}{\rightarrow} I(Y_n;\ Y_n + W_n) = r_n - \varepsilon \tag{61}$$

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

Decoding: Take $A_\varepsilon^T(\boldsymbol{X}, \boldsymbol{U})$ to be the set of

$$((x^1, \ldots, x^t), (\boldsymbol{u}^1, \ldots, \boldsymbol{u}^t)) \in \mathcal{X}^t \times \left(\prod_{n=0}^N \mathcal{L}_n^t\right)^t \tag{62}$$

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

$$\left(\boldsymbol{X}, (\boldsymbol{U}^1, \ldots, \boldsymbol{U}^t)\right) \tag{63}$$

where

$$\boldsymbol{U}^\ell = (U_0^\ell, \ldots, U_N^\ell). \tag{64}$$

Weak- and joint-typicality are defined in Reference [18].

At the base, find $\hat{x} = (\hat{x}^1, \ldots, \hat{x}^t) \in \mathcal{X}^t$ where $(\hat{x}, (\boldsymbol{U}^1, \ldots, \boldsymbol{U}^t)) \in A_\varepsilon^T(\boldsymbol{X}, \boldsymbol{U})$. Declare error events $\mathcal{E}_0$ if $\boldsymbol{X}$ is not found to be typical with $(\boldsymbol{U}^1, \ldots, \boldsymbol{U}^t)$, and $\mathcal{E}_1$ if there is some $\hat{x} \in \mathcal{X}^t$ where $\hat{x} \neq \boldsymbol{X}$ and $(\hat{x}, (\boldsymbol{U}^1, \ldots, \boldsymbol{U}^t)) \in A_\varepsilon^T(X, \boldsymbol{U})$.

Error analysis: By typicality and the law of large numbers, $P(\mathcal{E}_0) \to 0$ as $t \to \infty$. Also,

$$P(\mathcal{E}_1) \le \sum_{\substack{\hat{x}\in\mathcal{X}^t,\\ \hat{x}\neq\boldsymbol{X}}} P\left(\{(u_0^t, u_1^t, \ldots, u_N^t) : \hat{x} \in A_\varepsilon^T(\boldsymbol{X}, \boldsymbol{U})\}\right) \tag{65}$$

$$\le \sum_{\substack{\hat{x}\in\mathcal{X}^t,\\ \hat{x}\neq\boldsymbol{X}}} 2^{-t\cdot(I(\boldsymbol{X};\boldsymbol{U}^1,\ldots,\boldsymbol{U}^t)-3\varepsilon)} \tag{66}$$

$$< 2^{-t\cdot(I(\boldsymbol{X};\boldsymbol{U}^1,\ldots,\boldsymbol{U}^t)-3\varepsilon-tR)} \tag{67}$$

$$< 2^{-t\cdot(t\cdot I(\boldsymbol{X};\boldsymbol{Z}(\boldsymbol{r}-\boldsymbol{1}\cdot\varepsilon))-3\varepsilon-tR)} \tag{68}$$

Equation (68) follows from Equation (67) because both have the same mean and covariance matrix, and the Gaussian distribution maximizes entropy under these constraints (Reference [24]). So if R is chosen less than $I(X; \boldsymbol{Z}(\boldsymbol{r} - \boldsymbol{1}\cdot\varepsilon)) - 3\varepsilon$ then $P(\mathcal{E}_0 \cup \mathcal{E}_1) \to 0$ as $T \to \infty$. For small enough ε, by lower semi-continuity of mutual information $I(x; \boldsymbol{Z}(\boldsymbol{r} - \boldsymbol{1}\cdot\varepsilon)) - 3\varepsilon$ can be made arbitrarily close to $R_G(\boldsymbol{r})$. ∎

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 $\boldsymbol{\Sigma}$ 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 $\boldsymbol{\Sigma}$ and LAN constraint L then $R_{\overline{DC}}(L)$ is achievable.*

Proof: It is enough to show that any component in the maximization from (33) is achievable. Fix $\lambda \in \mathbb{R}$, a helper rate vector $\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L)$, and a compression rate vector $\boldsymbol{\rho} \in \mathcal{R}_{DC}^\lambda(\boldsymbol{r})$. If for some $n \in [1:N]$ 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 $n = [1:N]$. Fix some rate $R < R_G(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon) - \lambda$ and a block length $T = t^2 \in \mathbb{N}$.

Operation:

Transmitter setup: Generate a codebook $\mathcal{X} = \{X_1, \ldots, X_{2^{tR}}\} \subset \mathbb{C}^{t\times 1}$ 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, $n \in [1:N]$, generate t distortion vectors $\{\tilde{W}_{Q,n}^\ell\}_{\ell=1}^t \subseteq \mathbb{C}^{t\times 1}$, each uniform in the base region of the Voroni partition $\mathcal{P}_n^t$ of a regular, white $n-$dimensional lattice $\mathcal{L}_n^t$ scaled to have a normalized-second-moment:

$$G^t(\mathcal{L}_n^t, \mathcal{P}_n^t) = \frac{\|h_n\|^2 + \boldsymbol{\Sigma}_{n,n}}{2^{\rho_n - \varepsilon} - 1}. \tag{69}$$

These terms are detailed at the beginning of Reference [19]. Share $\{\tilde{W}_{Q,n}^\ell\}_{\ell=1}^t$ with the corresponding helper and the base.

At the base do the same, generating t distortion vectors $\{\tilde{W}_{Q,0}^\ell\}_{\ell=1}^t \subseteq \mathbb{C}^{t\times 1}$ with ρ_0 chosen large enough so that the statement in Lemma 2 holds. These are generated for the purpouse of the base quantizing its own receptions.

At receiver n, $n \in [0:N]$, choose a random mapping, $\mathsf{Index}_n : \mathcal{U}_n \to [1 : 2^{tr_n}]$:

$$\mathsf{Index}_n \sim \mathsf{Unif}\left(\{\phi | \phi : \mathcal{U}_n \to [1 : 2^{tr_n}]\}\right) \tag{70}$$

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

Have each receiver distribute its chosen Index_n to the base.

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

$$\boldsymbol{X} = (X_{m_1}, \ldots, X_{m_t}) \in \mathcal{X}^t \subseteq \mathbb{C}^{T\times 1}. \tag{71}$$

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

$$Y_n^\ell \triangleq (Y_{n,\ell\cdot(t-1)+1},\ldots,Y_{n,\ell\cdot(t-1)+t}) \in \mathbb{C}^{t\times 1}, \quad \ell \in [1:t] \tag{72}$$

Form a set of quantizations $\{\tilde{U}_n^\ell\}_{\ell=1}^T \subseteq \mathcal{L}_n^t$ by finding the point in $\mathcal{L}_n^t$ corresponding to the region in $\mathcal{P}_n^t$ in which $Y_n^\ell - \tilde{W}_{Q,n}^\ell$ resides. The properties of such $\tilde{U}_n^\ell$ are the subject of References [26], [19]. By Lemma 1 and the chain rule for entropy and mutual information, then for large enough t for any $S \subseteq [0:N]$,

$$\frac{1}{t}H(\tilde{U}_S^\ell|\tilde{U}_{S^C}^\ell) \underset{t\uparrow\infty}{\to} I(\boldsymbol{Z}_S(\boldsymbol{\rho});\boldsymbol{Y}_S|\boldsymbol{Z}_{S^C}(\boldsymbol{\rho})) < \sum_{n\in S} r_n \tag{73}$$

where convergence follows from (38) and Theorem 3.1 in Reference [27], and the inequality comes from choice of $\boldsymbol{\rho}$.

Helper encoding truncation: There is a small technicality: the idea is for helpers to convey $\{\tilde{U}_n^\ell\}_{\ell=1}^t$, $n \in [1:N]$ to the base by joint compression through binning, but binning is a strategy for finite-alphabet random variables and the $\tilde{U}_n^\ell$ have countable alphabet. This is resolved by truncating the low-probability parts of their alphabets to produce finite-alphabet variables $U_n^\ell \in \mathcal{U}_n$ with approximately the same statistics.

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

$$H(\boldsymbol{A}) = \sum_{a\in\mathsf{Dom}(\boldsymbol{A})} -\log_2(P_{\boldsymbol{A}}(a))P_{\boldsymbol{A}}(a) < \infty. \tag{74}$$

Then for each $\ell \in [1:t]$ and $n \in [0:N]$, finite sets

$$\mathcal{U}_0 \subseteq \mathcal{L}_0^t, \ \ldots, \ \mathcal{U}_N \subseteq \mathcal{L}_N^t, \tag{75}$$

can be chosen large enough that clipped random variables:

$$U_n^\ell \triangleq \begin{cases} \tilde{U}_n^\ell & \tilde{U}_n^\ell \in \mathcal{U}_n^\ell \\ \infty & \text{otherwise} \end{cases}, \ n \in [0:N] \tag{76}$$

satisfy (for any $S \subseteq [0:N]$)

$$0 < H(\tilde{U}_S^\ell) - H(U_S^\ell) < \varepsilon', \quad 0 < H(X_{m_\ell}, \tilde{U}_S^\ell) - H(X_{m_\ell}, U_S^\ell) < \varepsilon'. \tag{77}$$

By forming linear combinations of appropriate forms of Equation 77,

$$\|H(U_S^\ell|U_{S^C}^\ell) - H(\tilde{U}_S^\ell|\tilde{U}_{S^C}^\ell)\| < \varepsilon \tag{78}$$

so by Equation 73 for large enough t and $\mathcal{U}_n$'s,

$$\frac{1}{t}H(U_S^\ell|U_{S^C}^\ell) < \sum_{n\in S} r_n. \tag{79}$$

Similarly, finite alphabets $\mathcal{U}_n$ can be chosen large enough that

$$P(U_n^\ell = \infty) < \frac{1}{(N+1)\cdot t^2}. \tag{80}$$

Fix alphabets $\mathcal{U}_n$ and truncated variables U_n satisfying both these properties.

Helper joint-compression and forwarding: At receiver n for $n \in [0:N]$, form random variables

$$V_n^\ell \triangleq \mathsf{Index}_n^*(U_n^\ell), \ \ell \in [1:t], \tag{81}$$

where

$$\mathsf{Index}_n^*(U_n^\ell) = \begin{cases} 1 & U_n^\ell = \infty \\ \mathsf{Index}_n(U_n^\ell) & \text{otherwise} \end{cases} \tag{82}$$

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

Decoding: Take $A_\varepsilon^T(\boldsymbol{X},\boldsymbol{U})$ to be the set of

$$\left((x^1,\ldots,x^t),(\boldsymbol{u}^1,\ldots,\boldsymbol{u}^t)\right) \in \mathcal{X}^t \times \left(\prod_{n=0}^N \mathcal{U}_n\right)^t \tag{83}$$

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

$$\left(\boldsymbol{X},(\boldsymbol{U}^1,\ldots,\boldsymbol{U}^t)\right) \tag{84}$$

where

$$\boldsymbol{U}^\ell = (U_0^\ell,\ldots,U_N^\ell). \tag{85}$$

Weak- and joint-typicality are defined in Reference [18].

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

$$\boldsymbol{v} = (\boldsymbol{v}^1,\ldots,\boldsymbol{v}^t) \in \mathcal{V} \triangleq \left(\prod_{n=0}^N \mathsf{Range}(\mathsf{Index}_n)\right)^t \tag{86}$$

with $\boldsymbol{v}^\ell = (v_0^\ell,\ldots,v_N^\ell) \in \prod_{n=0}^N \mathsf{Range}(\mathsf{Index}_n)$ where $\mathcal{B} \triangleq \{B_{\boldsymbol{v}}|\boldsymbol{v} \in \mathcal{V}\}$ and

$$B_{\boldsymbol{v}} \triangleq \{\boldsymbol{u}|\mathsf{Index}_n(u_n^\ell) = v_n^\ell, \quad \forall n \in [0:N], \ell \in [1:t]\} \subseteq \left(\prod_{n=0}^N \mathcal{U}_n\right)^t. \tag{87}$$

Each $B_{\boldsymbol{v}}$ is the set of helper encodings represented by the compressed messages $\boldsymbol{v}$.

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

Error analysis:

We have the following error events:

- $\mathcal{E}_{\text{Clipped},\ell,n}$: $U_n^\ell = \infty$
- $\mathcal{E}_{JT}$: $\boldsymbol{X}$ is not typical with any $\boldsymbol{u} \in B_{\boldsymbol{V}}$
- $\mathcal{E}_{\hat{m},S}$: For $S \subseteq [0:N]$, and $\hat{m} \neq M$, then there is some $\hat{\boldsymbol{u}} = (\hat{\boldsymbol{u}}_S, \hat{\boldsymbol{u}}_{S^C}) \in B_{\boldsymbol{V}}$ where $\forall n \in S$ $\hat{u}_n = U_n$, $\forall n \notin S$ $\hat{u}_n \neq U_n$, and $(\hat{\boldsymbol{x}}_{\hat{m}}, \hat{\boldsymbol{u}}) \in A_\varepsilon^T(\boldsymbol{X},\boldsymbol{U})$. $\mathcal{E}_{\hat{m},S}$ denotes the situation where the system cannot uniquely decompress encodings from receivers S, causing the broadcast message to look like $\hat{m}$.

Take $E_{\text{Clipped}} = \cup_{\ell=1}^t \cup_{n=0}^N \mathcal{E}_{\text{Clipped},\ell,n}$ By choice of alphabets in Equation (80),

$$P(E_{\text{Clipped}}) \leq \sum_{n,\ell} P(\mathcal{E}_{\text{Clipped},\ell,n}) \tag{88}$$

$$< \frac{Nt}{t^2} \to 0 \text{ as } t \to \infty. \tag{89}$$

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

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

$$\begin{aligned}&\sum_{\hat{m}\neq 1} P(\mathcal{E}_{\hat{m},S} \cap \mathcal{E}_{\text{Clipped}}^C \cap \mathcal{E}_{JT}^C)\\ &\quad\leq 2^{TR} \sum_{\substack{\tilde{\boldsymbol{u}} \in B_{\boldsymbol{V}} \cap A_\varepsilon^T(\boldsymbol{U}):\\ \tilde{\boldsymbol{u}}_{S^C} = \boldsymbol{U}_{S^C}}} P\left((\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{u}}) \in A_\varepsilon^t(\boldsymbol{X}, \boldsymbol{U})\right)\end{aligned} \tag{90}$$

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

$$\begin{aligned}&\sum_{\hat{m}\neq 1} P(\mathcal{E}_{\hat{m},S} \cap \mathcal{E}_{\text{Clipped}}^C \cap \mathcal{E}_{JT}^C)\\ &\leq 2^{TR} \cdot \left|\left\{\substack{\tilde{\boldsymbol{u}} \in B_{\boldsymbol{V}} \cap A_\varepsilon^T(\boldsymbol{U}):\\ \tilde{\boldsymbol{u}}_{S^C} = \boldsymbol{U}_{S^C}}\right\}\right| \cdot 2^{-t\cdot(I(X;\boldsymbol{U})-3\varepsilon)}\\ &\leq 2^{TR}\left(2^{t\cdot(H(\boldsymbol{U}_S|\boldsymbol{U}_{S^C})+\varepsilon)}/2^{t\cdot(t\cdot\sum_{n\in S} r_n - \varepsilon)}\right) 2^{-t\cdot(I(X;\boldsymbol{U})-3\varepsilon)}.\end{aligned} \tag{91}$$

Taking the log,

$$\begin{aligned}&\log\left(\sum_{\hat{m}\neq 1} P(\mathcal{E}_{\hat{m},S} \cap \mathcal{E}_{\text{Clipped}}^C \cap \mathcal{E}_{JT}^C)\right) \leq \ldots\\ &\qquad TR + t\cdot(H(\boldsymbol{U}_S|\boldsymbol{U}_{S^C}) + \varepsilon) - \ldots \qquad (92)\\ &\qquad t\cdot(t\cdot\sum_{n\in S} r_n - \varepsilon) - t\cdot(I(X;\boldsymbol{U}) - 3\varepsilon)\\ &\leq t\cdot(H(\boldsymbol{U}_S|\boldsymbol{U}_{S^C}) - t\cdot\textstyle\sum_{n\in S} r_n + tR - I(X;\boldsymbol{U}) + 5\varepsilon) \qquad (93)\\ &\leq t^2\cdot(I(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon)|\boldsymbol{Z}_{S^C}(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon)) - \ldots\\ &\qquad \textstyle\sum_{n\in S} r_n + R - I(X; \boldsymbol{Z}(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon)) + 7\varepsilon/t). \qquad (94)\end{aligned}$$

Since R was chosen so that $R \leq R_G(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon) - \lambda = I(X; \boldsymbol{Z}(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon)) - \lambda$, then (94) approaches $-\infty$ as $t \to \infty$ (Thereby the left side of Equation (90) approaches 0) if for any $S \subseteq [0 : N]$ then:

$$I(\boldsymbol{Y}_S; \boldsymbol{Z}_S(\boldsymbol{\rho}-\boldsymbol{1}\cdot\varepsilon)|\boldsymbol{Z}_{S^C}(\boldsymbol{\rho}-\boldsymbol{1}\cdot\varepsilon)) < \lambda + \sum_{m\in S} r_m - 7\varepsilon. \tag{95}$$

By assumption that $\boldsymbol{\rho} \in \mathcal{R}_{DC}^\lambda(\boldsymbol{r})$ and for small enough ε, then (95) holds for each S. Since all error events approach 0, then a rate of $R_G(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon) - \lambda$ is achievable. For small enough ε, by lower semi-continuity of mutual information $R_G(\boldsymbol{\rho} - \boldsymbol{1}\cdot\varepsilon) - \lambda$ can be made arbitrarily close to $R_G(\boldsymbol{\rho}) - \lambda$. ∎

Appendix C
Proof of Conditional Capacity

Remark 6. *(Due to Reference [1]) The capacity of the system is $R_{\overline{DC}}(L)$ under the following restrictions:*

- *$\boldsymbol{\Sigma}$ is diagonal (no interference).*
- *The base does not have its own full-precision observation of the broadcast ($h_0 = 0$)*
- *The broadcaster must transmit a Gaussian signal*
- *Helper messages are independent of the transmitter's codebook $\mathcal{X}$.*

Proof: By Theorem 5 in Reference [1], the capacity of the system under the assumed restrictions is:[1]

$$\max_{\substack{\boldsymbol{r} \in \mathcal{R}_{\text{LAN}}(L),\\ \boldsymbol{V} \in \mathcal{V}}} C_{\boldsymbol{r},\boldsymbol{V}} \tag{96}$$

$$C_{\boldsymbol{r},\boldsymbol{V}} \triangleq \min_{S\subseteq[1:N]} \left\{\sum_{m\in S} [r_m - I(V_m; Y_m|X)] + I(\boldsymbol{V}_{S^C}; X)\right\}. \tag{97}$$

In (97), $\mathcal{V}$ is the collection of random vectors $\boldsymbol{V} = (V_1, \ldots, V_N)$ whose components are of the form $V_n = Y_n + \tilde{W}_n$ with independently distributed $\tilde{W}_n$:

$$\tilde{W}_n \sim \mathcal{CN}\left(0, \sigma_n^2 \cdot \frac{2^{-2v_n}}{1 - 2^{-2v_n}}\right) \tag{98}$$

for any $v_n \geq 0$. Since both (98) as a function of v_n and (23) as a function of r_n are injective on $(0, \infty)$, then for any $\boldsymbol{V} \in \mathcal{V}$ there is some $\boldsymbol{\rho}$ (in the context of Equation (34); possibly either inside or outside $\cup_{\lambda\in\mathbf{R}} \mathcal{R}_{DC}^\lambda(L)$) which will yield a variable $\boldsymbol{Z}(\boldsymbol{\rho})$ with identical distribution to $\boldsymbol{V}$.

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

$$\begin{aligned}C_{\boldsymbol{r},\boldsymbol{V}} &= \max\left\{c : \forall S \subseteq [1:N],\right.\\ &\qquad\left. c \leq I(\boldsymbol{V}_{S^C}; X) + \sum_{m\in S} r_m - I(V_m; Y_m|X)\right\} && (99)\\ &\leq \max\left\{c : \forall S \subseteq [1:N],\right.\\ &\qquad\left. c - I(\boldsymbol{V}_{S^C}; X) + I(\boldsymbol{V}_S; \boldsymbol{Y}_S|X) \leq \sum_{m\in S} r_m\right\} && (100)\\ &= \max\left\{c : \forall S \subseteq [1:N],\right.\\ &\qquad\left. [c - I(X; \boldsymbol{V})] + I(\boldsymbol{Y}_S; \boldsymbol{V}_S|\boldsymbol{V}_{S^C}) \leq \sum_{m\in S} r_m\right\} && (101)\\ &= \max_{\lambda\in\mathbb{R}} \max\left\{I(X;\boldsymbol{V}) - \lambda : \forall S \subseteq [1:N]\right.\\ &\qquad\left. I(\boldsymbol{Y}_S; \boldsymbol{V}_S|\boldsymbol{V}_{S^C}) \leq \sum_{m\in S} r_m + \lambda\right\} && (102)\end{aligned}$$

where $I(\boldsymbol{V}_S; \boldsymbol{Y}_S|X) = \sum_{m\in S} I(V_m; Y_m|X)$ follows from the fact that for each n, V_n given (X, Y_n) is conditionally independent of $\boldsymbol{V}_{\{j\neq n\}}$ and $\boldsymbol{Y}_{\{j\neq n\}}$. Maximizing over $\boldsymbol{r}$ this is identical to the set of rates shown in Theorem 3. ∎

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

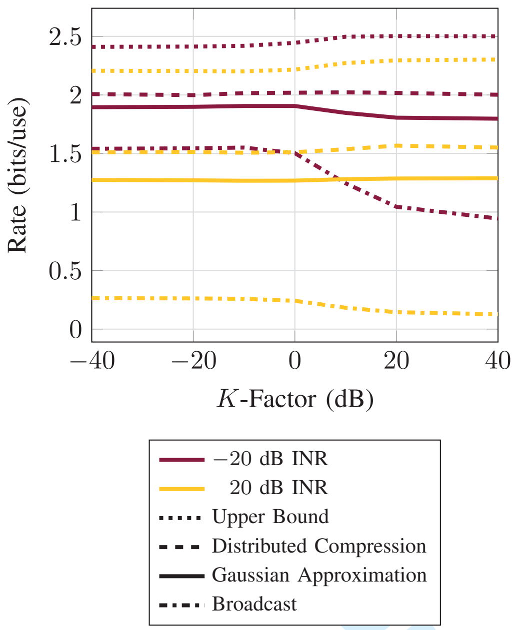


Fig. 6. Achievable rates versus K. 4 helpers with an observation available at the base and LAN constraint $L = 5$ bits shared among helpers. Averaged over 500 channel realizations with uniform independent phase and average receiver SNR fixed at 0 dB. Scattering environment has negligible impact on achievable rates for these bounds, except in the Broadcst bound where receive diversity is not well utilized. Similar results when the base does not have its own observation.

References

[1] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 3008–3023, 2008.

[2] A. Sanderovich, S. Shamai, and Y. Steinberg, "On upper bounds for decentralized MIMO receiver," in *IEEE International Symposium on Information Theory*, pp. 1056–1060, IEEE, 2007.

[3] C. Chapman, A. R. Margetts, and D. W. Bliss, "Bounds on the information capacity of a broadcast channel with quantizing receivers," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, pp. 281–285, IEEE, 2015.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[5] D. W. Bliss, S. Kraut, and A. Agaskar, "Transmit and receive space-time-frequency adaptive processing for cooperative distributed mimo communications," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5221–5224, IEEE, 2012.

[6] F. Quitin, A. Irish, and U. Madhow, "Distributed receive beamforming: a scalable architecture and its proof of concept," in *IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, IEEE, 2013.

[7] E. D. Vagenas, P. Karadimas, and S. A. Kotsopoulos, "Ergodic capacity for the SIMO nakagami-m channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, p. 8, 2009.

[8] D. J. Love, J. Choi, and P. Bidigare, "Receive spatial coding for distributed diversity," in *2013 Asilomar Conference on Signals, Systems and Computers*, pp. 602–606, IEEE, 2013.

[9] J. Choi, D. J. Love, and P. Bidigare, "Coded distributed diversity: A novel distributed reception technique for wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1310–1321.

[10] D. R. B. III, M. Ni, U. Madhow, and P. Bidigare, "Distributed reception with coarsely-quantized observation exchanges," in *47th Annual Conference on Information Sciences and Systems*, pp. 1–6, 2013.

[11] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Transacitons on Information Theory*, vol. 42, no. 3, pp. 887–902, 1996.

[12] S. Tavildar, P. Viswanath, and A. B. Wagner, "The Gaussian many-help-one distributed source coding problem," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 564–581, 2010.

[13] A. Yedla, H. D. Pfister, and K. R. Narayanan, "Code design for the noisy slepian-wolf problem," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2535–2545, 2013.

[14] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[15] I. E. Telatar *et al.*, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[16] H. Weingarten, S. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.

[17] Y. Wei and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, 2004.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[19] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Transactions on Information Theory*, vol. 42, no. 4, pp. 1152–1159, 1996.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[22] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[23] A. Goldsmith, *Wireless Communications*. Cambridge university press, 2005.

[24] D. W. Bliss and S. Govindasamy, *Adaptive Wireless Communications: MIMO Channels and Networks*. Cambridge University Press, 2013.

[25] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001–. [Online; accessed ¡today¿].

[26] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 428–436, 1992.

[27] F. Piera and P. Parada, "On convergence properties of shannon entropy," *Problems of Information Transmission*, vol. 45, no. 2, pp. 75–94, 2009.